ABSTRACT

HUNTSINGER, LETA FAY. Temporal Stability of Trip Generation Models: An Investigation of the Role of Model Type and Life Cycle, Area Type, and Accessibility Variables. (Under the direction of Nagui M. Rouphail.)

Transportation plays a significant role in the mobility, economic health, and quality of life of our communities. The transportation planning process is a complex process of developing and evaluating strategies to meet an area's long-term goals. Planning is inherently a public process that connects us to the future. In the field of transportation planning, travel demand models are often the tool used to make this connection. Travel demand models forecast future travel demand based on forecast input variables related to land use and demographic factors. A fundamental assumption of travel demand models is that model parameters remain stable over time. A violation of this assumption could lead to transportation analyses and travel forecasts that either over- or underestimate travel demand and associated transportation deficiencies, which could in turn lead to poorly allocated investments in transportation infrastructure. Developing a better understanding of the factors that influence travel behavior, the changes in travel behavior over time, and the explanatory variables that best capture these changes may lead to the development of models that are more temporally stable.

This research focuses on advanced statistical data analysis and the development of enhanced trip generation models. The focus of the data analysis is on developing a better understanding of trip making behavior, changes in this behavior over time, and the explanatory variables that contribute to these changes. Model development provides insights into the role that model type and explanatory variables defining life cycle, area type, and accessibility have on temporal stability. Three household travel surveys from Baltimore, Maryland administered in 1977, 1993, and 2001, and two from the Research Triangle region of North Carolina administered in 1995 and 2006 form the basis of the data analysis tasks. The 1995 and 2006 Triangle surveys supplemented with supporting land use and transportation network data form the basis of the model development tasks.

The findings from this analysis show that both the generation choice and cumulative logistic regression models are good models for trip generation. There is also evidence of temporal stability for these two models, perhaps even better than the more widely used cross-classification model based on findings reported in the literature. The introduction of explanatory variables defining life cycle, area type, and accessibility do not noticeably improve model fit, but there is evidence of improved model verification and temporal stability. Finally, analysis shows that trip rates that change over time can have implications for systems and project level planning, resulting in unexpected changes in vehicle miles traveled and associated emissions, transit ridership, traffic forecasts, and localized travel.

This research contributes to the field of travel demand modeling in several ways. It provides important information for model developers and contributes to the conversation on the temporal stability of trip generation models. An underlying assumption of trip generation models is temporal stability, but little information exists on modeling techniques or explanatory variables that improve temporal stability. This research explores three such variables and makes recommendations on their definition and application. This research also explores two modeling techniques for trip generation and makes recommendations regarding application and temporal stability. The findings outlined in this research support the development of advanced trip generation models for medium- and large-sized communities in order to better capture trip making behavior and improve temporal stability. This research also informs the debate on whether it is worth the additional expenditure of time and resources to develop advanced models.

Temporal Stability of Trip Generation Models: An Investigation of the Role of Model Type and Life Cycle, Area Type, and Accessibility Variables

by
Leta Fay Huntsinger

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Civil Engineering

Raleigh, North Carolina

2012

APPROVED BY:

Dr. Nagui Rouphail, Chair	Dr. John Stone
Dr. Billy Williams	Dr. Peter Bloomfield

DEDICATION

This research effort is dedicated to my family and friends. Without their love, support, and faith in me, I could not have achieved this goal.

BIOGRAPHY

Ms. Leta F. Huntsinger was born in Columbia, South Carolina to Joyce Jewell Stokes Huntsinger and Calvin Lee Huntsinger. In 1980, she moved to Wilmington, North Carolina with her mother, stepfather Ravon N. Spencer, and siblings Allan and Shelley. Here she graduated from New Hanover High School where she ran cross-country track, and played basketball. In 1987, she graduated from the University of North Carolina at Wilmington with a bachelor in Physics. Upon graduation, she went to work for the City of Wilmington Traffic Engineering Division. It was through this experience, under the leadership of her first mentor John Bauerlein, that she first developed her passion for transportation and engineering. In 1989, she moved to Raleigh, North Carolina to pursue a Masters Degree in Civil Engineering at North Carolina State University (NCSU) under the direction of Dr. John Stone. While attending school she worked for the North Carolina Department of Transportation (NCDOT), first with the Traffic Engineering Branch and then with the Transportation Planning Branch. It is in that role that Leta discovered an interest and talent for travel demand modeling and a passion for teaching others. Leta completed her Master's degree in 1994 and celebrated with a bicycle tour of the north coast of California. While with the Transportation Planning Branch Leta led the Model Research and Development Team where she was responsible for the oversight of three regional models, the development of best practice guidelines, various travel modeling applications, and model related training. Leta was an adjunct professor at Duke University from 2003 to 2006 where she taught an introductory course in transportation planning and engineering. Leta has also been a co-instructor of CE 701 Urban Transportation Planning at NCSU. In November 2003, Leta left NCDOT to serve as the program manager for the Triangle Regional Model Service Bureau at the Institute for Transportation Research and Education, NCSU. As Program Manager, Leta was responsible for leading and managing the model development activities of a nine-member team. It was in this role that Leta was blessed to be a part of the academic careers of three outstanding individuals: Dr. JinKi Eom, Dr. Yang (Johnny) Han, and Ms. Liza Runey Amar. In order to expand her skills and knowledge in travel modeling she left ITRE in November 2006 to join the Systems Analysis Group at Parsons Brinkerhoff where she had the good fortune to work with some of the most talented (and fun!) travel modelers in the country. Leta entered the doctoral program in Civil Engineering at NCSU in 2009 under the direction of Dr. Nagui Rouphail. During her years in the program, she worked part time with the Durham-Chapel Hill-Carrboro Metropolitan Planning Organization as the team leader for technical services.

Leta is a registered engineer in North Carolina and a certified public manager. She is actively involved in the Transportation Research Board holding committee membership positions with the Planning Applications Committee and Transportation Planning for Small and Medium Communities Committee, serving as the research sub-committee chair for the latter. Her research interests include travel forecasting, multimodal planning and analysis, transportation land use interactions, data methods, process improvement, and transportation planning and traffic operations integration. Leta is an outdoor enthusiast spending much of her time away from work hiking, biking, and kayaking; she is an avid runner and has completed two marathons and numerous half marathons. Leta is also very active in her community, serving previously on the Town of Cary Advisory Board for Parks, Recreation, and Cultural Resources, chairing the Town's Greenway Committee, and volunteering at the North Carolina Museum of Art.

ACKNOWLEDGMENTS

From dream, to concept, to completion, I did not undertake this effort without the influence, support, and guidance of mentors, professors, family, and friends. I have been blessed to have many mentors who have guided me along my journey. The first of these is John Bauerlein, former Traffic Engineer for the City of Wilmington, my first manager and mentor. John made me believe that I could do anything and put into place a support system to make it so. Over the course of my career, I have tried to emulate the things he taught me with those that I have been fortunate to lead and work with. Next along my path, I must acknowledge Dr. M.R. (Ron) Poole, the father of transportation planning in North Carolina. Ron saw in me a passion for learning and encouraged me to pursue my Ph.D. Even after a couple of false starts, Ron was always persistent, but never pushy, in encouraging me not to give up on my goal or my dream. Blake Norwood, former manager of the Transportation Planning Branch taught me courage and the power and payoff of teamwork. Blake fed my curiosity and thirst for knowledge by giving me challenging and interesting projects. I also want to thank Dr. Rick Donnelly for this friendship and support. While my pursuit for higher education did not afford me the opportunity to work with Rick for long, I value tremendously the time we spent together and the things he taught me that go beyond technical knowledge. I especially want to thank him for the time he invested in reading my research proposal, and for the advice he gave that made it stronger. Finally, I want to thank my committee chair, Dr. Nagui Rouphail for his support and guidance in the completion of my dissertation. My decision to finally commit to finishing my PhD is in large part due to his encouragement. Words cannot express my gratitude. All of these individuals have contributed to where I am today through their guidance, insights, accessibility, support, and respect.

Next, I want to acknowledge all the teachers and professors who have sustained my pursuit of knowledge over the years and who have challenged me to go beyond the limits I set for myself. I want to thank my committee members Dr. Peter Bloomfield, Dr. John Stone, and Dr. Billy Williams for their input and guidance on my research and dissertation. In particular, I want to acknowledge Dr. John Stone. Over the years, John has grown from professor to advisor to friend. I have enjoyed the time we have spent working on research projects together, co-instructing CE 701, and discussing life beyond the 8 to 5. Thank you for your friendship and support over the years. I also want to thank Dr. Peter Bloomfield for the assistance he provided on the statistics and model estimation work.

Finally, I want to thank my family and friends for all their encouragement and support, and for reminding me of the things that really matter in life. In particular, I want to thank Dr. Stacey Bricka, Dr. Yang (Johnny) Han, Dr. JinKi Eom, soon to be Dr. Bing Mei, Dr. KoSok Chae, Yanping Zhang, Hong Qi Lu and of course the Monday night Flying Saucer crowd: Dan, Terry, Rhett, Lori, Craig, Pam, Tim, and Nathan, it wouldn’t have been nearly as much fun without you!

TABLE OF CONTENTS

LIST OF TABLES xiv

LIST OF FIGURES xvii

CHAPTER 1. INTRODUCTION 1

1.1 Research Need 1

1.2 Research Objectives 1

1.3 Trip Generation Models – State of the Practice 2

1.4 Motivation 4

1.5 Approach 6

1.5.1 Phase 1: Data Analysis Approach 6

1.5.2 Phase 2: Model Development Approach 7

1.5.3 Phase 3: Comparative Analysis Approach 8

1.5.4 Phase 4: Case Study Approach 8

1.6 Significance of the Research 9

1.7 Research Scope and Limitations 9

1.8 Dissertation Organization 10

CHAPTER 2. DATA DESCRIPTION 11

2.1 Introduction 11

2.2 Baltimore Metropolitan Commission Data 12

2.3 Research Triangle Region Data 14

2.4 Cleaning and Processing Data for the Baltimore Region 17

2.5 Cleaning and Processing Data for the Triangle Region 18

2.5.1 48-hour versus 24-hour Activity Diary 18

2.5.2 Assigned Travel Days - Weekly versus Weekdays Only 19

2.5.3 Household Members Age 5 and Older versus All Household Members 19

2.5.4 Differences in geographic coverage 19

2.5.5 Under-sampling of Durham County 19

CHAPTER 3. URBAN TRENDS AND CHANGES IN TRIP MAKING 21

3.1 Introduction and Motivation 21

3.2 Literature Review 21

3.3 Methodology 23

3.3.1 Data Collection 23

3.3.2 Data Preparation 25

3.3.3 Data Analysis 26

3.4 Results 28

3.4.1 Demographic and Urban Change 28

3.4.2 Trip Rates 30

3.4.3 Comparisons with NHTS 35

3.5 Summary of Findings 36

3.6 Conclusions and Recommendations 41

CHAPTER 4. INFLUENCE OF LIFE CYCLE 42

4.1 Introduction and Motivation 42

4.2 Literature Review 42

4.3 Data Description 45

4.4 Life Cycle Definitions 46

4.5 Methodology 51

4.5.1 Life Cycle and Travel Behavior ... 51
4.5.2 Life Cycle and Temporal Effects ... 52
4.6 Analysis and Results ... 53
4.6.1 Life Cycle and Travel Behavior ... 53
4.6.2 Life Cycle and Temporal Effects ... 59
4.7 Summary and Recommendations ... 62
CHAPTER 5. INFLUENCE OF AREA TYPE AND ACCESSIBILITY ... 65
5.1 Introduction and Motivation ... 65
5.2 Literature Review ... 65
5.2.1 Accessibility ... 66
5.2.2 Area Type ... 67
5.3 Methodology ... 67
5.3.1 Data ... 68
5.3.2 Statistical Tests ... 69
5.4 Accessibility Definitions ... 71
5.5 Area Type Definitions ... 74
5.6 Analysis and Results ... 76
5.6.1 Accessibility, Area Type and Travel Behavior ... 76
5.6.2 Temporal Effects ... 82
5.7 Summary and Conclusions ... 83
CHAPTER 6. GENERATION CHOICE MODELS ... 87
6.1 Introduction and Motivation ... 87
6.2 Literature Review ... 88

6.2.1 Trip Generation Models ... 88
6.2.2 Transferability ... 89
6.3 Methodology ... 90
6.3.1 Data ... 91
6.3.2 List of Variables ... 91
6.3.3 Model Estimation ... 94
6.3.4 Tests of Temporal Stability ... 95
6.4 Analysis and Results ... 98
6.4.1 Model Specification ... 98
6.4.2 Model Estimation Results ... 98
6.4.3 Temporal Stability ... 103
6.5 Summary and Conclusions ... 109
CHAPTER 7. CUMULATIVE LOGISTIC REGRESSION MODELS ... 111
7.1 Introduction and Motivation ... 111
7.2 Literature Review ... 112
7.2.1 Trip Generation Models ... 112
7.2.2 Temporal Stability ... 113
7.3 Methodology ... 114
7.3.1 Data ... 114
7.3.2 List of Variables ... 115
7.3.3 Cumulative Logistic Regression Model ... 118
7.3.4 Tests of Temporal Stability ... 120
7.4 Analysis and Results ... 122

7.4.1 Model Specification 122
7.4.2 Model Estimation Results 122
7.4.3 Model Verification 126
7.4.4 Temporal Stability 127
7.4.5 Backwards Temporal Stability 134
7.5 Summary and Conclusions 138
CHAPTER 8. ASSESSING THE IMPACT OF SAMPLE SIZE, MODEL TYPE, AND EXPLANATORY VARIABLES ON TEMPORAL STABILITY 141
8.1 Introduction and Motivation 141
8.2 Methodology 142
8.2.1 Survey Data and Variable Description 142
8.2.2 Model Estimation and Verification 145
8.2.3 Temporal Stability 147
8.2.4 Experimental Design 149
8.3 Results 150
8.3.1 Generation Choice 150
8.3.2 Cumulative Logistic Regression 156
8.4 Comparative Analysis 160
8.4.1 Model Performance 160
8.4.2 Temporal Stability 161
8.5 Findings and Recommendations 164
8.5.1 Model Performance Findings 164
8.5.2 Temporal Stability Findings 164
8.5.3 Recommendations 165

CHAPTER 9. EVALUATING THE IMPLICATIONS OF TEMPORAL INSTABILITY . 166
9.1 Introduction and Motivation ... 166
9.2 Literature Review ... 167
9.3 Study Methodology ... 168
9.3.1 Case Study 1 - Research Triangle Region, North Carolina ... 169
9.3.2 Case Study 2 - Jackson County, North Carolina ... 170
9.3.4 Trip Rate Development ... 170
9.3.5 Trip Generation Performance Measures ... 172
9.3.6 System Level Performance Measures ... 172
9.3.7 Key Project Performance Measures ... 173
9.4 Analysis and Results ... 173
9.4.1 Research Triangle Region, NC ... 173
9.4.2 Jackson County, NC ... 177
9.4.3 Implications of Temporal Instability over the Longer Term ... 180
9.5 Summary of Findings ... 182
9.6 Conclusions and Recommendations ... 184
CHAPTER 10. SUMMARY, FINDINGS AND RECOMMENDATIONS ... 185
10.1 Summary ... 185
10.1.1 Phase 1: Data Analysis ... 185
10.1.2 Phase 2: Model Development ... 186
10.1.3 Phase 3: Comparative Analysis ... 186
10.1.4 Phase 4: Case Studies ... 187
10.2 Findings and Conclusions ... 187

10.2.1 Phase 1: Data Analysis 187
10.2.2 Phase 2: Model Development 188
10.2.3 Phase 3: Comparative Analysis 188
10.2.4 Phase 4: Case Studies 189
10.3 Recommendations for Professional Practice 190
10.4 Recommendations for Future Research 191
REFERENCES 193
APPENDIX 201
APPENDIX A: EXAMPLE BIOGEME OUTPUT 202
APPENDIX B: EXAMPLE STATA OUTPUT 205

LIST OF TABLES

Table 1. Baltimore Region Trip Statistics 33
Table 2. Triangle Region Trip Statistics 34
Table 3. NHTS Trip Rate Statistics 35
Table 4. Percent Change for NHTS Trip Rate Statistics 36
Table 5. NHTS Life Cycle Definition 47
Table 6. Kermanshah Life Cycle Definition 47
Table 7. Sun Life Cycle Definition 48
Table 8. Marker Life Cycle Definition 49
Table 9. Vadarevu Life Cycle Definition 50
Table 10. Zimmerman Life Cycle Definition 50
Table 11. Household and Person Mean Trip Rates and Sample Size (n) by Life Cycle 53
Table 12. Calculated p-value and eta-square (η^2) for Life Cycle Mean Trip Rates 57
Table 13. NHTS Variable-to-Variable Comparisons using 2006 Household Trip Rate 58
Table 14. Number and Percent of Cells Showing a Significant Difference in Mean Trip Rate for Variable-to-Variable Comparisons by Life Cycle Definition 59
Table 15. Life Cycle Mean Trip Rate Temporal Analysis between 1995 and 2006 61
Table 16. Calculated p-value and eta-square (η^2) for Accessibility Mean Trip Rates 77
Table 17. Calculated p-value and eta-square (η^2) for Area Type Mean Trip Rates 78
Table 18. Final Stratifications for Area Type and Accessibility Definitions 80
Table 19. Household and Person Mean Trip Rates and Sample Size (n) 81
Table 20. Results of Mean Trip Rate Temporal Analysis between 1995 and 2006 83
Table 21. Model Estimation Results for Case 1 Models 101
Table 22. Model Estimation Results for Case 2 Models 102
Table 23. Measures of Temporal Stability for 1995 Case 1 Models Applied to 2006 Socio-economic Data 104
Table 24. Measures of Temporal Stability for 1995 Case 2 Models Applied to 2006 Socio-economic Data 107

Table 25. Model Estimation Results for HBW Case 1 Models Using 1995 Survey 123
Table 26. Model Estimation Results for HBO Case 1 Models Using 1995 Survey 124
Table 27. Model Estimation Results for HBW Case 2 Models Using 1995 Survey 125
Table 28. Model Estimation Results for HBO Case 2 Models Using 1995 Survey 126
Table 29. Model Verification Measures for Models Estimated Using 1995 Survey 127
Table 30. Model Estimation Results for HBW Case 1 Models Using 2006 Survey 128
Table 31. Model Estimation Results for HBO Case 1 Models Using 2006 Survey 129
Table 32. Model Estimation Results for HBW Case 2 Models Using 2006 Survey 130
Table 33. Model Estimation Results for HBO Case 2 Models Using 2006 Survey 131
Table 34. Measures of Temporal Stability for 1995 Case 1 Models Applied to 2006 Socio-economic Data 132
Table 35. Measures of Temporal Stability for 1995 Case 2 Models Applied to 2006 Socio-economic Data 134
Table 36. Model Estimation Results for HBW and HBO Case 3 and Case 4 Models using 2006 Survey 135
Table 37. Model Verification Measures for Models Estimated Using 2006 Survey 136
Table 38. Measures of Temporal Stability for 2006 Models Applied to 1995 Socio-economic Data 137
Table 39. Generation Choice HBW Worker Model Estimation Results 151
Table 40. Generation Choice HBO Worker Model Estimation Results 152
Table 41. Generation Choice HBO Non-Working Adult Model Estimation Results 153
Table 42. Generation Choice HBO Child Model Estimation Results 153
Table 43. Model Verification Measures for Generation Choice Models 154
Table 44. Measures of Temporal Stability for Estimated Generation Choice Models Applied to Socio-economic Data 156
Table 45. Model Estimation Results for HBW Models using 1995 and 2006 Survey 157
Table 46. Model Estimation Results for HBO Models using 1995 and 2006 Survey 158
Table 47. Model Verification Measures for Cumulative Logistic Regression Models 159

Table 48. Measures of Temporal Stability for Estimated Cumulative Logistic Regression Models Applied to Socio-economic Data 160
Table 49. Percent Change in Trip Rates by Trip Purpose 171
Table 50. Typical Ranges for Trip Generation Performance Measures 172
Table 51. Typical Ranges for VMT Performance Measures 173
Table 52. Generalized Annual Average Daily Volumes by Level of Service (LOS) 173
Table 53. TRM Performance Measures 177
Table 54. Jackson County Performance Measures 180
Table 55. Jackson County Performance Measures – 32 Year Horizon 182

LIST OF FIGURES

Figure 1. Research Phases 6
Figure 2. Baltimore Survey Boundaries 14
Figure 3. Triangle Survey Boundaries 16
Figure 4. Percent Change in Trips per Household 37
Figure 5. Percent Change in Trips per Person 38
Figure 6. Percent Change in Regional Census Demographics 39
Figure 7. Triangle Survey Boundaries 69
Figure 8. Experimental Design 150
Figure 9. Percent Change in Trip Rates by Trip Purpose and Rate Set 171
Figure 10. TRM Percent Trips by Trip Purpose by Trip Rate Set 175
Figure 11. TRM Project Demand by TR Set and Level of Service 176
Figure 12. Jackson Percent Trips by Trip Purpose by Trip Rate Set 178
Figure 13. Jackson Project Demand by TR Set and Level of Service 179

CHAPTER 1. INTRODUCTION

Transportation plays a significant role in the mobility, economic health, and quality of life of our communities. It shapes growth by providing access to land and shapes public policy in areas related to "air quality, environmental resource consumption, social equality, land use, urban growth, economic development safety, and security" (FHWA 2007). The transportation planning process is a complex process of "developing strategies for operating, managing, maintaining, and financing an area's transportation system in such a way as to advance the area's long-term goals" (FHWA 2007). Planning is inherently a public process that connects us to the future. In the field of transportation planning, this connection to the future is often made through the development and application of travel demand models that are used to forecast future demand in travel based on forecast input variables related to land use, demographics, and socio-economic factors. Primary among the steps of a travel demand model is that for trip generation, the topic of this research. If the trip generation sub-model is inaccurate, the results of subsequent steps (trip distribution, mode choice and trip assignment) will be wrong.

1.1 Research Need

A peer exchange held in December 2004 to discuss issues of data transferability identified temporal transferability (stability) as a concept that is regularly assumed by modelers, while the validity of this concept has not be sufficiently studied (TMIP 2004). Without a better understanding of temporal stability, it is difficult to defend the use of model parameters developed in one point in time to forecast behavior many years into the future. This research adds to that understanding for trip generation models, and addresses the question of whether survey sample size, model form, and explanatory variables defining life cycle, area type, and accessibility contribute to temporal stability.

1.2 Research Objectives

This research seeks to evaluate the temporal stability of discrete choice models and to identify and evaluate key factors that influence temporal stability. The hypothesis to be tested is that advanced models such as the generation choice and cumulative logistic regression models are temporally stable, and that the temporal stability of these models improves with the inclusion of key explanatory variables such as lifecycle, area type, and accessibility. Lifecycle can be defined in a number of different ways, but is generally used to describe changes that people undergo from infancy, to childhood, to adulthood, to old age as a means of better understanding how these various cycles influence travel behavior (Zimmerman 1981). Area type describes land use and development type with respect to population and employment, while accessibility captures characteristics of the supply side of the transportation system. These are conceptual definitions; this research will explore and recommend definitions that seek to improve the explanatory power and temporal stability of trip generation rates and parameters.

The specific objectives of this research are to:

1. Better understand changes in travel behavior and the factors that influence travel behavior over time, especially with respect to lifecycle, accessibility, and area type;
2. Evaluate the usefulness of a new family of discrete choice models for trip generation;
3. Provide insights into the temporal stability of discrete choice trip generation models;
4. Compare the models estimated with and without lifecycle, accessibility, and area type variables to determine whether these variables improve temporal stability of trip generation rates; and
5. Conduct two case studies to evaluate the implication of trip rates that do not remain stable over time.

1.3 Trip Generation Models – State of the Practice

A travel demand model is a series of mathematical equations used to describe travel and travel choices. In its most basic form, this series of models is broken into a 4-step process, trip generation, trip distribution, mode choice, and trip assignment. Best practice models use locally collected travel survey data to estimate and calibrate the models. The trip generation model provides an estimate of the number of trips generated and attracted to each traffic analysis zone (TAZ) in the study area.

Trip generation models have taken many forms over the years, including zonal regression models, household regression models, and cross-classification models. Early travel forecasts consisted primarily of the extrapolation of “desire lines” developed from origin-destination (OD) surveys (FHWA 1975). This practice advanced in the early 1950s to consider land use and socio-economic factors in quantifying urban trip volumes, providing an analytical approach for using future land use plans to estimate future travel demand (FHWA 1975). Regression models of trip generation became commonplace in the late 1950s and early 1960s opening the door for a greater insight into travel and the factors influencing it (FHWA 1975). Regression models have the advantage of allowing the analyst to consider multiple independent variables, but the disadvantage of treating trip rates as continuous rather than discrete.

The 1970s marked a shift away from aggregate zonal level regression analysis to disaggregate household cross-classification procedures. Cross-classification models estimate an average number of trips as a function of two or more household attributes (Ortuzar and Willumsen 2011). This method has long been the most established model for estimating trips in a travel demand model. Cross-classification models overcome the limitations of regression models, but introduce another shortcoming with respect to the number of variables and stratifications considered before violating the minimum sample size requirements (about 30 samples per stratification), or conversely making the survey sample size prohibitively expensive. Another disadvantage of cross-classification is the lack of goodness of fit measures.

New model forms are becoming more common in the toolbox of models considered for trip generation. These disaggregate models based on discrete choice analysis are considered by some to be a major innovation in the field (Ben-Akiva and Lerman 1985). While commonly used for mode choice modeling, recent applications have also considered destination choice, and even more recently generation choice. Generation choice models estimate the frequency of daily person trips. Models that estimate person trips are an improvement over household based models as they allow for a greater use of important variables and are more compatible with other components of the modeling system (Ortuzar and Willumsen 2011).

In addition to choice-based models, another form of disaggregate model for trip generation is the cumulative logistic regression model. Cumulative logistic regression models, also known as ordered logistic regression, estimate relationships between an ordered categorical dependent variable and a set of independent variables.

Generation choice and cumulative logistic regression models offer several advantages over the commonly used cross-classification model, including the flexibility to consider more independent variables, the ability to include continuous variables in addition to classification variables, and statistical measures for evaluating the significance of the independent variables. Also, unlike the cross-classification model, where sample size quickly limits the number of explanatory variables due to the requirement that any given cell have at least 30 observations, a disaggregate model can capture multiple explanatory variables, making it possible to capture relationships that are not possible with the standard cross-classification approach (PB 2007).

1.4 Motivation

Transportation agencies use travel demand models to make investment decisions about transportation infrastructure improvements. These decisions include widening projects, new roadway construction, toll road investment, and major transit investments such as regional rail. The output from the models informs user benefit calculations, project design decisions, and in the case of toll roads, revenue forecasting. An implicit assumption of a travel demand model is that model parameters, such as trip generation rates, remain stable over time (Ortuzar and Willumsen 2011). A violation of this assumption could lead to transportation analyses and travel forecasts that either over- or underestimate travel demand and associated transportation deficiencies, which could in turn lead to poorly allocated investments in transportation infrastructure.

Trip generation models are the first in the sequence of models used to forecast travel demand. As the first step in the model chain, forecasting errors in this step may compound errors in the remaining steps. An initial goal of these models is to represent observed travel behavior, but that is only part of the challenge. These models must also forecast future travel behavior based on a set of assumed future input demographics (land use and population scenarios) and the travel behavior relationships captured in the model specification. The second goal may be as important as the first since travel demand models play a significant role in the transportation planning process, providing transportation planners, highway designers, transit operators, and decision makers with critical data needed for the development and implementation of transportation plans, projects and policies.

Poor forecasts can be the result of many things including errors in model specification, model calibration, model verification, and input data. Forecasting error can also result from model parameters that change over time, or in other words, when the behavior captured by the base year parameters does not hold true in the future. Trip generation rates (parameters) that change over time could lead to traffic forecasts that result in over designing and building transportation projects, further leading to overspending limited public resources. In the case of revenue generating projects, such as a toll road, forecasts higher than those eventually observed could lead to revenue collections that are lower than anticipated, leading to financial challenges over the life of the project and loss of public confidence.

Predicting the future with certainty is beyond our reach, but we can use data and mathematical tools to improve the forecasting process. Experienced model developers know that well-specified and calibrated models can minimize the impact of errors resulting from model specification, calibration, and verification issues, and that better data forecasting techniques can improve the quality of the input data. Improving our understanding of the factors that influence travel behavior and how these factors change over time can lead to better temporal stability of travel models.

The purpose of this research is to investigate changes in trip rates over time in order to develop a better understanding of these changes and the factors that influence these rates. The question considered in this investigation is whether certain variables related to households and individuals provide greater insight into temporal stability of trip generation models.

1.5 Approach

The research is conducted in four phases 1) data analysis, 2) model development, 3) comparative analysis, and 4) case studies, see Figure 1.

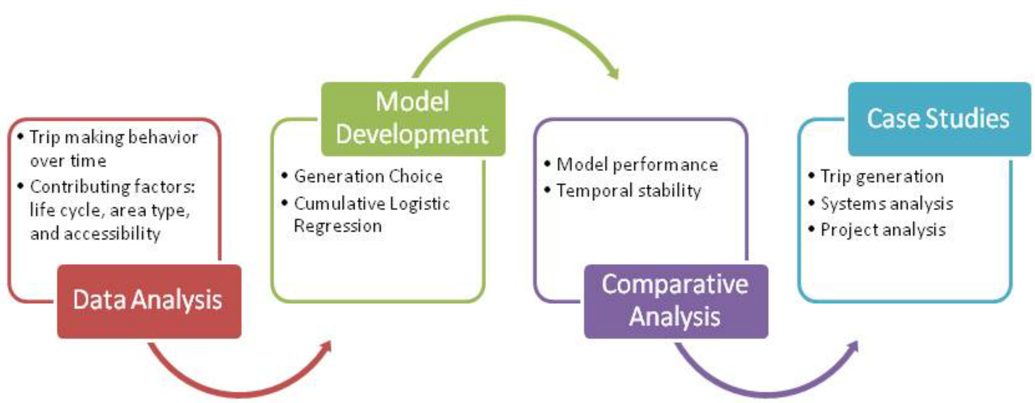


Figure 1. Research Phases

<u>1.5.1 Phase 1: Data Analysis Approach</u>

The data analysis phase is conducted in two parts:

1. A comparison of trip rates over the long and short-term horizon; and

2. An investigation of multiple definitions for life cycle, area type, and accessibility in order to assess the value of each in explaining trip making behavior between strata, trip purpose, and across time.

The comparison of trip rates uses both the Baltimore and Triangle datasets. The examination includes the percent change across time for various demographic and travel statistics as well as the t-statistic to test the null hypothesis of no significant difference between trip rates. The investigation of life cycle, area type, and accessibility is limited to survey data from the Triangle region, as supplemental data needed to calculate the variables for life cycle, area type and accessibility is not available for the Baltimore region. The research explores definitions of life cycle, area type, and accessibility documented in previous research. This included six definitions for life cycle, six different methods for assessing area type, and 24 accessibility definitions. ANOVA is used to evaluate the performance of the various definitions in explaining differences in travel behavior.

1.5.2 Phase 2: Model Development Approach

Using the findings from Phase 1, generation choice and cumulative logistic regression models for home-based work (HBW) and home-based other (HBO) trips are estimated first considering explanatory variables widely used in trip generation models, and second supplementing these variables with the variables from Phase 1 defining life cycle, area type, and accessibility.

The generation choice model is a logit model that estimates the daily trip frequency (n) by trip purpose (y) that a person (x) will make. The cumulative logistic regression model is a type of discrete choice model that estimates relationships between an ordered dependent variable, for example person trip generation, and a set of independent variables, for example household size, income, and workers.

Model performance is assessed through an examination of model verification statistics including person trips per person, person trips per household, HBW trips per worker, and a comparison of estimated to observed total trips for both trip purposes.

The measures of temporal stability include:

1. How well the models estimated in one year predict total trips by trip purpose observed in another year;
2. How well the models estimated in one year predict the fractions of trips by trip purpose and stratification observed in another year;
3. How well the individual model parameters compare between years; and
4. How well the overall contribution of the model parameters compares between years.

1.5.3 Phase 3: Comparative Analysis Approach

The Phase 3 approach is similar to Phase 2, but whereas Phase 2 focused on models estimated using 1995 survey data applied in the 2006 context, Phase 3 expanded the analysis to include models estimated using 2006 survey data applied in the 1995 context. This second round of model development allows for a detailed comparison of the models exploring sample size, model form, and temporal stability.

1.5.4 Phase 4: Case Study Approach

The case study analysis involves a series of model runs using two different case studies. The first case study considers a regional travel demand model and the second a traditional trip based travel demand model for a small urban area. The first model run is the baseline model and reflects the original trip rates for the validated model. Subsequent model runs include one model run for each of three trip rate sets based on the Baltimore and Triangle surveys, and then a fourth trip rate set capturing changes observed between the 1977 and 2001 National Household Travel Survey (NHTS). Trip generation, system level, and project performance measures quantify the effect of the rate changes on the model results.

1.6 Significance of the Research

This research provides important information for travel model developers and contributes to the conversation on the temporal stability of trip generation models. As noted earlier, an underlying assumption of trip generation models is temporal stability, but little information exists on modeling techniques or explanatory variables that support temporal stability. This research explores three such explanatory variables and makes recommendations on their definition and application. This research also explores two modeling techniques for trip generation and makes recommendations regarding application and temporal stability. The findings outlined in this research support the development of advanced trip generation models. In this context, advanced trip generation models refer to discrete choice models where trip generation is the selection of one trip alternative from among a set of mutually exclusive trip alternatives. This research also informs the debate on whether it is worth the additional expenditure of time and resources to develop advanced models.

1.7 Research Scope and Limitations

The scope of this research focuses on data analysis and the development of advanced trip generation models. The focus of the data analysis is on gaining a better understanding of trip making behavior, changes in trip making behavior over time, and the factors that contribute to these changes. Model development provides insights into the role that model type and explanatory variables defining life cycle, area type, and accessibility have on temporal stability. Three household travel surveys from Baltimore, Maryland, administered in 1977, 1993, and 2001, and two from the Research Triangle region of North Carolina administered in 1995 and 2006 form the basis of the data analysis tasks. The 1995 and 2006 Triangle surveys supplemented with supporting land use and transportation network data form the basis of the model development tasks.

This research is limited to the development of advanced trip based models for home-based work and home-based other trip purposes. Model development focuses only on generation choice and cumulative logistic regression models and is limited to two datasets for one geographic region approximately ten years apart.

1.8 Dissertation Organization

Following introductory material and a discussion of the data supporting this research effort, the presentation of this dissertation follows a series of journal articles discussing the major phases of the research. Chapter 3 presents an investigation and findings related to urban trends and changes in trip making over time. Chapters 4 and 5 investigate and present findings on the utility of life cycle (Chapter 4), area type, and accessibility (Chapter 5) for improving our understanding of the factors that influence trip making and how these factors may lead to improved temporal stability. The investigation of generation choice models is presented in Chapter 6 and cumulative logistic regression models in Chapter 7. Both chapters cover model estimation procedures and tests of temporal stability with and without explanatory variables defining life cycle, area type, and accessibility. A comprehensive comparative analysis of model type and temporal stability is presented in Chapter 8, and an evaluation of the implications of trip rates that change over time is presented in Chapter 9. The final chapter, Chapter 10, summarizes the research, findings, and makes recommendations for both professional practice and future research.

CHAPTER 2. DATA DESCRIPTION

The data sets supporting this research include five household travel surveys from five different points in time. Three surveys are from the Baltimore Metropolitan Region, Maryland, and two are from the Research Triangle Region, North Carolina. This chapter describes the various data sets and the cleaning and processing necessary to prepare the data for analysis.

2.1 Introduction

The ideal datasets for this research would be household travel surveys collected at two or more different points in time, covering the same geographic region, using the same survey methodology, and covering a period of at least 20 years. In order to test key factors such as lifecycle, area type, and accessibility and to test the application of the estimated models supporting census data, socio-economic and demographic data, traffic analysis zones, and transportation networks for each period is also required. Unfortunately, several factors contribute to the difficulty in identifying such data, including changes in survey methodology and changes in modeling technology.

The use of Computer-Assisted Telephone Interviewing (CATI) greatly expanded in the 1990s, increasing the efficiency of survey efforts and reducing the number under-reported trips by immediately flagging illogical or erroneous responses (TRB 2010). Most surveys prior to this time required manual coding and key punching tasks. Another methodological shift in the 1990s was the greater use of activity-based surveys where travel diaries report activities, not just trips (TRB 2010). These surveys not only support the development of activity based travel models but also capture more information related to the behavioral information on the choices of "whether, when, and how" people travel (TRB 2010). This last component is critical for the estimation of choice based models.

Changes in modeling tools in the 1990s also impede acquiring prior data that are comparable to data developed in the 1990's; the main change being a shift from command line modeling software to GIS based modeling software. Command line software greatly limited the number of attributes coded to describe the transportation system as well as the true spatial representation of the transportation system. These differences would likely influence the calculation of accessibility variables. Perhaps the biggest impediment to finding these ideal datasets is that few agencies saw the benefit of archiving all the development elements of historic data and model sets.

The previously listed factors present a challenge for the investigation of temporal stability, given that most travel demand models are used to forecast travel 20 to 30 years out. One could argue that changes measured within the first 10 years of model application are far less radical than changes 20 to 30 years out. However, with few datasets spanning 20 to 30 years, the ability to analyze temporal changes was limited in this research. The research overcomes the dataset limitation through the implementation of the two-tiered approach that uses datasets from two different urban areas as discussed below.

The key datasets supporting this research are the 1977 Baltimore Travel Data (Harvey 1980), the 1993 Baltimore Travel Data (Minnesota 2010), the 2001 Household Travel Survey: Baltimore Region Analysis (BMC 2005), the 1994/1995 Triangle Travel Behavior Survey (NuStats 1995), and the 2006 Greater Triangle Travel Survey (NuStats 2006).

2.2 Baltimore Metropolitan Commission Data

The Baltimore region includes Baltimore City, Anne Arundel County, Baltimore County, Carroll County, Harford County, and Howard County (BMC 2011). Travel survey data for the Baltimore region is available as a part of the Metropolitan Travel Survey Archive, an online database housed at the University of Minnesota and funded by the Bureau of Transportation Statistics and the Federal Highway Administration (Minnesota 2010). The intent of this archive is to "store, preserve, and make publicly available, via the internet, travel surveys conducted by metropolitan areas, states and localities" (Minnesota 2010). Documentation is available for 1977 and 2001, but not 1993. An xml codebook is available for key variables in the 1993 survey. Baltimore datasets for 1977, 1993, and 2001 include household, person, and trip data for each year. The availability of raw data from 1977, 1993, and 2001 and documentation from 1977 and 2001 is one of the key reasons survey data for the Baltimore region was selected for this analysis. A prior working relationship with staff at BMC and familiarity with the region was a secondary reason for selecting the BMC data.

The 1977 survey, conducted in May and June, covered only Baltimore, Maryland (Harvey 1980) and included travel data for 967 households. See Figure 2. The survey recorded one day of travel and activities for survey participants (Harvey 1980). Data files include household characteristics, vehicle characteristics, person characteristics, and trip data. The 1993 survey covered a larger geographic region including Baltimore City, Baltimore County, Carroll County, Harford County, Anne Arundel County, and Howard County (Maritz Marketing Research 1993). See Figure 2. The survey was a weekday survey and captured travel during a specified 24-hour period for 2,692 households (Maritz Marketing Research 1993). The survey database includes household records, occupant records, vehicle records, and trip records. The 2001 survey covered the same geographic region as the 1993 survey, Figure 2, but unlike the previous two surveys the 2001 survey included weekday travel along with a small sample of weekend travel (BMC 2005). The survey was a yearlong effort beginning in June 2001, another difference from the 1977 and 1993 surveys that captured only a short snapshot in time (BMC 2005). The 2001 survey database included household records, person records, vehicle records, and weekday trip records for 3,131 households.

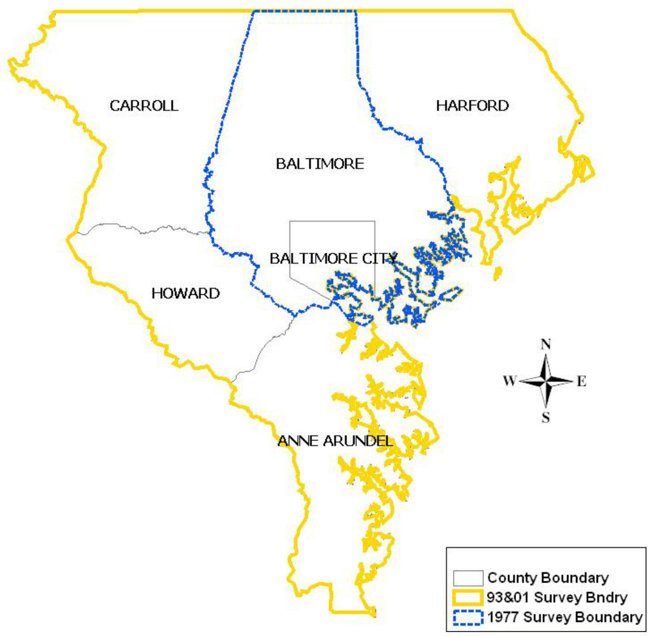


Figure 2. Baltimore Survey Boundaries

The Baltimore region has likely experienced significant growth and urbanization between 1977 and 2001. To understand changes and the possible affect on travel behavior, this research will include not only an analysis of the survey data, but also of available census and demographic data.

2.3 Research Triangle Region Data

The Triangle Region describes the geographic area covered by the Capital Area Metropolitan Planning Organization (CAMPO) and the Durham-Chapel Hill-Carrboro Metropolitan Planning Organization (DCHC MPO). Together the boundaries of the two MPOs cover eight counties; the currently adopted Triangle Regional Travel Demand Model covers nine counties. As an employee of the DCHC MPO the Triangle travel survey data is readily available to me and I have familiarity with both datasets through previous model development work.

The 1995 Triangle Travel Behavior Survey (1995 Survey) covered the period between November 1994 and April 1995 and the North Carolina counties of Wake, Durham, Orange and portions of Harnett, Chatham, Person, Granville, Franklin, and Johnston. See Figure 3. The 1995 Survey was an activity-based survey, favored over a more traditional trip-based survey because it leads to a better understanding of travel as a "derived demand" and places travel within the context of activities that the traveler participates in over the course of the day (NuStats 1995). The sample was a stratified random sample that included 1,778 households. Stratification reflected geographic location defined as urban (greater than 1,920 households per square mile), suburban (133 – 1,920 households per square mile), and rural (zero – 133 households per square mile). The 1995 Survey collected activity and travel data for all household members five and older over a 48-hour period; travel days covered weekday and weekend travel. The 1990 Census STF-3A data file formed the universe for the survey sample. Data files include household characteristics, person characteristics, activities, and vehicles.

The 2006 Greater Triangle Travel Survey (2006 Survey), also an activity-based survey, covered the period between January and June 2006 and the full North Carolina counties of Wake, Durham, Orange, Chatham, Franklin, Granville, Johnston, Lee, Person, and Vance; and portions of Harnett and Nash. See Figure 3. The 2006 Survey collected activity and travel data for all household members over a 24-hour period for a selected travel day between Monday and Friday. The final sample included 5,107 completed household survey records. The sample was a stratified random sample with both geographic and demographic targets. Individual counties formed the boundary for the geographic targets while demographic targets included household size and vehicle ownership. The 2000 Census data for the 12-county region formed the sample universe. Data files include a household file, person file, vehicle file, trip file, and a location file.

Figure 3. Triangle Survey Boundaries

While both surveys used CATI technology, several other methodological differences exist between the 1995 and 2006 datasets as summarized below; Section 2.5 provides a recommended approach for addressing, as necessary, these differences.

1. The 1995 48-hour survey versus the 2006 24-hour survey
2. The 1995 survey was a 7-day survey versus weekdays only for the 2006 survey
3. Household members age 5 and older were sampled in the 2006 survey as compared to all household members in the 1995 survey
4. The 2006 survey covered a larger geographic region
5. The 1995 under-sampling of Durham county due to area type stratification

There are several advantages of using these two datasets to answer the proposed research question regarding temporal stability of trip generation rates. The first noted advantage is that the datasets are easily available and already familiar to me. A second and more important advantage is that both surveys are activity-based surveys of person trips collected by the same survey firm using CATI technology. The activity-based surveys should minimize response differences that might result from different survey instruments or techniques. Recording person trips instead of vehicle trips supports the estimation of person based generation choice models. As previously noted, the surveys cover an eleven-year period, which capture the typical 10-year time horizon when most agencies undertake a major data collection and travel model update exercise. Both surveys include the geographic region of interest and have the added advantage of corresponding with a geographic region that includes secondary supportive data as outlined below.

Secondary data that will support this research include highway networks from 1995 and 2006, socio-economic data from 1995 and 2006, and 1990 and 2000 decennial census data including SF3 and Census Transportation Planning Package (CTPP) data. All data are readily available from the Capital Area Metropolitan Planning Organization (CAMPO), the Durham-Chapel Hill-Carrboro MPO (DCHC MPO), and the Triangle Regional Model Service Bureau (TRMSB). As with the Baltimore region, data analysis includes census and other demographic data to understand factors influencing urban trends and changes in travel behavior.

2.4 Cleaning and Processing Data for the Baltimore Region

Processing of the Baltimore data first involved the creation of match-up between the datasets to create consistent variable names and definitions for variables of interest. Survey documentation available through the Metropolitan Travel Survey Archive and provided by BMC staff facilitated this process. Initial processing of the survey data focused on the trip records and the development of trips by five trip purposes: home-based work (HBW), home-based shopping (HBSH), home-based school (HBSC), home-based other (HBO) and non home-based trips (NHB). The development and application of a consistent definition for each trip purpose helped assure that changes noted in the data analysis are not the result of using different definitions of trip purpose. This process was somewhat more limited than the one applied to the Triangle region (described below) due to limited familiarity with, and details within the data. Generation of a person trip file consisted of appending trip data from the trip record file to the person record file. Data processing included the generation of additional variables such as worker status, household size, income group, vehicle ownership group, household workers, and household children for the purposes of facilitating the reporting of statistics. The aggregation of trip records for each individual within a given household produced the household trip file. The final processing step involved flagging weekend travel records in order to exclude them from the analysis. Reasonableness checks of the data included the creation and review of various summaries of the survey data.

2.5 Cleaning and Processing Data for the Triangle Region

Section 2.3 noted several methodological differences that exist between the 1995 and 2006 datasets; summarized below is the recommended approach for reconciling these differences. The weighting and expansion of both datasets utilized census data by county stratified by household size and vehicles per household. Reasonableness checks using weighted and unweighted data include comparisons against household income, ethnicity, and owner status. Applying this consistent approach to both datasets helps assure that changes noted in the data analysis are not the result of using two different approaches to weighting and expansion.

2.5.1 48-hour versus 24-hour Activity Diary

Ideally, both travel days would be included in the estimation dataset in order to increase the number of trip records. However, the possibility exists that by doubling the representation of the 2-day households there will be an effect on the error term. To maintain as much consistency as possible between the datasets the analysis will be restricted to Day 1 records.

2.5.2 Assigned Travel Days - Weekly versus Weekdays Only

The weighting process will use demographic and geographic characteristics, therefore the removal of Saturday and Sunday records will not affect the estimation dataset.

2.5.3 Household Members Age 5 and Older versus All Household Members

Exclusion of trip records for household members under age 5 from the estimation dataset assures that the datasets match as closely as possible. This will not affect the weighting of the survey records.

2.5.4 Differences in geographic coverage

Both datasets will be re-weighted using a consistent method and geography. New expansion factors will also reflect a consistent method. This process will account for the original differences in geography.

2.5.5 Under-sampling of Durham County

The consistent weighting and expansion methodology for both the 1995 and 2006 datasets results in weighting and expansion factors that adjust for the under-sampling of Durham households in the 1995 survey.

The steps to weight and expand the survey data are:

1. Summarize 1990 and 200 census data
2. Remove weekend travel records from the 1996 dataset
3. Remove travel for household members under age 5
4. Remove records for all geographies outside Wake, Durham, and Orange counties
5. Develop new expansion factors
6. Develop new weights

Additional data processing is required to generate key variables for accessibility, area type, and lifecycle. For consistency, the 1995 Survey will be geo-coded to the 2006 traffic analysis zones. Data processing will include the development of new trip purpose codes to assure consistency between the datasets; this step includes processing unlinked trips into linked trips. An unlinked trip refers to a very short segment trip that is a part of a traveler's daily activity but not traditionally considered a separate trip for the purposes of travel demand modeling. An example of two unlinked trips processed to represent one linked trip would be a traveler stopping for fuel on the way to work, which would be coded as one work trip rather than a work-to-other and then other-to-work trip. Identical linking procedures for both surveys will eliminate any differences related to trip definitions. Finally, overall reasonableness checking will investigate, and repair as necessary, coding errors and inconsistencies.

A final comparison of survey data will include households surveyed, total expanded households, average household size, average vehicles per households, total person trips recorded, average daily household trip rate, and projected travel volume. The final estimation dataset for trip generation will be a record for each individual respondent in each survey with the number of trips by trip purpose and all explanatory variables.

CHAPTER 3. URBAN TRENDS AND CHANGES IN TRIP MAKING

3.1 Introduction and Motivation

A travel demand model is a series of mathematical equations used to describe travel and travel choices. In its most basic form, this series of models is broken into a 4-step process, trip generation, trip distribution, mode choice, and trip assignment. Best practice models use locally collected travel survey data to estimate and calibrate the models. Planners use these models to forecast travel demand 20 – 30 years into the future for the purposes of evaluating transportation strategies and investment. An implicit assumption of a travel demand model is that model parameters remain stable over time (Ortuzar and Willumsen 2011). A violation of this assumption could lead to transportation analyses and forecasts that over- or underestimate travel demand and associated transportation deficiencies, which could in turn lead to poorly allocated investments in transportation infrastructure.

Understanding how travel behavior changes over time can lead to the development of better models that capture this change within the model specification. The purpose of this paper is to investigate changes in trip rates over time in order to develop a better understanding of these changes and the explanatory variables that influence these rates. The question considered in this investigation is whether certain explanatory variables related to households or individuals provide greater insight into temporal stability of trip rates and if those variables can improve temporal stability if necessary.

3.2 Literature Review

An implicit assumption of travel demand models is that model parameters remain stable over time, (Ortuzar and Willumsen 2011) yet an investigation of historical travel data highlights changing trends in travel and travel behavior. The National Personal Travel Survey (NPTS), later called the National Household Travel Survey (NHTS), has tracked the nation's personal travel and travel trends since 1969, providing information related to the number of trips, trip purpose, travel mode, and trip duration (Hu and Young 1999). Between 1977 and 2001 person trips increased from 2.92 person trips per day to 4.09 person trips per day, and trips per household increased from 6.36 to 10.39 (Hu and Reuscher 2004). The biggest change was in person trips per day related to family or personal business. These trips increased from 0.91 in 1977 to 1.79 in 2001, a near doubling of the 1977 rate (Hu and Reuscher 2004). A study on changes in activity participation in the Puget Sound Region also documented an increase in the number of shopping trips taken over time (Yee and Niemeier 2000).

An examination of historical trip rate data clearly shows changes, posing a challenge for travel model developers, as one of the underlying assumptions of these models is that the parameters reflecting the behavior represented by explanatory variables remain stable over time. In this context, an explanatory variable might refer to household size or income, and the parameter the average number of trips for each individual or combination of explanatory variables.

Panel surveys, representing a longitudinal data collection effort, are a great resource for understanding changes in travel behavior over time, especially with respect to how behavior changes over the course of life. A panel survey is one in which surveys are administered to the same sample population at different points in time (Ortuzar and Willumsen 2011). Several studies have documented changes in behavior over time using panel data. One such study investigated changes in vehicle ownership as it relates to various life cycle events such as having a baby, moving, and changing employment (Yamamoto 2008). This study found that changes in life cycle do have an impact on vehicle ownership. In another panel survey administered over a 20-year period in Switzerland, researchers found that auto ownership and transit pass ownership substitute each other over the course of life (Beige and Axhausen 2008). Auto ownership is highest between the ages of 35 and 55, with men having higher auto ownership than women in the same age group (Beige and Axhausen 2008). While not directly evaluating changes in trip rates over time, these studies do document changes in demographic characteristics thought to influence trip making, which could lead us to infer that trip making also changes.

Other studies have focused on the temporal stability of trip production rates estimated from household survey data. In this context, temporal stability is concerned with how models developed during one period of time transfer to a future period. It also considers how trip rates estimated with data from one period compare to rates estimated with data from a future period. The literature indicates a heavy focus on the temporal stability of travel models during the 1970s (see Ashford, 1972 (Ashford and Holloway 1972); Kannel and Heathington, 1973 (Kannel and Heathington 1973); Smith and Cleveland, 1976 (Smith and Cleveland 1976); Yunker, 1976 (Yunker 1976); Doubleday, 1977 (Doubleday 1977)), while current work in this area has focused on spatial transferability (see Wilmot, 1995 (Wilmot 1995); Agyemang-Duah and Hall, 1997 (Agyemang-Duah and Hall 1997); Cotrus, 2005 (Cotrus, Prashker et al. 2005); Mohammadian, 2007 (Mohammadian and Zhang 2007); Everett, 2009 (Everett 2009)). The focus on spatial transferability is in response to both the lack of survey data in many urban areas as well as the increased cost for collecting such data (TMIP 2004; Everett 2009). While much remains for exploration in the area of spatial transferability, revisiting the issue of temporal stability is equally important. There are increasing constraints on funds available for infrastructure improvements (TRB 2006; Cambridge Systematics 2007). At the same time, growth, especially in urban areas, continues to outpace investment (Cambridge Systematics 2007). The combination of these factors places even more importance on improved analysis tools for better decision making.

3.3 Methodology

This section provides an overview of the datasets used for this research, the steps necessary to clean and prepare the data, and finally the analysis approach.

3.3.1 Data Collection

As discussed previously, panel surveys are useful in allowing researchers to observe changes in travel behavior over time, but they also have several disadvantages in comparison to the traditional travel survey, including panel fatigue and panel conditioning (Madre and Yamamoto 2009). This study relies on traditional household surveys taken at different points in time for two different urban areas. While panel surveys offer many advantages for understanding changes in behavior over time, the majority of travel surveys conducted in the United States continue to be non-panel surveys.

The datasets supporting this research include three datasets from the Baltimore Metropolitan Commission (BMC) and two datasets from the Research Triangle Region of North Carolina. Specifically, these data include the 1977 Baltimore Travel Data (Harvey 1980), the 1993 Baltimore Travel Data (Minnesota 2010), the 2001 Household Travel Survey: Baltimore Region Analysis (BMC 2005), the 1994/1995 Triangle Travel Behavior Survey (1995 Survey) (NuStats 1995), and the 2006 Greater Triangle Travel Survey (2006 Survey) (NuStats 2006).

Baltimore Metropolitan Commission Data

The Baltimore region includes Baltimore City, Anne Arundel County, Baltimore County, Carroll County, Harford County, and Howard County (BMC 2011). Travel survey data for the Baltimore region are available as a part of the Metropolitan Travel Survey Archive, an online database housed at the University of Minnesota and funded by the Bureau of Transportation Statistics and the Federal Highway Administration (Minnesota 2010). The intent of this archive is to "store, preserve, and make publicly available, via the internet, travel surveys conducted by metropolitan areas, states and localities" (Minnesota 2010). Baltimore datasets for 1977, 1993, and 2001 include household, person, and trip data for each year.

Triangle Region Data

The Triangle Region describes the geographic area covered by the Capital Area Metropolitan Planning Organization (CAMPO) and the Durham-Chapel Hill-Carrboro Metropolitan Planning Organization (DCHC MPO). Together the boundaries of the two MPOs cover eight counties; the currently adopted Triangle Regional Travel Demand Model covers nine counties.

The 1995 Survey covers the period between November 1994 and April 1995 and the North Carolina counties of Wake, Durham, Orange and portions of Harnett, Chatham, Person, Granville, Franklin, and Johnston. The 1995 Survey was an activity-based survey, an improvement over a more traditional trip-based survey because it leads to a better understanding of travel as a "derived demand" and places travel within the context of activities that the traveler participates in over the course of the day (NuStats 1995). The 2006 Survey, also an activity-based survey, covered the period between January and June 2006 and the full North Carolina counties of Wake, Durham, and Orange.

3.3.2 Data Preparation

Baltimore Metropolitan Commission

Processing the Baltimore data involved the creation of match-up between the datasets to create consistent variable names and definitions for variables of interest. Survey documentation available through the Metropolitan Travel Survey Archive and provided by BMC staff facilitated this process. Initial processing of the survey data focused on the trip records and the development of trips by five trip purposes: home-based work (HBW), home-based shopping (HBSH), home-based school (HBSC), home-based other (HBO) and non home-based trips (NHB). The development and application of a consistent definition for each trip purpose helped assure that changes noted in the data analysis are not the result of using different definitions of trip purpose. Generation of a person trip file consisted of appending trip data from the trip record file to the person record file. Data processing included the generation of additional variables such as worker status, household size, income group, vehicle ownership group, household workers, and household children for the purposes of facilitating the reporting of statistics. The aggregation of trip records for each individual within a given household produced the household trip file. Reasonableness checks of the data included the creation and review of various summaries of the survey data.

Triangle Region

Processing for the Triangle data followed a similar process as for the BMC data. The first step was the creation of a match-up between the two datasets to create consistent variable names and definitions for the variables of interest. Survey documentation for both datasets facilitated this process. Both the 1995 and 2006 Triangle surveys covered a geographic area larger than the three core counties of Wake, Durham, and Orange, but this coverage varied between the two survey years. Coding of the survey records facilitated the removal of records outside the core counties in order to create geographic consistency between the two years. All further references to the Triangle region refer to Wake, Durham, and Orange counties only.

Additional data processing included the development of new trip purpose codes to assure consistency between the datasets; this step included processing unlinked trips into linked trips, where an unlinked trip refers to a very short segment trip that is a part of a traveler's daily activity but not traditionally considered a separate trip for the purposes of travel demand modeling. Identical linking procedures for both surveys helps eliminate any differences related to trip definitions. Finally, overall reasonableness checking of the survey data lead to the identification and repair of coding errors and inconsistencies. Generation of a person trip file consisted of appending trip data from the trip record file to the person record file; as with the BMC surveys, data processing included the generation of additional variables to facilitate the reporting of statistics. The aggregation of trip records for each individual within a given household resulted in a household trip file.

3.3.3 Data Analysis

The analysis of travel behavior over time presented in this paper considers changes in trip rates over a short and long-term horizon. The Baltimore travel data spans a 24-year time horizon and provides an understanding of how key indicators have changed over the longer horizon, providing insights into whether changes in the first 10 years follow a similar trajectory in the subsequent decade, or whether change increases as time goes on. The Triangle data set spans an 11-year period. The completeness of this data set combined with the data processing designed to control for differences in survey methodology and original data processing attempts to eliminate differences that might exist due to the processing methodology differences alone.

For this analysis, the t-statistic was used to test the null hypothesis (H_0) of no significant difference between the means estimated from the three Baltimore surveys and the means estimated from the two Triangle surveys. The assumptions for this test are independent random samples from two populations, normally distributed with equal variances in the two populations:

$$t = \frac{(\bar{y}_1 - \bar{y}_2)}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_1}}} \qquad \text{(Equation 1)}$$

$$s_p = \sqrt{\frac{(n_1-1)s_1^2-(n_2-1)s_2^2}{n_1+n_2-2}} \qquad \text{(Equation 2)}$$

$$df = n_1 + n_2 - 2 \qquad \text{(Equation 3)}$$

where:

t = test statistic

$\bar{y}_1$ = mean of sample, population 1

$\bar{y}_2$ = mean of sample, population 2

n_1 = sample size, population 1

n_2 = sample size, population 2

s_p = pooled standard deviation

s_1^2 = sample variance, population 1

s_2^2 = sample variance, population 2

df = degrees of freedom.

The null hypothesis of no significant difference between means is rejected where the absolute value of the test statistic, t, is greater than or equal to $t_{\frac{\alpha}{2}}$. The test was applied at 95% confidence level with $t_{\frac{\alpha}{2}}$ equal to 1.96.

3.4 Results

3.4.1 Demographic and Urban Change

Trip generation models often include variables for household size, auto ownership, number of workers, and children per household as these variables are highly correlated with the number of household trips. Understanding how these demographic measures have changed over time can offer insight into the potential changes in trip making.

BMC's website describes the Baltimore region as "the nation's 19th largest market, with over 2.5 million people. The market ranks among the top 20 in the country in the number of households, total effective buying income and retail sales" (BMC 2011). The top five occupational groups for the Baltimore region are office/administration, sales, managerial, education/training, and production (BMC 2011). The growth in population and housing for the Baltimore region between 1980 and 2000 is similar to the national average. Population increased 8% between 1980 and 1990 and 7% over the next decade to 2000. During the same period, population in the United States ranged between 10 and 13 percent (Census 2011). The growth in households during the same period was higher at 15% between 1980 and 1990 and 10% between 1990 and 2000. This trend shows up in the average household size that is declining (2.87 in 1980 to 2.62 in 2000) and the percentage of one-person households, a 15% increase between 1990 and 2000. The total number of persons aged 65 or greater is growing, but as a percentage of the overall population, this age group is holding around 12 percent. This age group is also the only age group with growth in both decades as compared to the other age groups that moved up and down between age groups and across years. The breakdown of households by number of vehicles has remained constant for the Baltimore region though the number of zero vehicle households did decrease slightly between 1990 and 2000 with the number of one-vehicle households increasing slightly during that same period.

Three major research universities, North Carolina State University, Duke University, and the University of North Carolina in large part define the Triangle region. Another contributing dynamic for the region is Research Triangle Park, home to leading research and development organizations in pharmaceutical and IT industries. Unlike the Baltimore region where growth seems to have stabilized between 1980 and 2000, high growth better defines the Triangle region. Population grew 32% between 1980 and 1990 and 39% between 1990 and 2000. During this same period, households grew 45% during the first decade and 37% in the second. This is much higher than the national average cited earlier. As with the Baltimore region, the average household size dropped between 1980 and 1990, but the next decade saw an increase in the average household size. The increase in the percentage of households with four or more persons suggests that this change comes primarily from a growth in larger families. Employment grew at a higher rate than both population and households, with a 58% change between 1980 and 1990 and a 44% change between 1990 and 2000. Population increased within all age groups for both decades with the highest growth in the 0 to 4 age group for the first decade (48%) and the 45 to 64 age group in the second decade (67%). The age group for persons 65 or older also saw high increase in growth with a 39% increase in the first decade and 29% in the second, though as a proportion of the total population there was little change. The percentage of households with zero vehicles is much lower than for Baltimore. This difference is not surprising as Baltimore is a much larger city with a comprehensive transit system that includes bus and rail modes. The Triangle region is auto dominant and the suburban development pattern makes it much more difficult to be without a car.

These two regions are not unique in their changing demographics. An analysis of NHTS and Census data shows national trends towards changes in family structure, worker population, and vehicle availability. The family structure in 1960 was primarily households with a father working outside the home, a homemaker mother, and three children (McGuckin and Srinivasan 2004). Over the next 40 years this profile changed to represent one where 67% of the households were not a nuclear family (McGuckin and Srinivasan 2004). In 2000, 28% of households were married couples with no children, 26% were living alone, and 13% of the households were other or unrelated (McGuckin and Srinivasan 2004). The households are also getting smaller with more vehicles and a different worker profile with 61% of women working in 2000 as compared to 38% in 1960 (McGuckin and Srinivasan 2004).

National demographic trends are not the only changes that can affect travel; major national events can also affect travel. One such major event that occurred during the period covered by the BMC and Triangle surveys was 9/11. Research conducted by the Bureau of Transportation Statistics using NHTS data found that following 9/11 there was a reduction in discretionary travel, a lower percentage of trips made by persons under 25, and changes in travel by mode depending on the distance traveled (Harrison 2005). While important to keep these factors in mind, they are likely minimal for survey data evaluated in this study. The 2001 BMC survey was completed prior to September 11, 2001. The 2006 Triangle survey was far enough away from the event that any residual impacts are likely to be minor in nature.

3.4.2 Trip Rates

This section provides a summary of trip rates for the two regions over time, including household trip rates, person trip rates, rates by trip purpose, and rates by various strata. The hypothesis of no significant difference between the trip rates for the various years was tested using the t-statistic at a 95% confidence level.

Table 1 is a summary of trip statistics for the Baltimore region; there is no significant difference in the mean trip rates for the shaded cells, indicating acceptance of H_0 at the 95% confidence level. The analysis shows a significant increase in the household trip rate between 1977 and 1993, and a significant decrease in the trip rate between 1993 and 2001. While lower than the 1993 rate, the household trip rate in 2001 is still higher than the 1977 rate. This trend may reflect fast urban growth in the earlier years followed by slowing growth in the later years. The trip rate per person increased from 2.39 trips per person in 1977 to 4.15 trips per person in 1993. As with the household trip rate, the person trip rate dropped between 1993 and 2001, but was still much higher than the 1977 rate. The analysis shows that the percent change in trip rates per person between 1977 and 1993 was 50% or greater for most stratifications. The smallest increase was shopping trips per person which increased by 31%. Another stratification that showed a smaller change in comparison is the trip rate for the 20 to 44 age group. The null hypothesis of no significant difference between 1977 and 1993 at the 95% confidence interval was rejected for all person trip rate stratifications.

Household trip rates between 1977 and 1993 tell a slightly different story. Trip rates for several stratifications changed less than 10%, including the home-based other (HBO) trip rates, work trip rates by one worker and three or more worker households, and total trip rates of one and two vehicle households. Of these, there was no significant difference in trip rates between 1977 and 1993 for 3 stratifications, HBO trip rates, and trip rates by 1 and 2 vehicle households. The percent change for HBO trip rates was below 10% for all comparison years and showed no significant difference between 1977 and other years.

The comparisons between 1993 and 2001 show less difference between years for person trips over household trips. Of particular interest is household size, vehicles per household, and age group, which changed less than 10%, most changed less than 5%. Five of these showed no significant difference in rates, two-person households, zero and one-vehicle households, the 5 to 19 age group, and the 65 and over age group. For household trip rates, the stratification by household size showed the smallest percent change with all categories less than 10%. The only categories that showed no significant difference in trip rate was HBO trips and the two-person household.

Over the longer term, 1977 to 2001, the percent change in household trip rate was less than 10%. Other stratifications with a percent change less than 10% between 1977 and 2001 include the rates for HBW trips, HBO trips, one and two worker households (work trips) and two vehicle households. In addition to a percent change less than 10%, all of these stratifications also showed no significant difference in trip rates between 1977 and 2001.

Table 2 summarizes trip statistics for the Triangle region based on the 1995 and 2001 survey data. Person trip rates appear to be more stable than household trips for the Triangle data, including the person trip rate for total trips. Stratifications by household size, vehicles per household, and age group show the smallest percent change between survey years as well as showing no significant difference between trip rates for all stratifications except three-person households and age group 45 to 64. Work trips by worker decreased by 12% perhaps reflecting an increase in trip chaining, telecommuting, or flexible work arrangements. The home-based shopping (HBSH) trip rate changed only 1%, a change that showed no statistical difference between the two survey years. Person trip rates showed no statistical difference for all household vehicle stratifications, household trip rates showed no difference for the zero vehicle stratification only. One and two person household size stratifications showed no statistical difference in trip rate, this is not the case for the larger household sizes. All household trip rates by trip purpose changed more than 10% between the two survey years. The highest difference, both at the person level and household level, was for HBO trips.

In comparison, there is no stratification across surveys that show temporal stability between all survey years for either the household trip rates or the person trip rates. There are however, indications of stability for some survey years and stratifications across both regions. There is stability in household trip rates for the two person household stratification for both the 1993 and 2001 Baltimore surveys and both Triangle surveys. For person trips, three stratifications show no statistical difference in trip rates for both the Baltimore (1993-2001) and Triangle surveys; these include one-vehicle households, age group 5 to 19, and age group of persons 65 and older.

Table 1. Baltimore Region Trip Statistics

					Percent Change			t-statistic		
		1977	1993	2001	77-93	93-01	77-01	77-93	93-01	77-01
Household Trips										
Total Trips		7.8	9.3	8.3	19%	-11%	6%	9.70	-38.86	3.29
By Purpose	HBW	1.6	1.9	1.6	20%	-18%	-1%	7.79	-99.38	-0.60
	HBSH	1.1	1.0	1.6	-11%	54%	38%	-3.46	17.01	7.03
	HBSC	0.6	1.1	0.4	68%	-58%	-30%	8.85	-26.10	-8.10
	HBO	2.7	2.7	2.6	-1%	-1%	-3%	-0.43	-1.25	-0.82
By HH Size	1	2.5	4.4	4.1	77%	-7%	64%	10.97	-6.06	9.86
	2	5.2	7.5	7.6	44%	2%	47%	12.29	1.72	12.50
	3	6.9	9.8	10.6	43%	8%	54%	8.34	4.32	13.81
	4+	12.1	14.4	15.3	19%	6%	26%	9.35	4.21	29.79
By Workers (Work)	1	1.4	1.5	1.4	9%	-10%	-2%	3.31	-16.64	-0.88
	2	2.5	2.9	2.6	18%	-12%	4%	5.99	-15.66	1.05
	3+	4.6	4.9	4.0	7%	-18%	-12%	2.98	-14.73	-7.65
By Vehicles	0	4.6	6.3	5.1	35%	-18%	11%	8.48	-10.69	3.92
	1	7.1	7.2	6.2	1%	-15%	-13%	0.51	-13.04	-6.85
	2	9.7	10.2	9.9	5%	-3%	2%	1.56	-3.57	0.55
	3+	15.2	13.1	11.3	-13%	-14%	-26%	-3.04	-8.87	-8.16
Person Trips										
Total Trips		2.4	4.1	4.0	74%	-3%	68%	73.32	-9.71	121.3
By Purpose	HBW	0.5	0.9	0.8	75%	-11%	57%	27.51	-64.26	21.72
	HBSH	0.4	0.4	0.8	31%	68%	120%	9.91	28.09	23.11
	HBSC	0.2	0.5	0.2	143%	-54%	11%	21.22	-33.92	2.78
	HBO	0.8	1.2	1.3	44%	8%	55%	18.62	20.87	22.99
By HH Size	1	2.5	4.4	4.1	77%	-7%	64%	10.97	-6.06	9.86
	2	2.6	4.2	4.2	62%	-1%	61%	22.69	-1.19	20.86
	3	2.3	4.1	3.9	78%	-5%	70%	20.83	-3.48	38.37
	4+	2.4	4.1	3.9	74%	-6%	65%	99.71	-6.44	35.00
By Vehicles	0	1.6	3.3	3.3	103%	1%	105%	25.83	1.18	24.60
	1	2.3	4.1	4.1	74%	0%	74%	43.14	-0.14	183.3
	2	2.7	4.3	4.2	57%	-3%	52%	30.74	-21.43	32.12
	3+	3.2	4.3	3.9	36%	-9%	24%	12.61	-8.88	15.02
By Age Group	5 – 19	2.0	3.4	3.3	72%	-1%	69%	28.11	-1.19	18.76
	20 – 44	3.3	4.4	4.2	33%	-5%	27%	19.63	-5.70	38.52
	45 – 64	2.7	4.5	4.2	65%	-6%	55%	20.94	-22.98	21.66
	65+	1.5	4.2	4.2	183%	-1%	181%	33.84	-0.35	25.55

Accept H_0, no significant difference at the 95% confidence level

Table 2. Triangle Region Trip Statistics

		1995 Survey	2006 Survey	% Change 1995 to 2006	t-statistic
Household Trips					
Total Trips		8.4	9.2	10%	6.15
By Purpose	HBW	2.1	1.9	-12%	-7.45
	HBSH	0.9	1.0	11%	2.95
	HBSC	0.9	0.7	-16%	-4.79
	HBO	2.2	3.1	43%	11.29
By Size	1	4.5	4.5	-1%	-0.59
	2	7.9	7.9	1%	0.52
	3	9.8	11.4	16%	6.46
	4+	14.0	16.2	15%	6.24
By Workers (Work)	1	1.6	1.4	-13%	-7.21
	2	3.1	2.7	-12%	-7.20
	3+	4.5	4.2	-5%	-1.37
By Vehicles	0	4.3	4.5	6%	0.59
	1	5.6	6.1	8%	3.03
	2	9.4	10.2	8%	3.85
	3+	10.6	11.8	12%	3.91
Person Trips					
Total Trips		4.1	4.2	1%	1.05
By Purpose	HBW	1.0	0.8	-19%	-13.24
	HBSH	0.5	0.5	2%	0.66
	HBSC	0.4	0.3	-23%	-9.73
	HBO	1.1	1.4	31%	12.60
By Size	1	4.5	4.5	-1%	-0.60
	2	4.3	4.3	0%	0.17
	3	3.9	4.1	5%	2.88
	4+	3.9	4.0	1%	0.60
By Vehicles	0	3.0	3.0	0%	-0.05
	1	4.3	4.3	0%	0.16
	2	4.1	4.2	1%	0.60
	3+	4.1	4.2	3%	1.56
By Age Group	5 – 19	3.3	3.4	2%	1.13
	20 – 44	4.4	4.4	0%	-0.36
	45 – 64	4.3	4.5	5%	2.73
	65+	4.2	4.3	5%	1.34

Accept H_0, no significant difference at the 95% confidence level

3.4.3 Comparisons with NHTS

As noted previously, the NHTS has tracked the nation's personal travel and travel trends since 1969, making it an excellent barometer for tracking travel changes over time. Numerous studies have investigated these changes; the results are useful as a comparison against the findings from the Baltimore and Triangle data. NHTS data between 1969 and 2009 shows a steady decline in persons per household, 3.16 in 1969 to 2.50 in 2009; and a steady increase in vehicles per household, 1.16 in 1969 and 1.86 in 2009 (2009 NHTS). Workers per household and vehicles per worker have also increased. In 1969, the NPTS data showed 1.21 workers per household and 0.96 vehicles per worker. By 2009, these numbers had increased to 1.34 workers per household and 1.39 vehicles per worker. These demographic shifts have lead to changes in travel statistics. The average daily person trips in 1969 were only 2.02 trips per day. By 2009, the number of trips increased to 3.79 trips per day, see Table 3. Average daily person trips per household went from 6.36 to 9.50 during the same period. Work trip rates per person increased in the 1990s with a high of 0.76 trips per person, but over the longer term from 1977 to 2009 have remained stable. The largest increase in person travel between 1977 and 2009 are trip rates for personal business. The trip rate per person was less than one in 1977, saw a high of 1.97 in 1995, and measured at 1.61 trips per person in 2009. Trips for school and church have remained stable, while social and recreational trips have increased slightly from 0.71 in 1977 to 1.09 trips per person in 2009.

Table 3. NHTS Trip Rate Statistics

Person Trips	1977	1983	1990	1995	2001	2009
Trips/HH	7.69	7.20	8.94	10.49	9.66	9.50
Trips/Person	2.92	2.89	3.76	4.30	4.09	3.79
Rates by Purpose						
To/From Work	0.57	0.59	0.62	0.76	0.65	0.59
Family/Personal Errands	0.91	1.02	1.71	1.97	1.79	1.61
School/Church	0.35	0.34	0.35	0.38	0.40	0.36
Social/Recreational	0.71	0.80	1.01	1.07	1.09	1.04
Other	0.38	0.14	0.06	0.12	0.16	0.18

Source: (Santos, McGuckin et al. 2011)

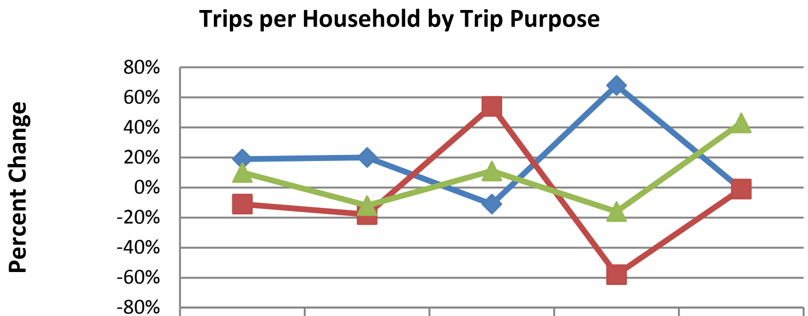
Considering percent change in the NHTS data, Table 4, the biggest change is over the longer term from 1977 to 2009, this period covers the typical 30-year span for transportation planning analysis. With the exception of 1983 to 1990, trip rates are more stable over the shorter term.

Table 4. Percent Change for NHTS Trip Rate Statistics

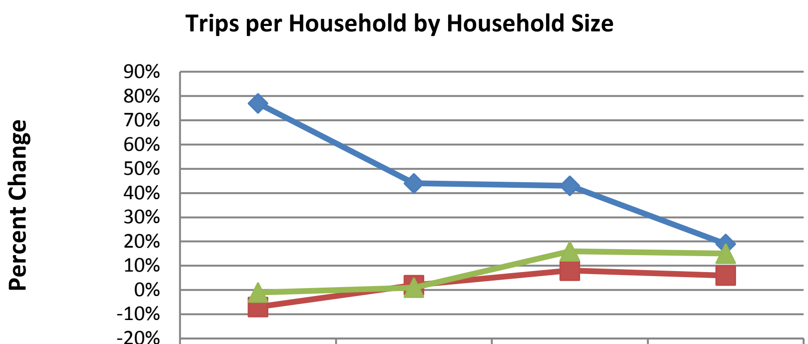
Person Trips	77 to 83	83 to 90	90 to 95	95 to 01	01 to 09	77 to 09
Trips/HH	-0.06	0.24	0.17	-0.08	-0.02	0.24
Trips/Person	-0.01	0.30	0.14	-0.05	-0.07	0.30
Rates by Purpose						
To/From Work	0.04	0.05	0.23	-0.14	-0.09	0.04
Family/Personal Errands	0.12	0.68	0.15	-0.09	-0.10	0.77
School/Church	-0.03	0.03	0.09	0.05	-0.10	0.03
Social/Recreational	0.13	0.26	0.06	0.02	-0.05	0.46
Other	-0.63	-0.57	1.00	0.33	0.13	-0.53

3.5 Summary of Findings

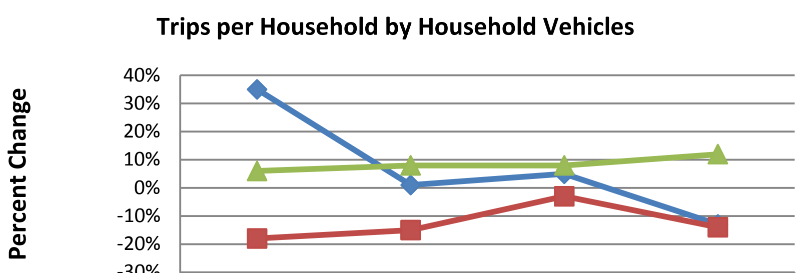
This analysis compared trip rates over a long and short-term horizon to improve understanding of changes and the stability of trip rates over time. Figure 4 shows the relationship of changing household trip rates over time for the survey data evaluated in this study. Figure 5 shows these same relationships for person trip rates. Finally, Figure 6 shows changes in various demographic characteristics using Census data from 1980, 1990, and 2000.



	Total	HBW	HBSH	HBSC	HBO
BMC 77-93	19%	20%	-11%	68%	-1%
BMC 93-01	-11%	-18%	54%	-58%	-1%
TRM 95-06	10%	-12%	11%	-16%	43%

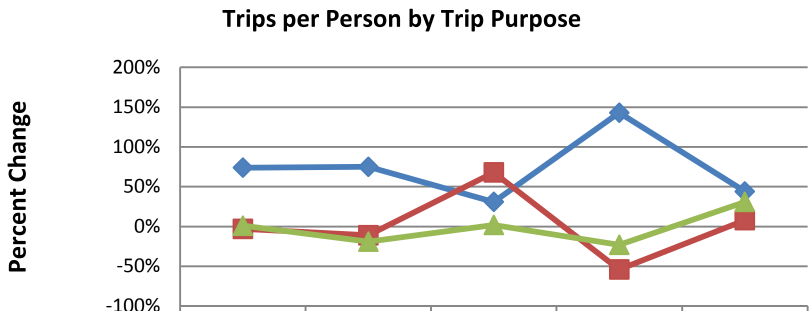


	HH Size 1	HH Size 2	HH Size 3	HH Size 4+
BMC 77-93	77%	44%	43%	19%
BMC 93-01	-7%	2%	8%	6%
TRM 95-06	-1%	1%	16%	15%

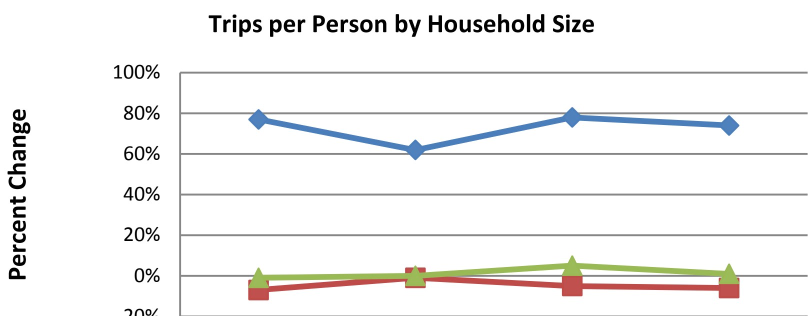


	HH Veh 0	HH Veh 1	HH Veh 2	HH Veh 3+
BMC 77-93	35%	1%	5%	-13%
BMC 93-01	-18%	-15%	-3%	-14%
TRM 95-06	6%	8%	8%	12%

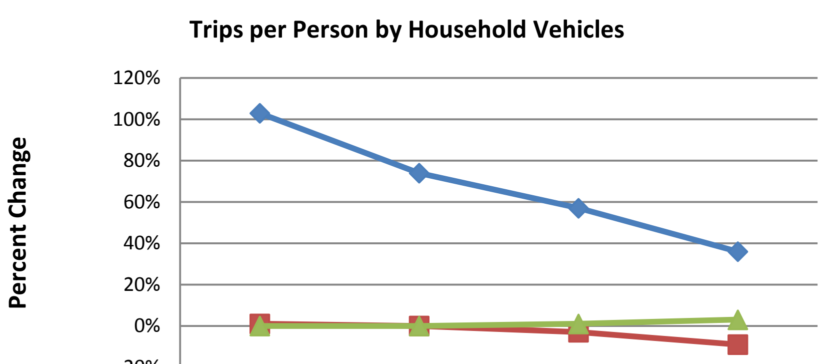
Figure 4. Percent Change in Trips per Household



	Total	HBW	HBSH	HBSC	HBO
BMC 77-93	74%	75%	31%	143%	44%
BMC 93-01	-3%	-11%	68%	-54%	8%
TRM 95-06	1%	-19%	2%	-23%	31%



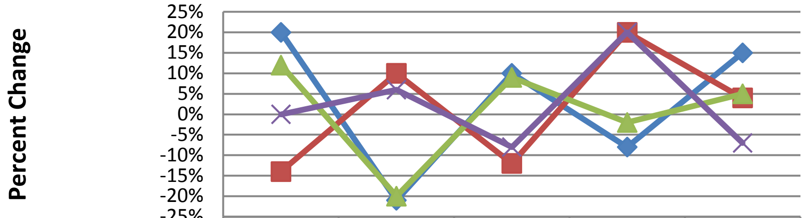
	HH Size 1	HH Size 2	HH Size 3	HH Size 4+
BMC 77-93	77%	62%	78%	74%
BMC 93-01	-7%	-1%	-5%	-6%
TRM 95-06	-1%	0%	5%	1%



	HH Veh 0	HH Veh 1	HH Veh 2	HH Veh 3+
BMC 77-93	103%	74%	57%	36%
BMC 93-01	1%	0%	-3%	-9%
TRM 95-06	0%	0%	1%	3%

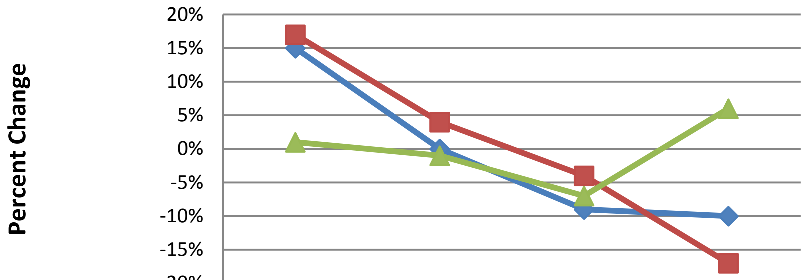
Figure 5. Percent Change in Trips per Person

Population by Age



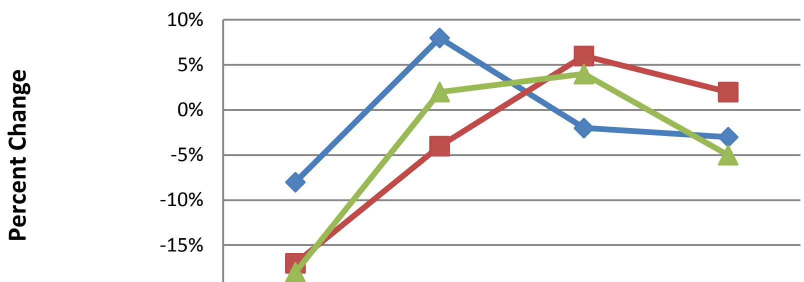
	0 to 4	5 to 19	20 to 44	45 to 64	65 or older
BMC 1980-1990	20%	-21%	10%	-8%	15%
BMC 1990-2000	-14%	10%	-12%	20%	4%
TRM 1980-1990	12%	-20%	9%	-2%	5%
TRM 1990-2000	0%	6%	-8%	20%	-7%

Households by Household Size



	HH Size 1	HH Size 2	HH Size 3	HH Size 4+
BMC 1990-2000	15%	0%	-9%	-10%
TRM 1980-1990	17%	4%	-4%	-17%
TRM 1990-2000	1%	-1%	-7%	6%

Households by Number of Vehicles



	HH Veh 0	HH Veh 1	HH Veh 2	HH Veh 3+
BMC 1990-2000	-8%	8%	-2%	-3%
TRM 1980-1990	-17%	-4%	6%	2%
TRM 1990-2000	-18%	2%	4%	-5%

Figure 6. Percent Change in Regional Census Demographics

Analysis of the Baltimore data from 1977 to 2001 offers insights into how trip rates have changed over the longer horizon, and whether changes in the first 10 years follows a similar trajectory in the subsequent decade. The results show a high increase in household and person trip rates across most strata in the first decade. Household and person trip rates for most strata dropped during the second decade, but were still higher than the rates observed in 1977. This finding may reflect fast growth in the early years followed by slower growth in the later years. Strata where rates increased between 1993 and 2001 include HBSH trips, HBO trips, and household trip rates by household size for two, three, and four plus person households. Applying the hypothesis test of no significant difference between trip rates shows mixed results. There is no difference in household trip rates within the first decade for some strata, but differences exist across all strata for person trip rates. The second decade shows no difference in person trip rates for some strata, but only one stratum for household trip rates.

Findings from the Research Triangle analysis show results that are more favorable in the shorter term, especially for person trip rates. The overall person trip rate shows no statistical difference between the 1995 and 2006 surveys, while the household trip rate increases significantly. The most stable person trip rate by purpose is for shopping trips, while no trip purposes show stability across time for household trip rates. The other finding in the Triangle data is the pattern of stability of the household size strata and the vehicle ownership strata, suggesting that variable selection may improve stability.

A review of NHTS data also shows changes in trip rates over time, with the smallest change occurring over the shorter time spans in all but one comparison, and the greatest change over the longer term. NHTS data shows greatest stability in work trip rates per person and school or church trip rates per person. The biggest change in person trip rates over the long term is trips related to personal business.

Overall, the analysis shows that trip rates change over time, but that there is evidence of stability depending on how the data is stratified, suggesting that the choice of variables may improve temporal stability. Results from the Triangle data suggest that person trip rates may be more stable than household trip rates. There is difficulty in predicting social factors that impact travel behavior over time, for example the impact of women in the work force, increases in auto ownership, the influence of the internet, and other trends that might influence mobility and trip making. There is however, some evidence that certain variables increase stability over time. From this, we could infer that improved models with the right combinations of variables will better capture changes in travel behavior over time.

3.6 Conclusions and Recommendations

It is clear from the analysis documented in this paper that trip rates change over time. A challenge then is to improve our models to capture that change through the selection of better explanatory variables and estimation of more stable trip rates. Results from this paper are encouraging, suggesting that certain stratifications and variables may improve temporal stability. The next step in this effort is to explore a wiser choice of explanatory variables, better describing household composition, worker status, etc. to assess whether temporal stability improves. One recommendation is for additional research that explores how the analysis in this paper transfers to model development, including an investigation of model specification, model form, and key variables. Another useful area to explore is the application of models developed for one period to a future period where observed trips can be compared against the model estimated trips. This research effort should inform survey data collection regarding variables that influence stability, model specification efforts, and an understanding of the sensitivity of models to change.

CHAPTER 4. INFLUENCE OF LIFE CYCLE

4.1 Introduction and Motivation

Travel demand models are valuable tools in the transportation planning process; based on sound theory they bring a quantitative measure to a political process. The forecasts output from these models guide decision makers in the evaluation and selection of transportation programs and projects. Developing a better understanding of the factors that influence travel behavior, the changes in travel behavior over time, and the variables that best capture these changes may lead to the development of models that are more stable over time, increasing our confidence in model results and leading to more cost effective investment decisions.

Previous research identified life cycle as a key variable for predicting travel behavior. "The concept of life cycle is based on the idea that as the family matures, it passes through a series of stages within each of which it behaves in a different manner" (Kermanshah 1997). Given the promising results, one could hypothesize that life cycle may also be a key variable leading to improved stability of trip generation models over time.

The aim of this paper is to evaluate several life cycle definitions identified in the literature in order to assess the value of life cycle for explaining trip making behavior and for improving the temporal stability of trip generation models.

4.2 Literature Review

Models are simplified representations of reality. A travel demand model is a series of mathematical equations used to describe travel and travel choices. A number of factors influence these choices, including characteristics of the household, the trip maker, the trip purpose, and the transportation system. Historically, the common factors considered when estimating the number of trips include income, auto ownership, and household size (FHWA 1975; Ortuzar and Willumsen 2011). Recent studies have also considered number of workers and children (PB 2007). While no model can take into account all of the factors that influence travel behavior, selecting factors that most influence the decision will lead to models that are more robust, increasing confidence in the model results. One such factor, seldom used in trip generation models developed within the context of a traditional 4-step model, is family life cycle.

In a study on transferability of the 2001 NHTS (Hu, Reuscher et al. 2007), life cycle status was identified as a significant factor influencing travel behavior. In a study using 1963 and 1974 household survey data from Rochester, New York, Kitamura and Kostyniuk found that life cycle stage accounts for as much or more variation in travel than did household size, income, number of workers, or number of cars (Kitamura and Kostyniuk 1986). Ortuzar and Willumsen also identify life cycle variables as important factors for explaining trip making behavior (Ortuzar and Willumsen 2011).

Zimmerman, in a FHWA sponsored study using the 1977 NPTS data, defined four life cycles as traditional family, single parent, childless couple, and the single person family (Zimmerman 1981). Given the richness of the NPTS dataset, these four life cycles were further broken down to define various stages that capture age of the head of household and age of children. Key findings of this study related to trip making indicate that, in the traditional family cycle, travel increases as the age of household children increases and then declines for elderly couples in the later years of life. In addition, the single parent family travels considerably less than two-parent families, and there is a general decline in travel in the single person life cycle with each later stage of life. Finally, while social and recreational travel declines with age, trips for family business remain an important component of travel throughout the life cycle.

In another study focusing on the influence that life cycle plays on travel behavior, Kermanshah found that family life cycle is a useful variable in explaining travel behavior, with a significant effect on trips by trip purpose (Kermanshah 1997). The study focus was on disaggregate trip generation models and on independent variables related to “socio-economic characteristics, family structure and life cycle, and land use attributes” (Kermanshah 1997).

Richardson (2006) expanded the concept of life cycle to consider age of the individual combined with the average age and size of the household. He found this alternative measure to be a stronger predictor of trips than the conventional life cycle descriptor alone, concluding that his construct has promise in describing life cycle change, and merits further study (Richardson 2006).

Both life cycle and the built environment were the focus of another study that investigated sustainable travel choice and the role these two variables play in the proportion of trips made by automobile (Sun, Waygood et al. 2009). The built environment was found to be a better predictor of sustainability regarding the fraction of automobile trips, concluding, “households in same the life cycle stage retain the same basic number of trips” (Sun, Waygood et al. 2009). This result suggests stability in the life cycle variable for predicting the number of trips made.

Marker (2000) further investigated life cycle within the context of cohort and period effects using five successive waves of panel data from the Puget Sound Transportation Panel. Marker notes that panel data provides the opportunity to investigate characteristics that might be unique to each cohort, and may indeed account for differences in travel behavior more than life cycle (Marker 2000). While noting that his results are not conclusive, citing the failure to control for other variables that influence activity, Marker calls into question the use of a life cycle variable alone, suggesting that cohort and period effects have a greater influence (Marker 2000). His analysis did not show the expected changes in activity as a household transitioned from one life cycle to another, though this finding merits further study given the short time between survey waves and the small number of transitions that were observed in the data.

4.3 Data Description

The datasets supporting this research come from household surveys administered in the Triangle Region of North Carolina in 1995 and 2006. Specifically, these data include the 1994/1995 Triangle Travel Behavior Survey (NuStats 1995), and the 2006 Greater Triangle Travel Survey (NuStats 2006). The Triangle Region describes the geographic area covered by the Capital Area Metropolitan Planning Organization (CAMPO) and the Durham-Chapel Hill-Carrboro Metropolitan Planning Organization (DCHC MPO).

The 1995 Survey covers the period between November 1994 and April 1995 and the counties of Wake, Durham, Orange and portions of Harnett, Chatham, Person, Granville, Franklin, and Johnston. The sample was a stratified random sample that included 1,778 households. Stratification reflected geographic location, defined as urban, suburban, and rural. The 1995 Survey collected activity and travel data for all household members five and older over a 48-hour period; travel days covered weekday and weekend travel. The 1990 Census formed the universe for the survey sample. Data files include household characteristics, person characteristics, activities, and vehicles.

The 2006 Survey covered the period between January and June 2006 and the counties of Wake, Durham, Orange, Chatham, Franklin, Granville, Johnston, Lee, Person, and Vance; and portions of Harnett and Nash. The 2006 Survey collected activity and travel data for all household members over a 24-hour period for a selected travel day between Monday and Friday. The final sample included 5,107 completed household survey records. The sample was a stratified random sample with both geographic and demographic targets. Individual counties formed the boundary for the geographic targets while demographic targets included household size and vehicle ownership. The 2000 Census data for the 12-county region formed the sample universe. Data files include a household, person, vehicle, trip, and location file.

Both surveys covered a geographic area larger than the three core counties of Wake, Durham, and Orange, but this coverage varied between the two survey years. Removal of the records outside the core counties created geographic consistency between the two survey years. All further references to the Triangle region refer to Wake, Durham, and Orange counties only. Day two records, Saturday and Sunday records, and trip records for household members under age five were removed from the data 1995 survey to improve data set consistency. Although not explored for this research effort, an alternative approach for addressing the missing records for household members under age five in the 2006 survey would be to synthesize the missing records.

The development of new trip purpose codes also assured consistency between the datasets. This included processing unlinked trips into linked trips, where an unlinked trip refers to a very short segment trip that is a part of a traveler's daily activity but not traditionally considered a separate trip for the purposes of travel demand modeling. Identical linking procedures for both surveys helps eliminate any differences related to trip definitions. Finally, overall reasonableness checking of the survey data led to the identification and repair of coding errors and inconsistencies.

4.4 Life Cycle Definitions

This research evaluated six different definitions of life cycle documented in the literature. These definitions were selected for further study as they all reflect slightly different classification schemes focusing on various family characteristics related to both the adult and, if present, child characteristics. The first, and perhaps most common, is the NHTS definition summarized in Table 5 (USDOT 2009). The NHTS identifies ten separate life cycle groups based on the number of adults, presence and age of children, retirement status.

Table 5. NHTS Life Cycle Definition

Code	Variable	Description
1	1ANC	one adult, no children
2	2ANC	2+ adults, no children
3	1AC5	one adult, youngest child 0-5
4	2AC5	2+ adults, youngest child 0-5
5	1AC15	one adult, youngest child 6-15
6	2AC15	2+ adults, youngest child 6-15
7	1AC21	one adult, youngest child 16-21
8	2AC21	2+ adults, youngest child 16-21
9	1ARNC	one adult, retired, no children
10	2ARNC	2+ adults, retired, no children

(Kermanshah 1997) uses only five categories to define life cycle focusing primarily on differences between families with children classified by age as summarized in Table 6.

Table 6. Kermanshah Life Cycle Definition

Code	Variable	Description
1	OC6	families with oldest child under 6
2	OC11	families with oldest child between 6 - 11
3	OC18	families with oldest child between 12 - 18
4	Y18O18	families with at least one child over 18, youngest child under 18
5	OTHR	all other types of families

The third classification considered in Table 7 is the scheme used by Sun that considers not only age of the children, but also age of the household adults and whether the family structure is nuclear family household or single parent household (Sun, Waygood et al. 2009).

Table 7. Sun Life Cycle Definition

Code	Variable	Description
1	SA	single adult younger than 60
2	CC	childless couple, oldest person <60
3	NF6	nuclear family, youngest child <6
4	NF12	nuclear family, youngest child 6 but younger than 12
5	NF17	nuclear family, youngest child 12 or older, but younger than 18
6	F18	nuclear and single parent families, all members of working age
7	OC	older childless couple, oldest person 60+
8	OA	older single adult, 60+
9	SP18	single parent, youngest child <18
10	OTHR	all other households

Marker (2000), Table 8, uses the largest number of classifications capturing smaller age ranges for adults, multiple adult households, female head of household, male head of household, and both younger and older households with no children. Full processing of the survey data for the Triangle region required the addition of a new variable to capture records where age of household adults was unknown.

Table 8. Marker Life Cycle Definition

Code	Variable	Description
1	NC1	multiple adults <35, no children
2	NC2	multiple adults 35-44, no children
3	NC3	multiple adults 45-54, no children
4	NC4	multiple adults 55-64, no children
5	NC5	multiple adults 65+, no children
6	SFM1	single female <45, no children
7	SFM2	single female 45-64, no children
8	SFM3	single female 65+, no children
9	SM1	single male <45, no children
10	SM2	single male 45+, no children
11	SP	single adult any age, at least one child any age
12	CH1	adults <35, all children <6
13	CH2	adults 35-39, all children <6
14	CH3	adults 40+, all children <6
15	MX1	adults <35, children <6 and 6-21
16	MX2	adults 35-39, children <6 and 6-21
17	MX3	adults 40+, children <6 and 6-21
18	SA1	adults <40, children 6-21
19	SA2	adults 40-49, children 6-21
20	SA3	adults 50+, children 6-21
21	OTHR[1]	all other households, age unknown[1]

[1] *added variable*

Table 9 summarizes the classification scheme used by Vadarevu and Stopher (1996); the presence of household worker is a key defining variable for this definition. Full processing of the Triangle survey data required the addition of a new variable to capture all household records.

Table 9. Vadarevu Life Cycle Definition

Code	Variable	Description
1	SA	single working adult
2	ANC	multiple adults, no children, at least one worker
3	YF	young families, at least one worker, children <6
4	OF	older families, at least one worker, at least one child 6 or older
5	NWA	non-working adults, no children
6	NWC[1]	non-working adults, with children[1]

[1] *added variable*

The final classification scheme was proposed by Zimmerman, Table 10, to reflect differences between the traditional family life cycle, single parent life cycle, and childless couple life cycle (Zimmerman 1981). Full processing required the addition of a new variable.

Table 10. Zimmerman Life Cycle Definition

Code	Variable	Description
1	NC30NA	young, not coupled, head of household <30, no children
2	NC55NA	not coupled, head of household 30-55, no children
3	NC69NA	not coupled, head of household 56-69, no children
4	YC30NA	young coupled, head of household <30, no children
5	CAA06	coupled, any age, oldest child 0-6
6	CAA710	coupled, any age, oldest child 7-10
7	CAA1115	coupled, any age, oldest child 11-15
8	CAA1618	coupled, any age, oldest child 16-18
9	CAA1922	coupled, any age, oldest child 19-22
10	C3049NA	coupled, head of household 30-49, no children
11	C5059NA	coupled, head of household 50-59, no children at home
12	C60NA	coupled, head of household 60+, no children at home
13	OP7079NA	not coupled older person, 70-79, no children at home
14	OP80NA	not coupled older person, 80+, no children at home
15	NC06	not coupled, oldest child 0-6
16	NC710	not coupled, oldest child 7-10
17	NC1115	not coupled, oldest child 11-15
18	NC1618	not coupled, oldest child 16-18
19	NC1922	not coupled, oldest child 19-22
20	OTHR[1]	all other unclassified households

[1] *added variable*

4.5 Methodology

The focus of this analysis is on the investigation of different operational definitions of life cycle as summarized in Tables 5 through 10 and an assessment of their usefulness in explaining trip making behavior between life cycle classes, across time horizons and trip purposes.

4.5.1 Life Cycle and Travel Behavior

An analysis of variance (ANOVA) using all life cycle definitions and the strata within each definition provided insight into which life cycle definition and strata best explained travel behavior. ANOVA tests the null hypothesis (H_0) that all means for a given measure are equal. Likewise, the alternative hypothesis (H_a) that at least one mean differs is also tested. If the test results indicate that the null hypothesis should be rejected this does not imply that all mean values are unequal, it is simply an indication that the differences between the means should not be ignored or attributed to chance alone. Using the p-value, we reject H_0 if the p-value is less than 0.05 at a 95% confidence level. ANOVA tests whether the means differ, but does not tell us which means differ. A multiple comparison test evaluates the differences between specific means in order to determine which means differ. For this research the Bonferroni multiple-comparison test was used with the critical value of p less than 0.05; in other words, if the calculated value for p is greater than 0.05 then the difference between any given set of means is not significant (Acock 2006). Finally, eta-square (η^2) quantifies the strength of the difference between the means:

$$\eta^2 = \frac{SSB}{SST} \qquad \text{(Equation 4)}$$

where:

η^2 = eta square, a measure of explained variance (ANOVA equivalent of r^2)

SSB = between group sum of squares

SST = total sum of squares.

4.5.2 Life Cycle and Temporal Effects

Another important characteristic to test is the temporal stability of the various life cycle definitions. The application of trip generation models for future forecasts depends heavily on the assumption that the travel behavior represented by the model remains stable over time. Including variables in the models that improve the temporal stability of the models will lead to improved modeling tools and more defensible forecasts. For this analysis, the t-statistic tests the null hypothesis of no significant difference between the means estimated from the 1995 survey and the 2006 survey. The assumptions for this test are independent random samples from two populations, normally distributed with equal variance in the two populations:

$$t = \frac{(\bar{y}_1 - \bar{y}_2)}{s_p(\sqrt{\frac{1}{n_1} + \frac{1}{n_1}})} \qquad \text{(Equation 5)}$$

$$s_p = \sqrt{\frac{(n_1 - 1)s_1^2 - (n_2 - 1)s_2^2}{n_1 + n_2 - 2}} \qquad \text{(Equation 6)}$$

$$df = n_1 + n_2 - 2 \qquad \text{(Equation 7)}$$

where:

t = test statistic

$\bar{y}_1$ = sample mean, population 1

$\bar{y}_2$ = sample mean, population 2

n_1 = sample size, population 1

n_2 = sample size, population 2

s_p = pooled sample standard deviation

s_1^2 = sample variance, population 1

s_2^2 = sample variance, population 2

df = degrees of freedom.

The null hypothesis of no significant difference between means is rejected when the absolute value of the test statistic, t, is greater than or equal to $t_{\frac{\alpha}{2}}$, df. The test was applied at 95% confidence level with, $t_{\frac{\alpha}{2}}$,df equal to 1.96 where degrees of freedom (df) are greater than or equal to the sum of n_1 and n_2 minus 2.

4.6 Analysis and Results

4.6.1 Life Cycle and Travel Behavior

Table 11 summarizes the 1995 and 2006 calculated mean trip rate and sample size for all life cycle definitions.

Table 11. Household and Person Mean Trip Rates and Sample Size (n) by Life Cycle

		Mean Trip Rates and Sample Size (n)							
Life Cycle Variables		Household Trips				Person Trips			
		1995		2006		1995		2006	
		Mean	n	Mean	n	Mean	n	Mean	n
NHTS	1AC15	9.14	52	9.79	86	3.68	129	3.99	211
	1AC21	9.13	8	8.94	32	4.56	16	4.61	62
	1AC5	5.89	9	11.19	27	3.79	14	3.51	86
	1ANC	4.68	326	4.58	705	4.68	326	4.58	705
	1ARNC	3.67	58	4.08	185	3.67	58	4.08	185
	2AC15	14.53	186	15.23	446	4.00	675	4.14	1643
	2AC21	10.08	39	13.13	112	3.36	117	4.24	347
	2AC5	10.18	165	14.23	449	4.01	419	3.92	1629
	2ANC	8.38	424	8.39	1120	4.26	832	4.25	2220
	2ARNC	7.54	85	7.83	282	4.37	149	4.30	505
Sun	CC	8.18	304	8.05	98	4.17	595	4.13	191
	F18	9.17	18	10.64	55	3.37	49	4.08	146
	NF12	14.87	135	15.84	289	4.08	491	4.12	1112
	NF17	12.88	65	14.36	197	3.74	223	4.25	667
	NF6	10.14	160	14.29	432	4.01	407	3.93	1570
	OA	3.75	79	4.32	331	3.75	79	4.29	331
	OC	7.46	117	8.37	1108	4.18	209	4.33	2110
	OTHR	9.58	106	8.36	269	4.50	226	3.94	601
	SA	4.73	303	4.58	536	4.73	303	4.58	536
	SP18	8.82	65	10.12	129	3.75	153	4.00	329

TABLE 11 (continued)

Life Cycle Variables		Household Trips 1995 Mean	Household Trips 1995 n	Household Trips 2006 Mean	Household Trips 2006 n	Person Trips 1995 Mean	Person Trips 1995 n	Person Trips 2006 Mean	Person Trips 2006 n
Marker	CH1	7.02	44	11.27	122	3.68	84	3.80	362
Marker	CH2	8.84	37	12.39	88	4.36	75	3.68	296
Marker	CH3	8.89	19	13.43	79	4.45	38	4.00	263
Marker	MX1	11.78	23	16.87	31	3.87	70	4.02	130
Marker	MX2	11.80	25	17.36	58	3.73	79	4.13	244
Marker	MX3	14.00	25	16.99	94	4.19	86	3.95	406
Marker	NC1	8.36	137	8.24	187	4.29	267	4.17	369
Marker	NC2	8.65	71	8.36	109	4.45	138	4.34	210
Marker	NC3	9.28	97	8.01	214	4.55	198	4.14	414
Marker	NC4	7.23	97	8.33	418	3.77	186	4.28	814
Marker	NC5	7.65	104	8.29	436	4.32	182	4.34	802
Marker	OTHR	8.00	9	7.64	72	4.10	20	4.08	166
Marker	SA1	11.80	65	11.31	58	3.60	213	3.62	181
Marker	SA2	13.49	163	14.71	329	4.04	545	4.16	1162
Marker	SA3	12.28	53	13.55	284	3.77	170	4.24	909
Marker	SFM1	4.42	106	4.40	146	4.42	106	4.40	146
Marker	SFM2	4.37	62	4.53	240	4.37	62	4.53	240
Marker	SFM3	3.60	45	4.13	180	3.60	45	4.13	180
Marker	SM1	5.01	128	4.50	112	5.01	128	4.50	112
Marker	SM2	4.59	41	4.75	187	4.59	41	4.75	187
Kermanshah	OC11	12.57	148	15.27	290	3.97	469	4.14	1070
Kermanshah	OC18	13.07	175	14.31	506	3.89	589	4.13	1754
Kermanshah	OC6	8.08	99	12.17	291	4.10	195	3.81	931
Kermanshah	OTHR	6.68	910	6.86	2325	4.31	1411	4.31	3702
Kermanshah	Y18O18	14.00	20	15.84	32	3.94	71	3.73	136
Vadarevu	ANC	8.39	434	8.44	1186	4.29	849	4.27	2341
Vadarevu	NWA	5.44	162	5.56	467	4.01	220	4.09	635
Vadarevu	NWC	8.71	7	9.95	21	4.36	14	3.87	54
Vadarevu	OF	12.85	351	14.66	844	3.90	1158	4.12	3004
Vadarevu	SA	4.73	299	4.66	639	4.73	299	4.66	639
Vadarevu	YF	8.08	99	12.21	287	4.10	195	3.81	920

Mean Trip Rates and Sample Size (n)

TABLE 11 (continued)

Life Cycle Variables		Mean Trip Rates and Sample Size (n)							
		Household Trips				Person Trips			
		1995		2006		1995		2006	
		Mean	n	Mean	n	Mean	n	Mean	n
Zimmerman	C3049NA	8.43	145	8.45	247	4.28	285	4.31	476
	C5059NA	8.25	95	8.28	347	4.15	189	4.29	667
	C60NA	7.46	117	8.39	520	4.18	209	4.37	978
	CAA06	8.68	114	12.63	319	4.04	245	3.84	1046
	CAA1115	15.07	104	15.63	259	4.21	369	4.14	979
	CAA1618	14.31	48	15.48	181	3.79	182	4.16	672
	CAA1922	11.96	27	14.64	45	3.44	94	3.95	167
	CAA710	12.70	81	16.22	154	3.82	272	4.19	597
	NC06	8.00	11	11.04	26	4.19	21	3.96	72
	NC1115	8.94	33	10.29	69	3.88	76	3.88	184
	NC1618	9.63	16	10.70	43	3.02	51	4.42	104
	NC1922	11.22	9	10.75	20	4.81	21	3.84	56
	NC30NA	6.42	137	5.72	123	5.15	171	4.40	160
	NC55NA	5.33	236	5.08	426	4.55	276	4.36	495
	NC69NA	4.33	45	4.74	285	3.98	49	4.33	311
	NC710	9.23	13	10.59	27	3.64	33	3.84	76
	OP7079NA	4.03	37	4.66	126	3.73	40	4.29	137
	OP80NA	3.91	11	3.97	63	3.46	11	3.67	70
	OTHR	8.00	9	7.64	72	4.10	20	4.08	166
	YC30NA	7.48	64	8.10	92	3.96	121	4.14	180

As indicated in Table 12, the mean value of total household trips varies significantly across life cycle for all definitions of life cycle. The η^2 value is higher for the 2006 data with several life cycle definitions explaining 40% of the variance in household trips. For household trips, life cycle best explains the variability in trip rates for school trips (HBSC) and work trips (HBW). Life cycle explains less than 10% of the variability in shopping trips (HBSH). Using the life cycle definitions, the 2006 data best explains the variability of trips classified as "other" (HBO) with ranges between 19 and 25%.

While the mean value of total person trips varies significantly across life cycle, this variable explains only 3% or less of the variability in person trips. The life cycle definitions proposed by Marker and Zimmerman explain 21% of the variation in person school trips for the 2006 data, while the NHTS definition better explains the variability in person work trips. Clearly, life cycle explains household trips better than person trips, which makes sense because life cycle is a household construct. However, a review of the individual trip purposes for person trips provides evidence that life cycle is still a good predictor of person trip making. From this, the analyst can infer that understanding life cycle at the household level can help describe the trip dynamics at an individual level, supporting intuition that family composition influences activities, and therefore trip making at the individual level.

Based on this comparison, the Kermanshah life cycle definition is the poorest performing in the group. There is no significant difference for shopping trips in the 1995 household level analysis and HBO in the 1995 person level analysis. Additionally, this definition explains very little of the variation in trip rates for most trip purposes. In comparison to the other life cycle definitions, Kermanshah has only four distinct classifications that focus only on families with children, grouped by age of the children. Likewise, the Vadarevu life cycle definition has six distinct classifications with the primary distinction in the classifications being working and non-working adults. This proves to be a good stratification for explaining the variability in work trips, but not as good for explaining variability for other trip purposes. In contrast, the number of classifications for the other definitions is much higher, 10 for both NHTS and Sun, and 20 for Zimmerman and Marker. All four of these life cycle definitions explain a high degree of variability in total trips rates, and trip rates by trip purpose.

Table 12. Calculated p-value and eta-square (η^2) for Life Cycle Mean Trip Rates

	NHTS		*KSHAH*		*SUN*		*MARKR*		*VADAU*		*ZIMRMN*	
	p	η^2	p	η^2	p	η^2	p	η^2	p	η^2	p	η^2
1995 Household Trips												
Total		0.32	<.001	0.23		0.32		0.32		0.30		0.31
HBW		0.29		0.05		0.25		0.22		0.28		0.25
HBSH	<.001	0.08	0.16	0.00	<.001	0.08	<.001	0.07	<.001	0.05	<.001	0.09
HBSC		0.41	<.001	0.45		0.42		0.45		0.43		0.47
HBO		0.15		0.11		0.14		0.15		0.14		0.15
2006 Household Trips												
Total		0.40		0.32		0.39		0.40		0.38		0.40
HBW		0.27		0.06		0.16		0.21		0.28		0.22
HBSH	<.001	0.05	<.001	0.01	<.001	0.04	<.001	0.05	<.001	0.04	<.001	0.04
HBSC		0.49		0.55		0.48		0.54		0.53		0.55
HBO		0.25		0.19		0.23		0.25		0.2		0.24
1995 Person												
Total		0.02		0.01		0.02		0.02		0.01		0.03
HBW		0.13	<.001	0.03		0.09		0.09		0.14		0.1
HBSH	<.001	0.07		0.02	<.001	0.06	<.001	0.05	<.001	0.05	<.001	0.06
HBSC		0.14		0.15		0.15		0.16		0.14		0.18
HBO		0.04	0.88	0		0.02		0.02		0.03		0.03
2006 Person												
Total		0.01		0.01		0.01		0.01		0.01		0.01
HBW		0.14		0.03		0.05		0.09		0.13		0.09
HBSH	<.001	0.04	<.001	0.01	<.001	0.02	<.001	0.03	<.001	0.03	<.001	0.03
HBSC		0.18		0.2		0.18		0.21		0.19		0.21
HBO		0.04		0.01		0.02		0.04		0.02		0.03

Table 13 provides an example of the variable-to-variable comparisons conducted for each life cycle definition using the Bonferroni multiple-comparison test described previously. A trip purpose code identifies cells with a p-value less than 0.05 indicating a significant difference in the mean trip rate for the represented trip purpose. For example, a cell with "TWC" indicates a significant difference for total, work, and school trips between the two variables classified by that cell. As expected, reviewing this analysis across all life cycle definitions shows patterns of differences in school trips for households with school age children, and between household work trips for workers versus non-workers. The number and age of children and the number of adults is a better predictor of differences for HBO trips. A surprising finding was significant differences in shopping trips for life cycle categories defined by age greater than 65. These trip rates were significantly higher than other life cycle categories. Single adults and families also influence the number of shopping trips. There are significant differences in total, work, school, and other trips between single adults with and without children, but no differences between single adults classified by gender. Finally, while there are differences between life cycle strata by age greater than 65, there are no significant differences between age groups within that classification, for example between 60-69, 70-79, and age 80 or greater.

Table 13. NHTS Variable-to-Variable Comparisons using 2006 Household Trip Rate

	1AC15	**1AC21**	**1AC5**	**1ANC**	**1ARNC**	**2AC15**	**2AC21**	**2AC5**	**2ANC**
1AC21	C								
1AC5	O	C							
1ANC	TCO	TCO	TCO						
1ARNC	TWC	TWC	TCO	TWO					
2AC15	TWCO	TWCO	TWC	TWHCO	TWCO				
2AC21	TWC	TWC	TW	TWHCO	TWCO	TWCO			
2AC5	TWHCO	TO	TWC	TWHCO	TWHCO	CO	WCO		
2ANC	WC	C	WCO	TWHO	TW	THCO	TWCO	THCO	
2ARNC	TWHC	WC	TC	TWHO	THO	TWCO	TWC	TWCO	WHO

T=total H=shop O=other
W=work C=school

Another way to investigate the value of using life cycle to describe variations in trip making is to summarize the variable-to-variable comparisons by percent and number of cells that show a significant difference between the mean trip rate by trip purpose. Table 14 shows this for household and person trips using both the 1995 and 2006 survey data. It is best to focus the comparisons on the pairs of definitions that have a similar number of cells. Between NHTS and Sun, differences between total, work, and school trips are best explained using the Sun definition of life cycle. The NHTS definition best captures the differences between shopping trips, while both perform similarly for other trips. Between Marker and Zimmerman, the Marker definition better explains differences between strata for total trips, work trips at the person level, shopping trips, and other trips. The Zimmerman definition better explains differences for work trips at the household level and school trips.

Table 14. Number and Percent of Cells Showing a Significant Difference in Mean Trip Rate for Variable-to-Variable Comparisons by Life Cycle Definition

Life Cycle (# of cells)	Total		Work		Shop		School		Other	
	%	#	%	#	%	#	%	#	%	#
1995 Household Trips										
NHTS (45)	47%	21	62%	28	29%	13	53%	24	27%	12
SUN (45)	67%	30	73%	33	24%	11	69%	31	42%	19
MARKR (190)	48%	92	37%	70	8%	16	47%	90	28%	53
ZIMR (190)	36%	68	39%	74	8%	15	49%	93	16%	30
2006 Household Trips										
NHTS (45)	78%	35	69%	31	27%	12	80%	36	64%	29
SUN (45)	80%	36	76%	34	27%	12	82%	37	69%	31
MARKR (190)	77%	147	45%	86	22%	42	49%	93	59%	112
ZIMR (190)	63%	120	63%	119	12%	23	67%	127	43%	82
1995 Person Trips										
NHTS (45)	13%	6	53%	24	38%	17	56%	25	27%	12
SUN (45)	18%	8	64%	29	36%	16	64%	29	27%	12
MARKR (190)	4%	7	29%	55	9%	18	41%	77	3%	6
ZIMR (190)	7%	14	26%	50	8%	16	40%	76	6%	12
2006 Person Trips										
NHTS (45)	16%	7	67%	30	44%	20	73%	33	42%	19
SUN (45)	13%	6	78%	35	36%	16	69%	31	42%	19
MARKR (190)	6%	11	55%	105	27%	51	56%	106	38%	73
ZIMR (190)	2%	3	49%	93	23%	43	59%	112	28%	53

4.6.2 Life Cycle and Temporal Effects

The goal of the temporal analysis is to acquire a greater understanding of how life cycle influences trip making over time, in order to inform the hypothesis that the inclusion of these variables in trip generation models will contribute to the temporal stability of travel forecasts. A t-statistic value less than 1.96 (corresponding to a 5% significance level) suggests that there is not enough evidence to reject the null hypothesis, and we can infer temporal stability between the means. Table 15 provides a summary of results; shaded cells indicate H_0 cannot be rejected thus providing evidence of temporal stability. In comparison to the total household and person trip rates, stratification by life cycle improves the temporal stability of the calculated rates. Considering the number of stratifications that are temporally stable, total person trips show greater stability than household trips. In comparison, trip rates using the Sun definition of life cycle are less stable than the other definitions with no more than 50% of the strata remaining stable for household or person trips for any given trip purpose. Only 20% strata are stable for total household trips, household work trips, and person work trips. The Zimmerman definition of life cycle shows the greatest stability as measured by looking at the highest percentage of strata with stability for the majority of the trip purpose categories for both household and person trips. Temporal stability for all life cycle definitions is best for shopping trips while total household trips are less stable for households with children.

Table 15. Life Cycle Mean Trip Rate Temporal Analysis between 1995 and 2006

		Calculated t-statistic									
Life Cycle Variables		Household Trips					Person Trips				
		Total	HBW	HBSH	HBSC	HBO	Total	HBW	HBSH	HBSC	HBO
NHTS	1AC15	1.79	-1.09	1.40	-2.44	1.15	2.16	-0.77	1.55	-2.31	0.63
	1AC21[1]	-0.15	-0.33	1.57	-3.33	0.23	0.09	-0.17	2.22	-2.96	0.35
	1AC5[1]	3.16	-1.30	1.54	1.09	2.28	-0.53	-3.88	1.60	-0.83	0.89
	1ANC	-1.24	-5.88	3.30	-10.09	4.67	-1.24	-5.88	3.30	-10.09	4.67
	1ARNC	1.62	NA	-1.24	-11.22	0.91	1.62	NA	-1.24	-11.22	0.91
	2AC15	2.13	-3.74	-0.06	0.08	4.62	1.96	-3.96	-0.33	-0.37	5.12
	2AC21	3.73	-4.14	1.95	1.50	3.40	5.07	-4.27	1.83	0.94	4.20
	2AC5	8.52	-1.89	2.82	-1.03	9.49	-0.99	-8.75	0.28	-6.10	7.36
	2ANC	0.07	-4.18	-0.74	-12.44	3.38	-0.13	-4.62	-0.99	-10.99	4.00
	2ARNC	0.79	0.34	-3.64	-3.89	1.75	-0.54	-3.03	-4.92	-4.49	1.79
Sun	CC	-0.35	-1.34	-2.28	9.29	1.66	-0.28	-1.50	-2.58	7.69	2.03
	F18	1.29	-3.37	0.60	1.98	1.04	2.72	-2.20	0.91	1.71	1.74
	NF12	2.75	-2.65	-1.98	1.31	5.07	0.56	-3.57	-4.57	-0.29	4.70
	NF17	2.36	-2.80	2.32	-0.30	2.67	3.89	-2.66	2.60	-0.14	3.67
	NF6	8.61	-2.05	3.06	-0.83	9.43	-0.94	-8.88	0.33	-5.69	7.48
	OA	2.40	0.49	0.11	-15.84	0.90	2.33	0.42	0.06	-15.84	0.78
	OC	2.48	6.75	-6.57	-0.05	-0.55	1.02	7.28	-8.71	-0.93	-1.90
	OTHR	-3.22	-3.91	0.99	-9.35	0.86	-5.41	-5.06	1.63	-10.69	1.14
	SA	-2.21	-4.24	1.94	-7.76	3.70	-2.22	-4.24	1.94	-7.76	3.70
	SP18	3.10	-3.00	2.76	-2.46	2.52	1.67	-3.44	3.24	-3.63	2.20
Marker	CH1	5.21	-0.88	2.07	-2.57	4.85	0.63	-5.26	1.09	-6.61	3.71
	CH2	4.32	1.06	0.58	-2.03	5.23	-3.89	-4.21	-2.09	-4.93	3.83
	CH3	3.77	-1.23	1.08	-0.00	4.58	-1.45	-4.40	-0.27	-1.02	3.49
	MX1	5.05	-0.21	1.41	1.10	3.31	0.87	-3.09	0.92	-1.65	2.19
	MX2	3.75	-4.39	1.79	0.14	3.46	1.73	-5.78	1.80	-2.18	2.97
	MX3	2.34	-1.00	1.32	0.49	2.63	-1.11	-2.44	-0.02	-1.16	1.76
	NC1	-2.40	-6.10	-1.67	-3.30	3.63	-6.76	-7.09	-5.17	-7.06	3.55
	NC2	-1.45	-1.89	1.09	-3.48	-2.05	-1.64	-1.92	1.48	-3.17	-2.26
	NC3	-4.98	-0.24	-3.48	-4.35	-3.65	-3.46	1.31	-2.68	-4.20	-3.30
	NC4	2.94	1.57	-0.77	-0.19	1.30	3.43	1.52	-1.05	-0.22	1.36
	NC5	1.61	0.71	-4.19	-1.04	3.03	0.12	0.33	-6.16	-6.27	3.00
	OTHR	-0.22	1.48	-2.64	0.54	-0.19	-0.03	0.86	-1.15	1.63	-1.22
	SA1	-0.77	-5.81	-2.69	-1.94	6.75	0.22	-2.08	-1.68	-1.19	7.37
	SA2	3.31	-4.65	1.02	2.09	5.08	1.77	-5.56	0.49	0.57	4.75

For shaded cells, accept H_0, evidence of temporal stability

Table 15 (continued)

Life Cycle Variables		Household Trips Total	HBW	HBSH	HBSC	HBO	Person Trips Total	HBW	HBSH	HBSC	HBO
		Calculated t-statistic									
Marker	SA3	1.71	-2.53	1.92	-0.71	1.44	2.94	-2.30	2.07	-0.67	2.26
Marker	SFM1	-0.21	-3.74	-1.06	-2.53	3.72	-0.21	-3.74	-1.06	-2.53	3.72
Marker	SFM2	0.70	0.97	-1.41	0.68	0.77	0.70	0.97	-1.41	0.68	0.77
Marker	SFM3	1.89	-0.61	-0.17	NA	0.89	1.89	-0.61	-0.17	NA	0.89
Marker	SM1	-3.19	-2.15	1.24	-0.43	9.41	-3.19	-2.15	1.24	-0.43	9.41
Marker	SM2	0.46	-1.38	2.41	-3.67	0.38	0.46	-1.38	2.41	-3.67	0.38
Zimmerman	C3049NA	0.07	-0.92	-0.69	2.47	0.47	0.41	-0.46	-1.15	2.38	0.62
Zimmerman	C5059NA	0.07	-0.72	-1.87	-0.24	0.09	0.98	-0.15	-1.82	-0.20	0.47
Zimmerman	C60NA	2.59	2.42	-4.66	-0.50	2.31	1.37	2.25	-6.49	-6.09	1.72
Zimmerman	CAA06	7.96	-1.54	2.24	-1.86	8.73	-1.80	-8.43	-0.50	-5.61	6.81
Zimmerman	CAA1115	1.06	-4.22	0.46	-0.01	3.16	-0.73	-4.80	-0.19	-1.33	3.15
Zimmerman	CAA1618	1.51	-2.65	-0.47	-0.61	2.22	2.75	-1.85	-11.63	-0.36	3.11
Zimmerman	CAA1922	3.23	-1.23	1.04	1.87	2.45	3.78	-2.57	0.76	1.58	4.09
Zimmerman	CAA710	6.44	-1.49	2.22	-1.57	5.65	3.69	-3.55	1.80	-3.38	6.15
Zimmerman	NC06[1]	1.76	-1.77	1.17	-1.72	2.60	-0.60	-5.22	1.19	-3.71	2.49
Zimmerman	NC1115	2.27	0.85	1.40	-0.29	-0.38	-0.01	-0.29	0.91	-1.86	-1.76
Zimmerman	NC1618	0.95	-3.32	2.50	-5.08	1.74	4.66	-1.11	3.79	-2.94	3.27
Zimmerman	NC1922[1]	-0.46	0.77	0.13	-0.35	-1.33	-2.73	-0.46	-0.26	-1.28	-2.77
Zimmerman	NC30NA	-2.59	-2.38	0.94	-0.72	0.19	-4.99	-3.88	-2.31	-1.46	-0.71
Zimmerman	NC55NA	-2.39	-6.46	1.61	-4.76	2.35	-2.16	-6.61	1.70	-5.93	2.65
Zimmerman	NC69NA	1.24	1.63	0.83	-0.84	-0.29	1.27	1.77	0.86	-0.84	-0.35
Zimmerman	NC710[1]	1.14	-3.10	0.33	-1.09	2.39	0.56	-4.25	0.18	-1.37	2.53
Zimmerman	OP7079NA	1.58	0.28	-0.48	-1.86	0.09	1.54	0.30	-0.52	-1.86	0.04
Zimmerman	OP80NA	0.09	-1.09	0.37	-2.48	0.14	0.41	-0.88	0.06	-2.61	0.91
Zimmerman	OTHR	-0.22	1.48	-2.64	0.54	-0.19	-0.03	0.86	-1.15	1.63	-1.22
Zimmerman	YC30NA	2.48	-5.84	-2.67	NA	3.52	2.03	-11.93	-3.20	-7.74	4.07
TOTAL TRIPS		6.15	-7.45	2.95	-4.79	11.3	1.05	-3.24	0.66	-9.73	12.6

For shaded cells, accept H_0, evidence of temporal stability

4.7 Summary and Recommendations

This paper investigated several definitions for describing the structure of households to assess the value of each in explaining trip making behavior between life cycle strata, trip purpose, and across time. Six life cycle definitions came from previous research. ANOVA was used to evaluate the performance of the six definitions in explaining differences in travel behavior. Eta-square provided insights into the strength of the difference between the means for each life cycle definition and variable-to-variable comparisons were made using the Bonferroni multiple-comparison test. Considering only a p-value less than 0.05 at a 95% confidence level, ANOVA shows that all life cycle definitions are useful in explaining differences in trip-making behavior for all six life cycle definitions. The NHTS definition showed good variability in trip rates across all life cycle categories, and the greatest variability for work and shopping trips compared to the other definitions. The variability across all trip purposes was greatest for the Sun definition as was the variability for total trips. The life cycle definition defined by Marker showed the greatest degree of variability for shopping and other trips, while the Zimmerman definition had the highest degree of variability for work and school trips. The Kermanshah categories displayed high variation for school trips, but little variation for other trip purposes while the Vadarevu categories capture the variance in work and school trips but not so for other trip purposes. Considering the strength of the difference between the means, as well as the differences between specific means, the NHTS, Sun, Marker, and Zimmerman definitions are superior in explaining trip making behavior. In variable-to-variable comparisons, the definition used by Sun outperformed the NHTS with the exception of shopping trips. The Marker definition performed better than Zimmerman did for total, shopping, and other trips, but Zimmerman is better for work and school trips.

Evaluating the temporal stability of the most promising life cycle definitions at a 95% confidence interval provides encouraging evidence regarding the stability of trip rates for most strata within each of the four definitions. Considering the number of stratifications that are temporally stable, total person trips show greater stability than household trips. In comparison, trip rates using the Sun definition of life cycle are less stable than the other definitions, with no more than 50% of the strata remaining stable for household or person trips for any given trip purpose. Only 20% of the strata are stable for total household trips, household work trips, and person work trips. Trip rates for the NHTS definition were stable for 60% of the strata for household trips and 80% for person trips. By trip purpose, shopping trips had the greatest number of strata that were temporally stable and person work trip the lowest number of strata that were stable. The Zimmerman definition shows the greatest stability of all four definitions, as measured by looking at the highest percentage of strata with stability for the majority of the trip purpose categories for both household and person trips. Total household and person trips were stable for 60% of the strata; shopping trips were stable for 70 and 75% of the strata for household and person trips respectively. The poorest performing was HBO trips showing only 45% of the strata as temporally stable. Trip rates using the Marker definition were most stable for household work and school trips, 65% and 70% respectively, and person shopping trips, 70%. Only 30% of the strata were temporally stable for HBO household and person trips. In summary, temporal stability for all life cycle definitions is best for shopping trips while total household trips are less stable for households with children. The results of these tests support the use of the life cycle as a key variable for trip generation models, as there is evidence from this research that life cycle has a strong influence on trip making while also providing improved stability in trip rates over time.

This research effort generally supports the development of synthetic populations and advanced trip generation models that allow more independent variables in the model specification thereby capturing more of the factors that influence travel behavior. Future research in this area should focus on model estimation using life cycle to see how well life cycle contributes to improved model fit, performance, and temporal stability. Future research should also explore how these variables influence tour rates and trip chaining, especially with the growing interest in tour-based models.

CHAPTER 5. INFLUENCE OF AREA TYPE AND ACCESSIBILITY

5.1 Introduction and Motivation

A travel demand model is a series of mathematical equations that represent how people make travel decisions. Travel models are important because the forecasts output from these models guide decision makers in the evaluation and selection of transportation programs and projects. Developing a better understanding of the factors that influence travel behavior, the changes in travel behavior over time, and the variables that best capture these changes may lead to the development of models that are more stable over time, increasing our confidence in model results and leading to more cost effective investment decisions.

Previous research identified accessibility and area type (or urban form) as key variables for predicting travel behavior (Kockelman 1997; Purvis 1998; Thill and Kim 2005; Froehlich 2008). Given the promising results, one could hypothesize that these variables may also be key variables leading to improved stability of trip generation models over time. Accessibility is a measure of one's ability to reach goods, services, activities, and destinations (McGurrin 2011). Area type is a term used to describe the character of the land use. In a most traditional sense, it can describe rural, suburban, or urban areas defined purely based on density measures, but can also refer to measures of land use balance or land use mix. This paper explores density based measures and land use mix.

The purpose of this research is to evaluate definitions of accessibility and area type identified in the literature in order to assess the value of these concepts for explaining trip making behavior and for improving the temporal stability of trip generation models.

5.2 Literature Review

Models are simplified representations of reality. A travel demand model is a series of mathematical equations used to describe travel and travel choices. A number of factors influence these choices including characteristics of the household, the trip maker, the trip purpose, and the transportation system. Historically, the common factors considered when estimating the number of trips included income, auto ownership, and household size (FHWA 1975; Ortuzar and Willumsen 2011). Recent studies have also considered number of workers and children (PB 2007). While no model can take into account all of the factors that influence travel behavior, selecting factors that most influence the decision will lead to models that are more robust, leading to increased confidence in the model forecasts. Area type and accessibility are two such factors.

5.2.1 Accessibility

The literature cites mixed results for including accessibility in trip generation models, noting models parameters that are insignificant or have the wrong sign (Thill and Kim 2005). In his research on the role of accessibility measures on trip making behavior, Thill notes that many research efforts that produced relationships that are either non-significant or produced the wrong sign were focused on aggregate modeling applications (Thill and Kim 2005). Results of this nature suggest that one's decision to travel is independent of infrastructure and the level of service it provides (Thill and Kim 2005). Thill argues for the testing and application of more complex accessibility measures or the development of disaggregate models in order to achieve more intuitive, and positive, results (Thill and Kim 2005). Specifically, his research investigates two different model types (cumulative and gravity), five spatial impedance functions (inverse power, negative exponential, Gaussian, rectangular, and negative linear), and a range of parameter values. The results of this research show a significant relationship between several measures of accessibility and trip generation models, achieved through the careful design and testing of a family of accessibility metrics with the inclusion of multiple measures within one model often resulting in greater explanatory power than the included socio-economic variables (Thill and Kim 2005). These findings lead to a conclusion and recommendation that model developers not search for a single "best" definition of accessibility, as only a portion of what accessibility represents is captured by any one definition, but rather to test and implement a family of measures (Thill and Kim 2005).

In another study, researchers investigated not only the sensitivity of accessibility measures on trip making, but also their influence on trip making over time (Froehlich 2008). The study found that trip making reacted with positive elasticity to accessibility, but that over time accessibility became less important for work trip generation (Froehlich 2008).

While the question of accessibility has not been fully resolved, it is clear from the studies outlined above, as well as several others that have documented the benefits of including accessibility measures in trip generation models (Purvis 1996; Kockelman 1997), that the literature will benefit from additional research investigating this measure.

5.2.2 Area Type

Density or area type is another variable that can better reflect trip-making behavior. (Liss, McGuckin et al.) documents the influential role that urban, suburban, and rural location have on travel behavior using the 2001 NHTS, noting that the average person trips per person increases as one moves from urban to rural settings. Purvis (1998) cited neighborhood density variables as leading to improved performance of travel demand models while Meurs (1990) in his work on the unobserved effects of trip generation models, notes that spatial characteristics related to city size are shown to affect trip generation. Unlike the previously cited studies, density was found to be insignificant when controlled for accessibility (Kockelman 1997). Land development characteristics were found to have a highly significant influence on travel behavior when measured as land use balance (Kockelman 1997). These results support the findings from other work that notes that measures such as residential density, employment density, and mix of land use influence trip generation (Ross and Dunning 1997; Kitamura, Akiyama et al. 2001; Litman 2011).

5.3 Methodology

The focus of this analysis is on the investigation of different operational definitions of accessibility and area type as defined in the previous section, and an assessment of their usefulness in explaining trip making behavior between strata, across time horizons and trip purposes. This analysis considers four trip purposes along with total trips for both household and person level trips. The trip purposes include home-based work trips (HBW), home-based shopping trips (HBSH), home-based school trips (HBSC), and home-based other trips (HBO) where other includes trip purposes not defined by work, shopping, or school.

5.3.1 Data

The datasets supporting this research come from activity based household travel surveys administered in the Research Triangle Region of North Carolina in 1995 and 2006. Specifically, these data include the 1994/1995 Triangle Travel Behavior Survey (NuStats 1995), and the 2006 Greater Triangle Travel Survey (NuStats 2006). The Triangle Region describes the geographic area covered by the Capital Area Metropolitan Planning Organization (CAMPO) and the Durham-Chapel Hill-Carrboro Metropolitan Planning Organization (DCHC MPO). Together the boundaries of the two MPOs cover eight counties with a combined MPO population of 1.4 million.

The 1995 Survey covers the period between November 1994 and April 1995 and the North Carolina counties of Wake, Durham, Orange and portions of Harnett, Chatham, Person, Granville, Franklin, and Johnston, see Figure 7.

The 2006 Survey covered the period between January and June 2006 and the full North Carolina counties of Wake, Durham, Orange, Chatham, Franklin, Granville, Johnston, Lee, Person, and Vance; and portions of Harnett and Nash, see Figure 7.

Both the 1995 and 2006 Triangle surveys covered a geographic area larger than the three core counties of Wake, Durham, and Orange, but this coverage varied between the two survey years. Coding of the survey records facilitated the removal of records outside the core counties in order to create geographic consistency between the two years. All further references to the Triangle region refer to Wake, Durham, and Orange counties only.

Figure 7. Triangle Survey Boundaries

The development of new trip purpose codes assured consistency between the datasets. This step included processing unlinked trips into linked trips, where an unlinked trip refers to a very short segment trip that is a part of a traveler's daily activity but not traditionally considered a separate trip for the purposes of travel demand modeling. Identical linking procedures for both surveys helps eliminate any differences related to trip definitions. Finally, overall reasonableness checking of the survey data led to the identification and repair of coding errors and inconsistencies.

5.3.2 Statistical Tests

An analysis of variance (ANOVA) using all accessibility and area type definitions and the strata within each definition provided insight into which definitions and strata that best explained travel behavior. ANOVA tests the null hypothesis (H_0) that all means for a given measure are equal. Likewise, the alternative hypothesis (H_a) that at least one mean differs is also tested. Using the p-value, we reject H_0 if the p-value is less than 0.05 at a 95% confidence level.

ANOVA tests whether the means differ, but does not tell us which means differ. A multiple comparison test evaluates the differences between specific means in order to determine which means differ. For this research the Bonferroni multiple-comparison test (Acock 2006) was used with the critical value of p less than 0.05; in other words, if the calculated value for p is greater than 0.05 then the difference between any given set of means is not significant (Acock 2006). Finally, eta-square (η^2) quantifies the strength of the difference between the means:

$$\eta^2 = \frac{SSB}{SST} \qquad \text{(Equation 8)}$$

where:
η^2 = eta square, a measure of explained variance (ANOVA equivalent of r^2)
SSB = between group sum of squares
SST = total sum of squares.

Another important characteristic to test is the temporal stability of the various definitions. The application of trip generation models for future forecasts depends heavily on the assumption that the travel behavior represented by the model remains stable over time. Including variables in the models that improve the temporal stability of the models will lead to improved modeling tools and more defensible forecasts. For this analysis, the t-statistic was used to test the null hypothesis of no significant difference between the means estimated from the 1995 survey and the 2006 survey. The assumptions for this test are independent random samples from two populations, normally distributed with equal variance in each population:

$$t = \frac{(\bar{y}_1 - \bar{y}_2)}{s_p(\sqrt{\frac{1}{n_1} + \frac{1}{n_2}})}$$ (Equation 9)

$$s_p = \sqrt{\frac{(n_1-1)s_1^2 - (n_2-1)s_2^2}{n_1+n_2-2}}$$ (Equation 10)

$$df = n_1 + n_2 - 2$$ (Equation 11)

where:

t = test statistic

$\bar{y}_1$ = mean of sample, 1995 survey

$\bar{y}_2$ = mean of sample, 2006 sample

n_1 = sample size, 1995 sample

n_2 = sample size, 2006 sample

s_p = pooled sample standard deviation

s_1^2 = sample variance, 1995 sample

s_2^2 = sample variance, 2006 sample

df = degrees of freedom.

The null hypothesis of no significant difference between means is rejected when the absolute value of the test statistic, t, is greater than or equal to $t_{\frac{\alpha}{2}}$,df. The test was applied at 95% confidence level with $t_{\frac{\alpha}{2}}$,df equal to 1.96.

5.4 Accessibility Definitions

This research focused on four functional forms, three impedance measures, and two opportunity measures for a combination of 24 accessibility measures tested. The selection of these various combinations is motivated by the fact that they all offer a slightly different measure of accessibility based on distance, travel time, and a generalized cost along with various functional forms.

The general form for measuring accessibility is (Ortuzar and Willumsen 2011):

$$A_i = \sum_j f(E_j^n, C_{ij}) \qquad \text{(Equation 12)}$$

where:
A_i = measure of accessibility for zone *i*
E_j^n = measure of attraction of zone *j*
C_{ij} = impedance separating zones *i* and *j*.

The measure of attraction, or opportunity, of zone j (E_j^n) is typically expressed by the number of jobs, number of jobs by type, total population, or some combination of these variables. This research explored total employment, total employment plus total population, and total retail employment. The measure of impedance (C_{ij}) considered in this research included zone-to-zone travel time (TT) in minutes, distance (DI) in miles, and generalized cost (GC) in minutes. The generalized cost considered in this research is:

$$GC = t + a * d \qquad \text{(Equation 13)}$$

$$a = \frac{aoc}{p*\frac{wr}{60}} \qquad \text{(Equation 14)}$$

where:
GC = generalized cost in minutes
t = link travel time in minutes

d = link distance in miles

a = distance coefficient

aoc = auto operating cost in dollars/mile (value used for research 0.094) (MTC 1998)

p = purpose factor (value used for research work trips = 0.5, other trips = 0.25) (PB 2008)

wr = hourly wage rate in dollars per hour (14.37 for 1995, 19.99 for 2006; www.bls.gov).

The four functional forms considered include a power function, two variations on an exponential function, and a combination power exponential function.

$$PowerFunction: C_{ij}^{-2}$$

$$Exponential_1: e^{-C_{ij}}$$

$$Exponential_2: e^{-0.1C_{ij}}$$

$$\mathrm{PowerExponential: C_{ij}^{0.5} * e^{-0.1C_{ij}}}$$

For simplicity each of the combinations of four functional forms, three impedance measures, and two opportunity measures were assigned a unique variable name as defined below.

Variables using power function (P2) where measure of opportunity equal to employment is E, opportunity equal to employment plus population is EP, and where measure of impedance equal to travel time TT, distance is DI, and generalized cost is GC:

- P2E_TT
- P2EP_TT
- P2E_DI
- P2EP_DI
- P2E_GC
- P2EP_GC

Variables using function exponential1 where measures of opportunity and impedance are as defined above:

- E1E_TT
- E1EP_TT
- E1E_DI
- E1EP_DI
- E1E_GC
- E1EP_GC

Variables using function exponential2 where measures of opportunity and impedance are as defined above:

- EE_TT
- EEP_TT
- EE_DI
- EEP_DI
- EE_GC
- EEP_GC

5.5 Area Type Definitions

The area type analysis included definitions for density, land use density index, and land use mix. Density is a simple calculation of units per acre where units represents the land use feature of interest. This research considered population density (PopD), residential density (ResD), employment density (EmpD), and a composite density of residential plus employment (CmpD). Unlike the simple density calculation of units per acre, the land use density index (LUDI) considers the relationship between the employment in a given zone (E_i) and the ratio (R) of total regional households (HH_r) to total regional employment (E_r) (Huntsinger, Nwoko et al. 1999):

$$LUDI = \frac{HH_i + (R * E_i)}{Area_i} \qquad \text{(Equation 15)}$$

where:

LUDI = land use density index

HH_i = number of households in zone i

Ei = total employment in zone i

$Area_i$ = area per square mile for zone i

R = ratio of total regional households (HH_r) to total regional employment (E_r).

Land use mix (LUM) captures the mix between different types of land use in a given zone. It varies between one and zero, where one indicates perfect mix and zero indicates no mix, for example a zone with only residential development. A measure of entropy is used to calculate land use mix. Entropy has been used previously to capture land use balance where the focus in on the proportion of developed land by land use type (Kockelman 1997). This same concept can be applied to reflect the mix of land use types as measured by housing units and employment:

$$LUM_i = -\sum_{j=1}^{6} \frac{[P_j * \ln(P_j)]}{\ln(J)} \qquad \text{(Equation 16)}$$

where:

LUM_i = land use mix for zone i

P_j = proportion of land use type in the jth use type

J = number of use types (six uses were considered for this research: residential, industrial, retail, high traffic retail, office, and service).

5.6 Analysis and Results

5.6.1 Accessibility, Area Type and Travel Behavior

Given the large number of accessibility definitions under consideration, an initial screening technique led to the identification and elimination of variables that had minimal contribution to variation in household trip rates by trip purpose between the various strata specified for each variable. Using the centile function in (StataCorp 2007) ten strata were generated for the accessibility variables. Household trip rates by trip purpose were calculated for each strata and ANOVA used to eliminate variables using the F-statistic at a 90% confidence level. Simply stated, we removed these definitions for accessibility from further consideration.

Tables 16 and 17 show p-value and eta-square values for household and person trips by trip purpose for the definitions retained for further investigation.

Table 16. Calculated p-value and eta-square (η^2) for Accessibility Mean Trip Rates

<table>
<tr><th rowspan="2">Trip Purpose</th><th colspan="2">P2EP_DI</th><th colspan="2">EE_DI</th><th colspan="2">PEEP_DI</th><th colspan="2">E1E_TT</th><th colspan="2">E1EP_GC</th></tr>
<tr><th>p</th><th>η^2</th><th>p</th><th>η^2</th><th>p</th><th>η^2</th><th>p</th><th>η^2</th><th>p</th><th>η^2</th></tr>
<tr><td colspan="11">1995 Household Trips</td></tr>
<tr><td>Total Trips</td><td>0.08</td><td>0.01</td><td>0.07</td><td rowspan="3">0.01</td><td>0.02</td><td rowspan="3">0.01</td><td>0.05</td><td rowspan="5">0.01</td><td>0.08</td><td rowspan="5">0.01</td></tr>
<tr><td>HBW</td><td>0.84</td><td>0.00</td><td>0.34</td><td>0.09</td><td>0.04</td><td>0.04</td></tr>
<tr><td>HBSH</td><td>0.27</td><td rowspan="3">0.01</td><td>0.06</td><td>0.03</td><td>0.16</td><td>0.36</td></tr>
<tr><td>HBSC</td><td>0.19</td><td rowspan="2">0.01</td><td rowspan="2">0.02</td><td><.001</td><td>0.02</td><td>0.25</td><td>0.18</td></tr>
<tr><td>HBO</td><td>0.10</td><td>0.18</td><td>0.01</td><td>0.04</td><td>0.32</td></tr>
<tr><td colspan="11">2006 Household Trips</td></tr>
<tr><td>Total Trips</td><td><.001</td><td rowspan="2">0.01</td><td><.001</td><td>0.01</td><td>0.75</td><td rowspan="5">0.00</td><td><.001</td><td rowspan="2">0.01</td><td><.001</td><td>0.02</td></tr>
<tr><td>HBW</td><td>0.01</td><td>0.29</td><td rowspan="3">0.00</td><td>0.22</td><td>0.02</td><td>0.05</td><td rowspan="2">0.00</td></tr>
<tr><td>HBSH</td><td>0.61</td><td>0.00</td><td>0.63</td><td>0.16</td><td>0.22</td><td>0.00</td><td>0.34</td></tr>
<tr><td>HBSC</td><td rowspan="2"><.001</td><td rowspan="2">0.01</td><td>0.23</td><td>0.53</td><td rowspan="2"><.001</td><td rowspan="2">0.01</td><td rowspan="2"><.001</td><td rowspan="2">0.01</td></tr>
<tr><td>HBO</td><td>0.11</td><td>0.90</td></tr>
<tr><td colspan="11">1995 Person Trips</td></tr>
<tr><td>Total Trips</td><td><.001</td><td rowspan="3">0.01</td><td><.001</td><td rowspan="5">0.01</td><td><.001</td><td>0.02</td><td><.001</td><td rowspan="3">0.01</td><td><.001</td><td>0.00</td></tr>
<tr><td>HBW</td><td>0.11</td><td>0.01</td><td>0.23</td><td>0.00</td><td>0.01</td><td>0.09</td><td rowspan="2">0.01</td></tr>
<tr><td>HBSH</td><td>0.02</td><td rowspan="2"><.001</td><td rowspan="2"><.001</td><td rowspan="3">0.01</td><td><.001</td><td>0.01</td></tr>
<tr><td>HBSC</td><td>0.58</td><td>0.00</td><td>0.90</td><td>0.00</td><td>0.39</td><td>0.00</td></tr>
<tr><td>HBO</td><td>0.02</td><td>0.01</td><td>0.01</td><td>0.02</td><td><.001</td><td>0.01</td><td>0.01</td><td>0.01</td></tr>
<tr><td colspan="11">2006 Person Trips</td></tr>
<tr><td>Total Trips</td><td>0.01</td><td rowspan="5">0.00</td><td><.001</td><td rowspan="5">0.00</td><td>0.62</td><td rowspan="5">0.00</td><td>0.01</td><td rowspan="4">0.00</td><td>0.02</td><td rowspan="5">0.00</td></tr>
<tr><td>HBW</td><td>0.06</td><td>0.24</td><td>0.08</td><td rowspan="2">0.02</td><td>0.03</td></tr>
<tr><td>HBSH</td><td>0.02</td><td>0.11</td><td>0.01</td><td>0.02</td></tr>
<tr><td>HBSC</td><td>0.01</td><td>0.45</td><td>0.33</td><td rowspan="2"><.001</td><td>0.01</td></tr>
<tr><td>HBO</td><td><.001</td><td>0.36</td><td>0.91</td><td>0.01</td><td><.001</td></tr>
</table>

As shown in Table 16 the only accessibility definition where the value of household trips varies significantly for both analysis years is E1E_TT. The accessibility definitions investigated capture more variation in the total person trips rates between analysis years with all definitions showing a p-value less than 0.05 except PEEP_DI. The accessibility definitions E1E_TT and E1EP_GC show the greatest variation in both household and person trip rates by trip purpose for both analysis years, while EE_DI and PEEP_DI capture less variation in comparison. P2EP_DI does not capture as much variation for household trips by trip purpose, but does well when considering person trips by trip purpose.

Table 17. Calculated p-value and eta-square (η^2) for Area Type Mean Trip Rates

<table>
<tr><th rowspan="2">Trip Purpose</th><th colspan="2">ResD</th><th colspan="2">PopD</th><th colspan="2">EmpD</th><th colspan="2">CmpD</th><th colspan="2">LUDI</th><th colspan="2">LUM</th></tr>
<tr><th>p</th><th>η^2</th><th>p</th><th>η^2</th><th>p</th><th>η^2</th><th>p</th><th>η^2</th><th>p</th><th>η^2</th><th>p</th><th>η^2</th></tr>
<tr><td colspan="13">1995 Household Trips</td></tr>
<tr><td>Total</td><td>0.03</td><td rowspan="3">0.01</td><td>0.30</td><td rowspan="3">0.01</td><td><.001</td><td>0.01</td><td>0.10</td><td>0.01</td><td rowspan="2"><.001</td><td>0.01</td><td>0.01</td><td>0.01</td></tr>
<tr><td>HBW</td><td rowspan="2">0.06</td><td>0.15</td><td>0.08</td><td rowspan="2">0.00</td><td>0.02</td><td>0.02</td><td>0.00</td><td>0.43</td><td rowspan="2">0.00</td></tr>
<tr><td>HBSH</td><td>0.12</td><td>0.56</td><td>0.13</td><td>0.01</td><td>0.02</td><td>0.01</td><td>0.31</td></tr>
<tr><td>HBSC</td><td>0.77</td><td>0.00</td><td>0.87</td><td>0.00</td><td><.001</td><td rowspan="2">0.01</td><td>0.82</td><td rowspan="2">0.00</td><td>0.12</td><td>0.00</td><td><.001</td><td>0.01</td></tr>
<tr><td>HBO</td><td>0.40</td><td>0.01</td><td>0.43</td><td>0.01</td><td>0.03</td><td>0.69</td><td>0.02</td><td>0.01</td><td>0.45</td><td>0.00</td></tr>
<tr><td colspan="13">2006 Household Trips</td></tr>
<tr><td>Total</td><td rowspan="2"><.001</td><td rowspan="2">0.01</td><td><.001</td><td rowspan="2">0.01</td><td><.001</td><td>0.01</td><td rowspan="2"><.001</td><td>0.02</td><td rowspan="2"><.001</td><td>0.01</td><td><.001</td><td>0.01</td></tr>
<tr><td>HBW</td><td>0.01</td><td>0.23</td><td rowspan="2">0.00</td><td>0.01</td><td rowspan="4">0.00</td><td>0.60</td><td rowspan="2">0.00</td></tr>
<tr><td>HBSH</td><td>0.83</td><td rowspan="2">0.00</td><td>0.89</td><td rowspan="2">0.00</td><td>0.21</td><td>0.44</td><td>0.00</td><td>0.97</td><td>0.44</td></tr>
<tr><td>HBSC</td><td>0.07</td><td>0.28</td><td rowspan="2"><.001</td><td rowspan="2">0.01</td><td rowspan="2"><.001</td><td rowspan="2">0.01</td><td>0.34</td><td rowspan="2"><.001</td><td>0.01</td></tr>
<tr><td>HBO</td><td><.001</td><td>0.01</td><td><.001</td><td>0.01</td><td><.001</td><td>0.00</td></tr>
<tr><td colspan="13">1995 Person Trips</td></tr>
<tr><td>Total</td><td><.001</td><td>0.01</td><td>0.02</td><td>0.01</td><td rowspan="2"><.001</td><td rowspan="5">0.00</td><td><.001</td><td>0.02</td><td><.001</td><td>0.01</td><td>0.28</td><td>0.00</td></tr>
<tr><td>HBW</td><td>0.86</td><td>0.00</td><td>0.41</td><td>0.00</td><td>0.19</td><td>0.00</td><td>0.19</td><td rowspan="3">0.00</td><td>0.01</td><td>0.01</td></tr>
<tr><td>HBSH</td><td>0.04</td><td>0.01</td><td>0.05</td><td>0.01</td><td>0.29</td><td>0.04</td><td>0.01</td><td>0.07</td><td>0.69</td><td>0.00</td></tr>
<tr><td>HBSC</td><td>0.19</td><td>0.00</td><td>0.44</td><td>0.00</td><td>0.10</td><td>0.87</td><td>0.00</td><td>0.76</td><td>0.01</td><td>0.01</td></tr>
<tr><td>HBO</td><td>0.04</td><td>0.01</td><td>0.06</td><td>0.01</td><td>0.02</td><td>0.09</td><td>0.01</td><td><.001</td><td>0.01</td><td>0.12</td><td>0.0</td></tr>
<tr><td colspan="13">2006 Person Trips</td></tr>
<tr><td>Total</td><td>0.01</td><td rowspan="5">0.00</td><td>0.01</td><td rowspan="5">0.00</td><td rowspan="2"><.001</td><td rowspan="5">0.00</td><td>0.01</td><td rowspan="5">0.00</td><td>0.02</td><td rowspan="5">0.00</td><td>0.02</td><td rowspan="5">0.00</td></tr>
<tr><td>HBW</td><td><.001</td><td>0.03</td><td><.001</td><td><.001</td><td><.001</td></tr>
<tr><td>HBSH</td><td>0.57</td><td>0.19</td><td>0.19</td><td>0.15</td><td>0.02</td><td>0.23</td></tr>
<tr><td>HBSC</td><td>0.14</td><td>0.43</td><td>0.02</td><td>0.03</td><td>0.95</td><td><.001</td></tr>
<tr><td>HBO</td><td>0.03</td><td>0.07</td><td><.001</td><td>0.06</td><td>0.07</td><td>0.02</td></tr>
</table>

As indicated in Table 17 the mean value of total household trips varies significantly for all definitions of area type with the exception of PopD and CmpD. The area type definitions show more variance in the average household trip rate by trip purpose for the 2006 data than for the 1995 data. Area type as defined by PopD and CmpD show no variation in 1995 trip rates by trip purpose across the various strata, while LUDI best captures differences in the 1995 trips rates by trip purpose. For 2006, EmpD, CmpD, and LUM are best for HBSC and

HBO trips. All area type definitions except EmpD and LUM capture differences in HBW trip rates by strata, while none of the definitions captures differences in HBSH trip rates based on the 2006 data.

The mean value of total person trips also varies significantly for all definitions of area type with the exception of LUM for the 1995 data only. As with the household trip rates by trip purpose, the various definitions reflect more differences in person trip rates by trip purpose for the 2006 data. EmpD best captures the differences in trip rates for HBW and HBO trip purposes in both analysis years, but only the HBSC trip purpose in the 2006 data. LUDI is best for HBW and HBO trips in the 1995 data, and better for HBSH trips in the 2006 data. LUM captures differences in HBW and HBSC trips for both analysis years, and HBO trips in the 2006 data. PopD and CmpD capture very little of the variation in person trips rates by trip purpose.

Based on this comparison of p-values, ResD, PopD, and CmpD were found to be the poorest performing in the area type group, while EE_DI, and PEEP_DI are the poorest performing in the accessibility group. There is little to no significant difference in trips by trip purpose for ResD, PopD, or CmpD in the 1995 household level analysis and for CmpD in the 1995 person level analysis. Similarly, there is little to no significant difference in trips by trip purpose for EE_DI and PEEP_DI in the 2006 household and person trip analysis. In contrast, the other definitions show variation across both analysis years and by trip purpose.

While p-value results are encouraging with respect to identifying area type and accessibility definitions that capture the differences in trip rates, the η^2 value tells a different story with respect to the strength of the differences. None of the cells in Tables 16 and 17 shows a value greater than 0.02, indicating that the strength of the difference is not very good between the original ten strata for the area type and accessibility definitions considered. This finding indicates a need for further refinement of the strata to perhaps fewer than ten. This iterative process used the output from the Bonferroni multiple-comparison test to identify candidate strata for combining until the final classification scheme best-reflected differences between the mean trip rates. Table 18 shows the final stratifications for the area type and accessibility definitions retained for further investigation. Table 19 shows trip rates and sample size for these definitions.

Table 18. Final Stratifications for Area Type and Accessibility Definitions

Area Type Definitions				Accessibility Definitions			
EmpD	1	0	29.28	P2EP_DI	1	0	6747
	2	29	680.00		2	6747	13649
	3	681	178000		3	13649	34320
	Units are employees per square mile				4	34320	1150000
LUDI	1	0	712	E1E_TT	1	0	15
	2	712	3161		2	15	102
	3	3161	151000		3	102	647
	Units are households + factored employees per square mile				4	647	14000
LUM	1	0.00	0.16	E1EP_GC	1	0	23
	2	0.16	0.59		2	23	134
	3	0.59	1.00		3	134	437
	Units are land use mix index ranging from 0, no mix, to 1, perfect mix				4	437	10600

Table 19. Household and Person Mean Trip Rates and Sample Size (n)

Area Type		Mean Trip Rates and Sample Size (n)							
		Household Trips				Person Trips			
		1995		2006		1995		2006	
		Mean	n	Mean	n	Mean	n	Mean	n
EmpD	1	9.19	390	9.40	478	4.00	895	3.99	1127
	2	8.12	619	9.61	2038	4.09	1228	4.18	4689
	3	7.87	343	8.17	928	4.41	612	4.27	1777
LUDI	1	8.70	303	9.45	578	3.76	701	4.04	1351
	2	8.79	690	9.68	1598	4.18	1453	4.24	3647
	3	7.26	359	8.47	1268	4.49	581	4.14	2595
LUM	1	8.93	570	9.94	410	4.09	1243	4.14	984
	2	8.01	572	9.30	2287	4.14	1106	4.16	5115
	3	7.80	210	8.46	747	4.25	386	4.23	1494
Accessibility		Mean Trip Rates and Sample Size (N)							
		Household Trips				Person Trips			
		1995		2006		1995		2006	
		Mean	n	Mean	n	Mean	n	Mean	n
P2EP_DI	1	8.68	81	8.52	165	3.80	185	3.96	355
	2	8.48	142	9.92	501	3.60	334	4.03	1233
	3	8.35	575	9.68	1444	4.12	1164	4.16	3360
	4	8.31	554	8.48	1334	4.37	1052	4.28	2645
E1E_TT	1	9.27	124	9.42	320	3.83	300	3.97	760
	2	8.20	152	9.91	695	3.68	339	4.07	1691
	3	8.38	619	9.44	1401	4.08	1271	4.20	3147
	4	8.15	457	8.31	1028	4.51	825	4.28	1995
E1EP_GC	1	8.82	84	8.96	325	3.92	189	4.02	724
	2	8.24	135	9.88	651	3.57	312	4.09	1574
	3	8.88	302	9.62	1008	4.00	670	4.12	2353
	4	8.15	831	8.65	1460	4.33	1564	4.29	2942

Variable-to-variable comparisons using the Bonferroni multiple-comparison test described previously highlighted significant differences between the mean trip rates for the various strata within each of the final area type and accessibility definitions. An interesting finding from this comparison is the variation that LUM captures in school trip rates, especially for the 2006 data. There is a significant difference between all categories for the 2006 data and between category one and three for the 1995 data. This definition appears to show promise in better explaining school related trip making. Between the three definitions of area type, LUDI best captures differences in HBW trips at the household level, while EmpD better explains differences at the person level. For the accessibility definitions, E1E_TT is best for explaining differences in HBW and HBO person trips, while P2EP_DI is best for HBSC household and person trips, but only for the 2006 data. In general, this analysis shows that area type and accessibility are better at capturing differences in trips by trip purpose at the person level than at the household level.

5.6.2 Temporal Effects

The goal of the temporal analysis is to acquire a greater understanding of how area type and accessibility influence trip making over time in order to inform the hypothesis that the inclusion of these variables in trip generation models will contribute to the temporal stability of travel forecasts. A t-statistic value less than 1.96 suggests that there is not enough evidence to reject the null hypothesis and thus we can infer temporal stability between the means. Table 20 provides a summary of results. In comparison to the total household and person trip rates, stratification by area type and accessibility improves the temporal stability of the calculated household trip rates, but does not appear to offer much in the way of temporal stability for person trips. Total person trips and HBSH trips are stable for multiple stratifications of both area type and accessibility, but this benefit simply mirrors the stability seen in the overall trip rates. HBSC person trip rates are more stable for LUM and for accessibility measures of E1E_TT and E1EP_GC. In comparison, the stability of household trip rates improves with both area type and accessibility stratification, most notably for HBSC trips stratified by LUM.

Table 20. Results of Mean Trip Rate Temporal Analysis between 1995 and 2006

		Calculated t-statistic									
Area Type Variables		Household Trips					Person Trips				
		Total	HBW	HBSH	HBSC	HBO	Total	HBW	HBSH	HBSC	HBO
EmpD	1	1.85	-10.19	2.10	-15.36	5.48	-1.30	-16.74	1.41	-16.19	7.48
	2	6.73	-5.50	2.76	-0.48	9.66	1.55	-11.57	0.25	-3.88	10.08
	3	1.22	-2.07	0.84	-2.89	3.77	-1.96	-4.40	-0.23	-4.79	3.98
LUDI	1	2.74	-10.45	1.97	-3.58	6.20	5.23	-12.00	2.37	-5.69	9.48
	2	4.66	-3.99	-0.08	-3.68	7.89	1.36	-7.79	-2.57	-7.11	8.91
	3	4.61	-0.54	3.45	-0.76	5.82	-4.56	-7.03	1.20	-4.67	3.48
LUM	1	7.47	-5.15	4.06	-1.39	73.52	1.25	-9.58	0.64	-4.25	28.89
	2	5.54	-4.62	2.34	-1.56	8.02	0.27	-10.13	-0.16	-5.02	7.98
	3	1.95	-1.00	1.63	-0.51	2.91	-0.16	-2.96	1.12	-1.57	2.69
Accessibility Variables		Household Trips					Person Trips				
		Total	HBW	HBSH	HBSC	HBO	Total	HBW	HBSH	HBSC	HBO
P2EP_DI	1	-0.36	-2.87	0.26	-8.79	1.44	1.40	-2.79	1.13	-13.49	3.00
	2	3.05	-4.08	2.45	-1.49	4.69	4.55	-5.08	2.74	-2.48	6.67
	3	6.26	-1.82	2.19	-1.94	8.28	0.66	-7.22	-0.44	-5.90	7.76
	4	0.92	-6.87	0.80	-3.99	5.65	-1.79	-9.76	-0.05	-6.63	7.14
E1E_TT	1	0.37	-2.28	2.19	-3.92	1.41	1.56	-2.27	2.88	-5.08	2.51
	2	3.46	-4.45	1.27	-0.18	5.82	3.71	-6.13	0.64	-1.16	7.71
	3	5.54	-2.97	3.41	-4.31	7.55	2.49	-6.80	2.45	-8.25	7.85
	4	0.80	-5.76	-1.88	-3.42	5.63	-3.88	-9.69	-4.17	-6.43	6.53
E1EP_GC	1	0.27	5.83	2.00	-3.29	1.60	0.78	-2.42	2.61	-4.41	2.29
	2	3.07	8.39	0.69	-0.64	4.68	4.60	-4.31	0.40	-1.31	6.67
	3	2.46	12.66	1.10	-2.99	5.52	1.65	-4.66	0.52	-4.83	6.93
	4	3.50	19.37	2.34	-6.38	8.00	-0.99	-12.04	0.88	-11.75	9.34
TOTAL		6.15	-7.45	2.95	-4.79	11.3	1.05	-3.24	0.66	-9.73	12.6

For shaded cells, cannot reject H_0 at $t<1.96$, evidence of temporal stability; [1] $df < 39$, appropriate t-value based on lookup table

5.7 Summary and Conclusions

This paper investigated several measures of area type and accessibility to assess the value of each in explaining trip making behavior between strata, trip purpose, and temporal variability. The analysis considered six area types and 24 accessibility definitions from previous research. ANOVA was used to evaluate the performance of the various definitions in explaining differences in travel behavior. Eta-square provided insights into the strength of the difference between the means for each definition and variable-to-variable comparisons were made using the Bonferroni multiple-comparison test. Considering only a p-value less than 0.05 for a 95% confidence level, statistical tests showed that several area type and accessibility definitions can be useful in explaining differences in trip making behavior. The exponential function, $e^{-C_{ij}}$, with employment as the measure of attraction and travel time as the measure of impedance (E1E_TT) showed good variability in trip rates for household and person trip purposes in both analysis years. The exponential function, $e^{-0.1C_{ij}}$, with employment as the measure of attraction and distance as the measure of impedance (EE_DI) was the poorest performer in capturing trip rate variation for either person trips or household trips. The power exponential function, $C_{ij}^{0.5} * e^{-0.1C_{ij}}$, with employment plus population as the measure of attraction and distance as the measure of impedance (PEEP_DI) performed well for the 1995 data capturing variation in total trips, and all other trips types except work trips, but captured no variation in trip rates for the 2006 data. The power function, C_{ij}^{-2}, with employment plus population as the measure of attraction and distance as the measure of impedance (P2EP_DI); and exponential function, $e^{-C_{ij}}$, with employment plus population as the measure of attraction and generalized cost as the measure of impedance (E1EP_GC) performed equally well. Both definitions captured variations in household and person trips by trip purpose across analysis years with the exception of 1995 household trips for P2EP_DI and 1995 person trips for E1EP_GC.

Area type definitions were found to be effective in explaining differences in both household and person total trips, while only a few perform well for trips by purpose, and this often varied by analysis year. While a definition that performs well in one year is worthy of investigation, capturing the trip rate variability in separate analysis years provides evidence that the results are not specific to a given data set, but associated with the underlying behavior. The residential density (ResD) and population density (PopD) definitions show no variability in trip rates in the 1995 household trips by purpose, and very little variation in the 1995 person data. Composite density (CmpD) is just the opposite, with no variability captured in the 1995 person trips by purpose and variability shown only for the 1995 household work trips. Employment density (EmpD) captures variability in both analysis years for household and person trips not classified as work, shopping or school, person work trips only, and household school trips only. The land use density index (LUDI) captures variability in both analysis years for household work and other trips, whether “other” refers to trips not classified as work, shopping or school. Finally, land use mix (LUM) captures variability in household and person school trips and person work trips.

Based on this analysis ResD, PopD, and CmpD are the poorest performers in the group of area type definitions while EE_DI and PEEP_DI are the poorest performers in the accessibility group of definitions.

Considering the eta-square (η^2) statistic, the strength of the difference between the means for the specified strata for any given definition is not very high, meaning that very little of the variability in the dependent variable can be explained by the specified definition. Collapsing the original ten strata did not improve these results.

Evaluating the temporal stability of the most promising area type and accessibility definitions at a 95% confidence interval provides encouraging evidence regarding the stability of the mean trip rates for household trip rates, but does not offer much in the way of stability for person trip rates. While the analysis shows temporal stability for total person trips and shopping person trips, this stability is no better than that seen for trips not stratified by area type or accessibility. The only trip purpose where area type and accessibility appear to offer improved temporal stability of person trip rates for at least one stratum is school trips, specifically for LUM, E1E_TT, and E1EP_GC. LUM provides temporal stability for school household trips for every stratum. Work household trips show temporal stability for the highest density measure of LUDI and LUM, while household trips classified as “other” show temporal stability for all definitions of accessibility at the lowest level of accessibility. The temporal analysis suggests that incorporating these variables into household models could significantly improve the temporal stability of the models, though the results for person trip models appears less promising.

The results of these tests support the use of area type and accessibility as variables for trip generation models, as there is evidence from this research that these variables capture differences in trip making while also providing improved stability in household trip rates over time. Future research in this area should focus on model estimation using area type and accessibility to see how well these variables contribute to improved model fit, performance, and temporal stability. This research effort generally supports the development of trip generation models that allow more independent variables in the model specification, thereby capturing more of the factors that influence travel behavior.

CHAPTER 6. GENERATION CHOICE MODELS

6.1 Introduction and Motivation

Transportation plays a significant role in the mobility, economic health, and quality of life of our communities. It shapes growth by providing access to land and shapes public policy in areas related to "air quality, environmental resource consumption, social equality, land use, urban growth, economic development safety, and security" (FHWA 2007). The transportation planning process is a complex process of "developing strategies for operating, managing, maintaining, and financing the area's transportation system in such a way as to advance the area's long-term goals" (FHWA 2007). Planning is inherently a public process that connects us to the future. In the field of transportation planning, this connection to the future is often made through the development and application of travel demand models that are used to forecast future demand in travel based on forecast input variables related to land use, demographics, and socio-economic factors.

A travel demand model is a series of mathematical equations used to describe travel and travel choices. In its most basic form, this series of models is broken into a 4-step process, trip generation, trip distribution, mode choice, and trip assignment. Best practice models use locally collected travel survey data to estimate and calibrate the models. Planners use these models to forecast travel demand 20 – 30 years into the future for the purposes of evaluating transportation strategies and investment. An implicit assumption of a travel demand model is that model parameters remain stable over time (Ortuzar and Willumsen 2011). A violation of this assumption could lead to transportation analyses and travel forecasts that either over- or underestimate travel demand and associated transportation deficiencies, which could in turn lead to poorly allocated investments in transportation infrastructure.

Trip generation is the first in a sequence of models used to forecast travel demand. As the first step in the model chain, forecasting errors in this step may compound errors in the remaining steps. The trip generation model estimates the total number of trips in the region. The most commonly used form of the trip generation model is the cross-classification model.

This approach estimates the number of trips produced as a function of various household characteristics such as income, household size, or auto ownership (Ortuzar and Willumsen 2011). A less common form of the trip generation model is a generation choice model. This model form has the advantage of allowing more independent variables in the model specification and of capturing person travel as opposed to household travel (PB 2007). Given the opportunity to include more variables that influence individual travel behavior one can hypothesize that a generation choice model is temporally stable.

This paper evaluates the temporal stability of generation choice models based first on widely used explanatory variables, and second on a set of supplemental variables defining life cycle, area type, and accessibility to determine whether temporal stability improves with the inclusion of these variables. In this context, life cycle is used to describe changes that people undergo from infancy, to childhood, to adulthood, to old age as a means of better understanding how these various cycles influence behavior (Zimmerman 1981). Area type describes land use and development type with respect to population and employment, while accessibility captures characteristics of the underlying transportation system.

6.2 Literature Review

6.2.1 Trip Generation Models

Trip generation models are the first step in the traditional 4-step process for travel demand modeling. They provide an estimate of the number of trips generated and attracted to each traffic analysis zone (TAZ) in the study area. Common forms of the trip generation model include regression, cross-classification, and more recently, but seldom used, discrete choice (Ortuzar and Willumsen 2011). Given that trip generation is the first step in the classic model chain, any over- or under-prediction of trips will result in errors in the remaining steps. Well-specified and calibrated models can minimize this error in the base year, but models that fail to capture changes in travel behavior over time can lead to poorly informed decisions related to infrastructure investment, benefits analysis, economic investment decisions, and mobile emissions.

The development of disaggregate travel demand models based on discrete choice analysis is cited in the literature as a major innovation in the analysis of travel demand (Ben-Akiva and Lerman 1985). While most commonly used for mode choice modeling, recent applications of discrete choice analysis have included the development of destination choice models for the trip distribution step of travel demand models (TRB 2007). Even less common is the application of discrete choice analysis for the development of generation choice models. A generation choice model is a logit model that estimates the daily trip frequency for a person rather than household (PB 2007). Person-level trip models offer several advantages over household-based models, including greater compatibility with other components of the modeling system and a greater use of important variables (Ortuzar and Willumsen 2011). A fairly unique model structure, generation choice models offer several advantages (PB 2007):

1. Allow more independent variables to be used
2. Allow continuous independent variables to be used, rather than just classification variables
3. Allow statistical measures to be made measuring the significance of the independent variables for the entire equation

6.2.2 Transferability

The question of transferability has been studied both in spatial and temporal terms. Spatial transferability refers to the practice of applying data or models developed in one geographic region to another geographic region. Temporal stability is concerned with how models developed during one period of time transfer to a future period. The focus of this paper is on temporal stability. The literature indicates a heavy focus on the temporal stability of travel models during the 1970s (see Ashford, 1972; Kannel and Heathington, 1973; Smith and Cleveland, 1976; Yunker, 1976; Doubleday, 1977), while current work in this area has focused on spatial transferability (see Wilmot, 1995; Agyemang-Duah and Hall, 1997; Cotrus, 2005; Mohammadian, 2007; Everett, 2009). The focus on spatial transferability is in response to both the lack of survey data in many urban areas as well as the increased cost for colleting such data (TMIP 2004; Everett 2009). While much remains for exploration in the area of spatial transferability, revisiting the issue of temporal stability is equally important. There are increasing constraints on funds available for infrastructure improvements (TRB 2006; Cambridge Systematics 2007). At the same time, growth, especially in urban areas, continues to outpace investment (Cambridge Systematics 2007). The combination of these factors places even more importance on improved analysis tools for better decision making. Furthermore, recommendations from a peer exchange held in December 2004 included recommendations for research related to the temporal stability and dynamics of travel behavior in order to understand how model parameters change over time (TMIP 2004).

Previous research on temporal stability of trip generation models has yielded mixed results. Several studies suggest temporal stability for model coefficients, model application results, or both (Hill and Dodd 1966; Kannel and Heathington 1973; Downes and Gyenes 1976; Yunker 1976; Walker and Peng 1991) while other studies have shown less than favorable results (Ashford and Holloway 1972; Smith and Cleveland 1976; Doubleday 1977; Cotrus, Prashker et al. 2005).

The debate among researchers investigating temporal stability of trip generation models has focused on regression and cross-classification models, with the exception of Cotrus et al. who also investigated a Tobit model formulation. The primary variables included in the cross-classification models were household size, income, and auto ownership. Several researchers included area type as a variable (Smith and Cleveland 1976; Walker and Peng 1991), concluding that the inclusion of this variable could improve the temporal stability of trip generation models. Smith et al. and Cotrus et al. included a life cycle variable although findings related to temporal stability were inconclusive. No studies found to date have investigated the temporal stability of generation choice models.

6.3 Methodology

This paper focuses on the development of generation choice models using household travel surveys from two different periods with the goal of evaluating the temporal stability of the models. Model estimation includes the consideration of widely used explanatory variables in a set of models referred to as Case 1 models. A separate estimation exercise then considers variables defining life cycle, area type, and accessibility in addition to the original variables; these models are the Case 2 models. Best-fit models are first estimated using the 1995 data. Next, this model specification is used to estimate models using the 2006 data. The temporal stability of the original 1995 models is evaluated using several measures for evaluating temporal stability including both statistical tests and measures of how well the models predict future observed trip making behavior.

6.3.1 Data

The datasets supporting this research come from activity based household travel surveys administered in the Research Triangle Region of North Carolina in 1995 and 2006. Specifically, these data include the 1994/1995 Triangle Travel Behavior Survey (NuStats 1995), and the 2006 Greater Triangle Travel Survey (NuStats 2006). The Triangle Region describes the geographic area covered by the Capital Area Metropolitan Planning Organization (CAMPO) and the Durham-Chapel Hill-Carrboro Metropolitan Planning Organization (DCHC MPO). Together the boundaries of the two MPOs cover eight counties with a combined MPO population of 1.4 million.

The estimation process for generation choice models requires an estimation file that includes one record per person for each person surveyed including all testing variables. An example record might include values or dummy variables for the following: Person ID, Income 1, Income 2, Income 3, Income 4, Household Size, Number of Vehicles, etc.

6.3.2 List of Variables

This section defines the list of variables tested during model estimation. The first group describes the explanatory variables used in the estimation of the Case 1 models; these variables are widely used in trip generation models. The second set describes the supplemental life cycle, area type and accessibility variables used in the Case 2 models, which also consider variables in the first list.

Variables Tested in Case 1 Models

- age - Age of respondent
- hhadlts – Number of household adults
- hhsize - Household size
- hhvehgrp - Household vehicle group
- hhwrk – Number of household workers
- incgrp - Income group
- inc1 - Low income flag (< 25k)
- inc2 - Low medium income flag (25k-50k)
- inc3 - Medium high income flag (50k-100k)
- inc4 - High income flag (> 100k)
- less100k - HH income less than 100k flag
- less25k - HH income less than 25k flag
- less50k- HH income less than 50k flag
- lessveh - Vehicles less than workers flag
- moreveh - Vehicles more than workers flag
- nnwa - Number of non-working adults
- noinc - Unreported income flag
- numveh - Number of vehicles
- otadult - Number of other adults
- pnwa - Presence of a non-working adult flag
- pwrk - Presence of a worker flag
- pchild - Presence of children flag
- totchld - Total children
- veh - Vehicle present flag

Supplemental Variables Tested in Case 2 Models

- c3049na - couple, age of head of household between 30 and 49, no children
- c5059na - couple, age of head of household between 50 and 59, no children at home
- c60na - couple, age of head of household over 60, no children at home

- caa06 - couple, any age, oldest child between 0 and 6
- caa1115 - couple, any age, oldest child between 11 and 15
- caa1618 - couple, any age, oldest child between 16 and 18
- caa1922 - couple, any age, oldest child between 19 and 22
- caa710 - couple, any age, oldest child between 7 and 10
- e1ett - Accessibility exponential function where the measure of attraction is employment and the measure of impedance is travel time
- e1e1 - Categorical variable for previously defined accessibility function, value is 0 to 15
- e1e2 - Categorical variable for previously defined accessibility function, value is 16 to 102
- e1e3 - Categorical variable for previously defined accessibility function, value is 103 to 647
- e1e4 - Categorical variable for previously defined accessibility function, value is 648 and above
- empd1 - Units are employees per square mile where value is 0 to 29
- empd2 - Units are employees per square mile where value is 30 to 680
- empd3 - Units are employees per square mile where value is 681 and above
- lud1 - Units are households plus employees per square mile where value is 0 to 712
- lud2 - Units are households plus employees per square mile where value is 713 to 3161
- lud3 - Units are households plus employees per square mile where value is 3162 and above
- lum1 - Units are land use mix index ranging from 0 (no mix) to 1 (perfect mix) where value is 0 to 0.16
- lum2 - Units are land use mix index ranging from 0 (no mix) to 1 (perfect mix) where value is 0.16 to 0.59
- lum3 - Units are land use mix index ranging from 0 (no mix) to 1 (perfect mix) where value is 0.59 to 1.00

- nc06 - not coupled, oldest child between 0 and 6
- nc1115 - not coupled, oldest child between 11 and 15
- nc1618 - not coupled, oldest child between 16 and 18
- nc1922 - not coupled, oldest child between 19 and 22
- nc30na - not coupled, head of household 30 or younger, no children
- nc55na - not coupled, head of household between 30 and 55, no children
- nc69na - not coupled, head of household between 56 and 69, no children
- nc710 - not coupled, oldest child between 7 and 10
- op7079na - not coupled, head of household between 70 and 79, no children at home
- op80na - not coupled, head of household 80 or older, no children at home
- p2epdi - Accessibility power function where the measure of attraction is employment plus population and the measure of impedance is distance
- yc30na – young couple, head of household 30 or younger, no children

6.3.3 Model Estimation

The generation choice model is a logit model formulation that estimates the daily trip frequency (n) by trip purpose (y) that a person (x) will make:

$$P_{x,y}(n) = \frac{e^{U_{x,y}(n)}}{\sum e^{U_{x,y}(n)}} \qquad \text{(Equation 17)}$$

$$U_{x,y}(n) = C_n + \beta_{1n} * var_1 + \beta_{2n} * var_2 + \cdots \beta_{zn} * var_z \qquad \text{(Equation 18)}$$

where:

$P_{x,y}(n)$ = probability of a person of type “x” making “n” daily trips of purpose “y”

$U_{x,y}(n)$ = utility of “n” trips for person type “x” and purpose “y”

C_n = constant for trip alternative “n”

β_{1n}, β_{2n}, … β_{zn}, = estimated coefficients for trip alternative “n” and independent variables var_1, var_1, … var_z

var_1, var_1, … var_z,= model independent variables.

This research explores home-based work (HBW) trips by workers (WRK), and home-based other (HBO) trips by workers, non-working adults (NWA), and children (CHD). Logit models are used to estimate the probability of a worker making 0, 1, 2, or 3+ HBW trips; a worker or non-working adult making 0, 1, 2, 3, or 4+ HBO trips; and a child making 0, 1, or 2+ HBO trips.

The estimation of the models considered logical combinations of the variables described in the previous section for each trip purpose and person type. The 1995 and 2006 survey data were the primary sources for this estimation exercise. Model estimation relied on BIOGEME 1.8, a free software package available for the estimation of discrete choice models (Bierlaire 2008). The estimated trips are daily person trips by trip purpose. The final model selection was based on coefficient reasonableness (sign and magnitude), check of logical relationships with respect to the trip choice, rho-squared values, and t-statistics.

6.3.4 Tests of Temporal Stability

The measure of temporal stability considers how well the estimated models predict observed total trips by trip purpose, how well the estimated models predict observed fractions of trips by trip purpose and stratification, and the difference between the model coefficients for the 1995 and 2006 estimations looking both at the individual parameters and at the model as a whole.

The measure of how well the estimated 1995 models predict observed 2006 total trips by trip purpose considers percent error. Comparisons include percent error summaries of 2006 estimated trips using the 1995 models against the 2006 observed trips:

$$\%E = \frac{|ET_j - OT_j|}{OT_j} * 100 \qquad \text{(Equation 19)}$$

where:

%E = percent error

ET_j = 2006 estimated trips using the 1995 model
OT_j = 2006 observed trips.

The measure of how well the estimated 1995 models predict observed 2006 fractions of trips by trip purpose and stratification considers the weighted root mean square error (WRMSE) and the root of sum of residual squared (RSRS) (Agyemang-Duah and Hall 1997). Both are measures of the aggregate prediction accuracy of the transferred model:

$$WRMSE = \left(\sum_k PS_k * \left(\frac{PS_k - OS_k}{PS_k}\right)^2\right)^{.5} \qquad \text{(Equation 20)}$$

and

$$RSRS = \left(\sum_k (PS_k - OS_k)^2\right)^{.5} \qquad \text{(Equation 21)}$$

where:
k = number of trip alternatives by trip purpose (0, 1, 2, …, n)
WRMSE = weighted root mean square error
RSRS = root of sum of residual squared
PS_k = predicted fraction for trips k
OS_k = observed fraction for trips k.

The t-test is used to evaluate the difference between the estimated parameters for the 1995 and 2006 model estimations:

$$t = \frac{\alpha_{xi} - \alpha_{xi}}{(var(\alpha_{xi}) + var(\alpha_{xj}))^{0.5}} \qquad \text{(Equation 22)}$$

where:
t = t-test statistic

α_{xi} = estimated coefficient for variable x in 1995 context
α_{xj} = estimated coefficient for variable x in 2006 context
var(α_{xi}) = variance of α_{xi}
var(α_{xj}) = variance of α_{xj}.

Two statistical tests are employed to measure the temporal stability of the models as a whole, one considering the forecast probabilities and the other considering the model coefficients as a whole (Agyemang-Duah and Hall 1997).

The measure of stability of the trips by stratification uses the aggregate prediction statistic to test the hypothesis that the observed choice probabilities in 2006 are given by the 1995 model (Agyemang-Duah and Hall 1997):

$$APS = \sum_k \frac{(PS_k - OS_k)^2}{PS_k} \qquad \text{(Equation 23)}$$

where:
APS = aggregate prediction statistic.

This statistic is equivalent to the chi-square distribution with degrees of freedom equal to the number of alternatives – 1(Agyemang-Duah and Hall 1997).

The measure of stability of the model coefficients uses the Transferability Test Statistic (TTS) (Koppelman and Wilmot 1982) to test the hypothesis that the coefficients are equal. The TTS is defined as (Koppelman and Wilmot 1982):

$$TTS_i(\beta_j) = -2[LL_i(\beta_j) - LL_i(\beta_i)] \qquad \text{(Equation 24)}$$

where:

TTS_i (β_j) = measure of transferability for model estimated in 1995 for application in 2006

LL_i (β_j) = log-likelihood that the behavior observed in 2006 was generated by the model estimated in 1995

LLi (β_i) = log-likelihood for a model estimated in 2006.

The TTS has a chi-square distribution with degrees of freedom equal to the number of parameters (Koppelman and Wilmot 1982), under the null hypothesis that the true parameters governing the 2006 data are the values estimated from the 1995 data.

6.4 Analysis and Results

6.4.1 Model Specification

The model specification process used 1995 survey data focused first on the development of models testing an initial set of variables widely used in trip generation models. These models are referred to as the Case 1 models. HBW trip models are specified for workers only and estimate the probability of a worker making 0, 1, 2, or 3+ HBW trips. HBO trip models are specified for workers, non-working adults, and children. These models estimate the probability of a worker or non-working adult making 0, 1, 2, 3, or 4+ HBO trips, and a child making 0, 1, or 2+ HBO trips. Following the estimation of the Case 1 models, additional models were estimated testing the initial variables supplemented by variables defining life cycle, area type, and accessibility. These models are the Case 2 models. The goal is to evaluate temporal stability with and without the supplemental variables to determine whether life cycle, area type, and accessibility improve temporal stability. After selecting the 1995 Case 1 and 2 models, the same model specifications were estimated using 2006 survey data. The results are reported in Tables 21 and 22.

6.4.2 Model Estimation Results

The base case for each model is zero trips, and the coefficients for all variables are relative to this base case. The variables in the model work in concert with one another, and as such, variable coefficients cannot be considered in isolation. Considering the statistics for the 1995 models, the rho-squared values are within the typical range for models of this type and coefficients have the correct sign with most being significant at the 95% level. Coefficients below the specified significance were retained in the model because of the interaction with other more significant variables, or because of their importance in explaining the model beyond basic statistics. With the exception of the HBW-WRK models, rho-squared values for the 2006 models are generally lower than that of the 1995 models. This variation in goodness of fit is a result of using the 1995 best-fit specification for the 2006 model estimation, rather than letting the data and estimation process lead to best fit models. There are several instances where the sign of the coefficient changes between the 1995 and 2006 models. This difference is a consequence of differences in sample size and the demographic breakdown of the survey sample. In all but one case, the variables where the sign changed are not significant in the 2006 model.

The inclusion of the supplemental variables in the Case 2 models yields mixed results. None of the accessibility variables considered was significant. The area type variable reflecting a high mix of land use is a significant variable for non-working adults making a HBO trip. The life cycle variable was by far the best performing supplemental variable for explaining trip behavior. Various life cycle variables were significant for all trip purposes. For workers, being part of a couple with no children where the age of the head of household is between 30 and 49 has a positive effect on HBW and HBO trip making. This variable did not show up as significant for non-working adults or children. For non-working adults being a part of a couple with children of any age has a significant effect on making 4 or more HBO trips; for workers this shows up only if the worker is part of a couple with children between the ages of 7 and 10. This reflects the role having children plays on making HBO trips. The estimation results show that having older children in the home increases HBO trips for workers in a single parent household, reflecting the influence of not having a second adult in the household to share trip making responsibilities. A child is more likely to make a HBO trip if between the ages of 0 and 10 and part of a single parent household, perhaps reflecting the fact that the child must travel more with the parent as it is less likely there is someone at home to care for the child.

Considering the Case 1 model variables, the model estimation results show that workers living in a household where the number of vehicles is less than the number of workers negatively affects HBW trips. This likely reflects the need to combine trips and share available vehicle(s). Workers that are a part of household with more vehicles than workers are more likely to make two work trips. Having sufficient vehicles available frees the worker to make direct trips from home to work. The presence of children in the household and the presence of other adults detract from the probability of a worker making a work trip, perhaps reflecting the need to stay home to care for a sick child and other adults who could make the work trip. It may also capture the complexity of inter-household trip scheduling.

Having other adults in the household has a negative effect on workers making HBO trips, as other adults are available for sharing these responsibilities. As the number of children in the household increases, so does the probability that a worker will make four or more HBO trips. The number of household vehicles also increases the probability of a worker making a higher number of HBO trips.

For non-working adults, higher income has a positive impact on the number of HBO trips made, as does the number of children in the household. As with the worker HBO trips, the presence of other adults in the household detracts from the probability of a non-worker making HBO trips, reflecting the sharing of trip making responsibilities in multiple adult households. A child is less likely to make a HBO trip as the number of non-working adults in the household increases, and if part of a low-income household.

Table 21. Model Estimation Results for Case 1 Models

Variable	1995 Parameter	1995 t-statistic	2006 Parameter	2006 t-statistic
HBW Worker				
inc2_3	*0.339*	1.86	***0.633***	4.08
lessveh_1	***-0.670***	-2.38	***-0.555***	-3.18
moreveh_2	*0.219*	1.10	*0.034*	0.27
otadult_1	*-0.134*	-1.62	***-0.157***	-3.02
pchild_1	***-0.305***	-2.56	*0.055*	0.75
Log likelihood				
At zero	-2656.14		-6627.87	
At constant	-2312.53		-5711.00	
At convergence	-2300.24		-5690.53	
Rho2	0.134		0.141	
HBO Worker				
nchild_4	***0.412***	4.31	***0.312***	6.86
otadult_1	*-0.005*	-0.05	***-0.137***	-2.48
otadult_2	*-0.071*	-0.8	*-0.089*	-1.67
vehgrp_3	***0.292***	1.96	*0.001*	0.01
Log likelihood				
At zero	-3083.68		-7694.72	
At constant	-2272.18		-6424.84	
At convergence	-2261.16		-6397.36	
Rho2	0.267		0.169	
HBO Non-Working Adult				
inc3_2	*0.400*	1.90	*-0.024*	-0.21
inc3_3	***0.642***	1.98	*-0.070*	-0.34
nchild_4	***0.465***	4.15	***0.301***	4.67
otadult_1	***-0.598***	-3.19	***-0.389***	-4.15
otadult_2	***-0.533***	-3.84	***-0.285***	-4.05
otadult_4	***-0.756***	-4.00	***-0.376***	-4.07
Log likelihood				
At zero	-975.32		-2697.42	
At constant	-877.64		-2506.14	
At convergence	-854.21		-2481.35	
Rho2	0.124		0.080	

Table 21 (continued)

Variable	1995 Parameter	1995 t-statistic	2006 Parameter	2006 t-statistic
HBO Child				
inc1_2	***-1.590***	-2.99	***-0.665***	-2.94
nnwa_1	***-0.704***	-2.51	***-0.510***	-3.53
Log likelihood				
At zero	-532.83		-1859.95	
At constant	-470.13		-1681.83	
At convergence	-459.88		-1670.74	
Rho2	0.137		0.102	

Bold coefficients are significant at the 95% confidence level, _x indicates the trip number (0, 1, 2, …, n)

Table 22. Model Estimation Results for Case 2 Models

Variable	1995 Parameter	1995 t-statistic	2006 Parameter	2006 t-statistic
HBW Worker				
inc1_3	***-0.637***	-1.98	*-0.323*	-0.98
lessveh_1	***-0.605***	-2.38	***-0.605***	-3.87
c3049na_1	***0.485***	2.56	*0.115*	0.80
c3049na_2	***0.405***	2.43	***0.337***	2.80
Log likelihood				
At zero	-2656.14		-6627.87	
At constant	-2312.53		-5711.00	
At convergence	-2303.07		-5697.70	
Rho2	0.133		0.14	
HBO Worker				
c3049na_3	***0.61***	2.41	***-0.58***	-2.13
caa710_4	***1.27***	4.4	***0.668***	3.71
nc1115_2	***1.14***	2.96	*-0.198*	-0.6
nc30na_1	***0.46***	2.25	*0.318*	1.53
Log likelihood				
At zero	-3083.68		-7694.72	
At constant	-2272.18		-6424.84	
At convergence	-2255.20		-6414.93	
Rho2	0.269		0.166	

Table 22 (continued)

Variable	1995 Parameter	1995 t-statistic	2006 Parameter	2006 t-statistic
HBO Non-Working Adult				
inc2_3	***-1.010***	-2.08	*0.032*	0.14
caa06_4	***1.530***	3.93	***1.080***	4.92
caa1115_4	***1.050***	2.25	***1.010***	3.53
caa710_4	***1.520***	3.62	***1.810***	6.35
lum3_1	***0.883***	2.69	*-0.172*	-0.95
lum3_2	***0.532***	1.96	*-0.227*	-1.59
otadult_1	***-0.477***	-2.59	***-0.375***	-4.1
otadult_2	***-0.384***	-2.95	***-0.276***	-4.04
otadult_3	*-0.103*	-0.53	***-0.352***	-2.85
Log likelihood				
At zero	-975.32		-2697.42	
At constant	-877.64		-2506.14	
At convergence	-849.99		-2459.48	
Rho2	0.129		0.088	
HBO Child				
inc1_2	***-1.530***	-2.86	***-0.611***	-2.69
nc06_1	***2.060***	2.64	*-0.619*	-1.02
nc710_1	***1.470***	2.87	***0.872***	2.5
Log likelihood				
At zero	-532.83		-1859.95	
At constant	-470.13		-1681.83	
At convergence	-456.99		-1674.26	
Rho2	0.142		0.100	

Bold coefficients are significant at the 95% confidence level, _x indicates the trip number (0, 1, 2, …, n)

6.4.3 Temporal Stability

Table 23 summarizes the statistical measures of temporal stability for the 1995 Case 1 models as applied to 2006 data. The relative and percent error shows that the 1995 Case 1 models are most useful for predicting 2006 HBW-WRK trips and HBO-NWA trips. The 1995 models perform poorly when predicting HBO-WRK and HBO-CHD trips. Considering the measures of aggregate prediction power for the 1995 Case 1 models, the WRMSE indicates that HBW-WRK models perform best, while the HBO-CHD model again performs the worst. The RSRS measure leads to the same findings.

Applying the t-test at the 0.05 significance level to the 1995 and 2005 Case 1 models shows parameter equality between the two models in nearly all cases. These results suggest that the 1995 models should perform well for estimating 2006 trip making behavior.

The Aggregate Prediction Statistic (APS) tests the hypothesis that the observed trip making probabilities in the 2006 data are predicted by the 1995 model. At a significance level of 0.05, we fail to reject this hypothesis for all but the HBO-CHD model. This finding suggests that the models are temporally stable and acceptable for forecasting trip making behavior.

The hypothesis that the set of coefficients in the 1995 Case 1 models are equal to the set of coefficients in the 2006 Case 1 models uses the Transferability Test Statistic (TTS). The results summarized in Table 23 show that we reject this hypothesis at the 0.05 significance level for all models. This finding implies that it is not necessary for the set of coefficients in an initial year model to equal the set of coefficients in a future year model in order for the model to be useful for predicting trip shares, and therefore useful in forecasting future travel behavior, as shown with the APS test.

Table 23. Measures of Temporal Stability for 1995 Case 1 Models Applied to 2006 Socio-economic Data

Trips	Predicted Fraction of Trips	Observed Fraction of Trips	Predicted Trips	Observed Trips
HBW Worker				
0	0.2201	0.2489	0	0
1	0.2351	0.2366	116491	119402
2	0.4636	0.4653	459335	469649
3+	0.0811	0.0493	149077	92341
Total Trips			724903	681392
%E	9.23			
WRMSE	0.1275			
RSRS	0.0429			
APS	**0.0162**			
TTS	31.33			

For bolded cells, H_0 cannot be rejected at the 95% confidence level

Table 23 (continued)

Trips	Predicted Fraction of Trips	Observed Fraction of Trips	Predicted Trips	Observed Trips
HBO Worker				
0	0.5855	0.4964	0	0
1	0.1636	0.1746	81029	88101
2	0.1567	0.1823	155291	183966
3	0.0482	0.0545	71605	82521
4+	0.0461	0.0923	110225	232572
Total Trips			418150	587160
%E	22.69			
WRMSE	0.2563			
RSRS	0.1043			
APS	**0.0657**			
TTS	**19.43**			
HBO Non-Working Adult				
0	0.3908	0.3322	0	0
1	0.1202	0.1456	23807	28151
2	0.2679	0.2951	106087	114096
3	0.0769	0.0631	45691	36627
4+	0.1442	0.1640	137907	156838
Total Trips			313492	335712
%E	2.56			
WRMSE	0.1487			
RSRS	0.0735			
APS	**0.0221**			
TTS	45.60			
HBO Child				
0	0.5629	0.4326	0	0
1	0.1440	0.1261	31547	27096
2+	0.2931	0.4412	166974	243856
Total Trips			198522	270952
%E	23.98			
WRMSE	0.3274			
RSRS	0.1981			
APS	0.1072			
TTS	16.19			

For bolded cells, H_0 cannot be rejected at the 95% confidence level

Table 24 summarizes the measures of temporal stability for the 1995 Case 2 models as applied to 2006 data. When life cycle variables are considered in model specification, the HBO-NWA model performs the best in consideration of the relative and percent error, while the HBW-WRK model performs second best. The HBO-CHD Case 2 model improves slightly as compared to the Case 1 model, but is still the worst model for predicting trips. The HBO-WRK model also performs poorly. WRMSE also points to the HBO-NWA model as having the best aggregate prediction power of all the 1995 Case 2 models. This finding changes when considering the RSRS measure. In this case, the HBW-WRK model performs slightly better than the HBO-NWA model. In all cases, the 1995 HBO-CHD Case 2 model is the poorest performer.

A comparison of the individual parameters for the 1995 and 2006 Case 2 models using the t-test at the 0.05 significance level indicates greater variability between the 1995 and 2006 parameters than that seen with the Case 1 models. The difference occurs for land use mix and several life cycle variables. All such instances are cases where the sign of the coefficient changed between the two model estimation years. As noted previously, this change in sign is likely due to differences in sample size and the demographic distribution of the survey sample.

The hypothesis that the observed trip making probabilities in the 2006 data are given by the 1995 Case 2 models cannot be rejected at the 0.05 significances level for all but the HBO-CHD model. As with the Case 1 models, we reject the hypothesis that the set of coefficients in the 1995 Case 2 models are equal to the set of coefficients in the 2006 Case 2 models at the 0.05 significance level for all except the HBW-WRK model.

Table 24. Measures of Temporal Stability for 1995 Case 2 Models Applied to 2006 Socio-economic Data

Trips	Predicted Fraction of Trips	Observed Fraction of Trips	Predicted Trips	Observed Trips
HBW Worker				
0	0.2082	0.2489	0	0
1	0.2215	0.2366	109730	119402
2	0.4910	0.4653	486482	469649
3+	0.0793	0.0493	145823	92341
Total Trips			742035	681392
%E	11.81			
WRMSE	0.1474			
RSRS	0.0587			
APS	**0.0217**			
TTS	**0.20**			
HBO Worker				
0	0.5882	0.4964	0	0
1	0.1611	0.1746	79831	88101
2	0.1608	0.1823	159323	183966
3	0.0473	0.0545	70295	82521
4+	0.0426	0.0923	101894	232572
Total Trips			411343	587160
%E	23.95			
WRMSE	0.2783			
RSRS	0.1077			
APS	**0.0775**			
TTS	124.02			
HBO Non-Working Adult				
0	0.3839	0.3322	0	0
1	0.1292	0.1456	25575	28151
2	0.2739	0.2951	108476	114096
3	0.0714	0.0631	42402	36627
4+	0.1417	0.1640	135491	156838
Total Trips			311943	335712
%E	2.05			
WRMSE	0.1231			
RSRS	0.0629			
APS	**0.0152**			
TTS	87.47			

For bolded cells, H_0 cannot be rejected at the 95% confidence level

Table 24 (continued)

Trips	Predicted Fraction of Trips	Observed Fraction of Trips	Predicted Trips	Observed Trips
HBO Child				
0	0.5535	0.4326	0	0
1	0.1568	0.1261	34364	27096
2+	0.2896	0.4412	164987	243856
Total Trips			199351	270952
%E	23.68			
WRMSE	0.3343			
RSRS	0.1963			
APS	0.1118			
TTS	56.42			

For bolded cells, H_0 cannot be rejected at the 95% confidence level

Considering the APS test, all models pass the test for temporal stability with the exception of the HBO-CHD model. Based on the temporal stability tests employed in this analysis the HBO-CHD model is the least temporally stable, whether considering the supplemental variables or not. This finding points to the need for more research in this area to improve our understanding of trip making by children.

The most temporally stable model overall considering the relative and percent error and the WRMSE is the HBO-NWA Case 2 model that includes variables related to life cycle and area type. This analysis also shows this model to be stable when considering the RSRS values and the APS statistic. It fails the stability test however when considering the TTS value. The HBW-WRK Case 2 model is the only model that is temporally stable considering all measures. In this particular case, the inclusion of life cycle variables appears to influence the temporal stability of the model.

In addition to the HBO-CHD model, the HBO-WRK performed poorly in the temporal stability tests, with no improvement in temporal stability with the inclusion of supplemental variables. The APS test suggests we cannot reject the hypothesis of temporal stability when considering the fraction of trips, but the percent error and WRMSE values are high.

6.5 Summary and Conclusions

The purpose of this investigation has been to evaluate the temporal stability of generation choice models, and whether temporal stability improves with the inclusion of supplemental variables related to life cycle, area type, and accessibility. Models were specified and coefficients estimated using an activity-based household travel survey conducted in 1995. The final model selection was based on coefficient reasonableness, check of logical relationships with respect to trip choice, rho-squared values, and t-statistics. Models were first developed testing widely used trip generation variables (Case 1 models), and then a second set of models were developed considering commonly used variables supplemented with variables defining life cycle, area type, and accessibility (Case 2 models). After selecting the 1995 Case 1 and Case 2 models, the same model specifications were estimated using an activity-based household travel survey conducted in 2006.

The model estimation results for the Case 1 models show that income, vehicle ownership, children, other adults, and non-working adults influence trip making. The Case 2 models also support the importance of income, vehicle ownership, and other adults in trip making decisions. The only Case 2 supplemental variables that passed the significance tests were life cycle, for all models, and area type for the non-working adult HBO trip model only. From this study, life cycle appears to be the most promising supplemental variable.

The measures of temporal stability considered in this study include an analysis of how well the estimated models predict observed total trips by trip purpose, how well the estimated models predict observed trip fractions by trip purpose and stratification, a comparison of individual model parameters, and an evaluation of the model parameters in total. The results of the temporal stability tests suggest temporal stability for several of the models tested.

The Case 2 home-based work worker (HBW-WRK) model is stable considering all measures of temporal stability. Both Case 1 and Case 2 HBO non-working adult (HBO-NWA) and Case 1 HBW-WRK trip models are temporally stable considering all measures except the

Transferability Test Statistic (TTS). While the hypothesis that the 2006 observed trip fractions are given by the 1995 model cannot be rejected for the Case 1 and 2 HBO-WRK trip models, the percent error and WRMSE are high suggesting that these models are not temporally stable. The least temporally stable models considering all measures are the Case 1 and 2 HBO models for children. This finding suggests that more work is needed to understand travel by children.

The encouraging finding from this research is that generation choice models, with their ability to estimate person trips and accommodate more variables that define the traveler and the trip, are temporally stable using various measures. With respect to the secondary hypothesis of life cycle, area type, and accessibility variables, this particular analysis shows that while these variables help explain travel behavior, they provide little, if any, additional benefit with respect to temporal stability.

The promising results from this study suggest that generation choice models may also be good candidate models for spatial transferability. The hypothesis of the spatial transferability of generation choice models should be tested using survey data from similar periods, but different geographic regions. Another area for future research deals with the issue of backwards transferability. This study considered the application of models developed in the 1995 context to observed data in the 2006 context. Investigating backwards transferability with models estimated in the 2006 context to observed data in the 1995 context would provide interesting insights, especially with respect to the role of survey data. Finally, this study considers only HBW and HBO trips. An exploration of home-based shopping and home-based school trips could answer the question of whether generation choice models are temporally stable for these trip purposes, and whether there are particular markets where additional research would benefit our understanding of individual travel behavior.

CHAPTER 7. CUMULATIVE LOGISTIC REGRESSION MODELS

7.1 Introduction and Motivation

Trip generation models have taken many forms over the years, including zonal regression models, household regression models, and cross-classification models. Regression models have the advantage of allowing the analyst to consider multiple independent variables, but the disadvantage of treating trip rates as continuous rather than discrete. Cross-classification models overcome this limitation, but introduce another shortcoming with respect to the number of variables and stratifications considered before violating the minimum sample size requirements (about 30 samples per stratification), or conversely making the survey sample size prohibitively expensive. Another disadvantage of cross-classification is the lack of goodness of fit measures. These limitations led the author to consider the cumulative logistic regression model, also known as ordered logistic regression, for estimating trips. The cumulative logistic regression model, often used in the social sciences, estimates the relationship between an ordered categorical dependent variable and a set of independent variables. The cumulative logistic regression model applied to trip generation provides for the estimation of discrete trips and allows for the consideration of multiple independent variables. In addition to allowing multiple independent variables, this model can capture trips made by an individual as opposed to trips made by the household.

This investigation evaluates the usefulness of the cumulative logistic regression model for estimating trip generation. In addition to estimating a set of models based on widely used explanatory variables, subsequent model estimation also considers variables capturing life cycle, area type, and accessibility in order to evaluate whether these variables improve model fit and performance.

A secondary focus of this investigation is on the issue of temporal stability. Temporal stability is concerned with how models developed during one period of time transfer to a future period. Given the stated benefits of cumulative logistic regression models, and the relative ease of including independent variables, the author evaluates temporal stability using several measures, and assesses whether the inclusion of life cycle, area type, and accessibility variables result in improved stability.

7.2 Literature Review

7.2.1 Trip Generation Models

Early travel forecasts consisted primarily of the extrapolation of "desire lines" developed from origin-destination surveys (FHWA 1975). This practice advanced in the early 1950s to consider land use and socio-economic factors in quantifying urban trip volumes, and it provided an analytical approach for using future land use plans to estimate future travel demand (FHWA 1975). Regression models of trip generation became commonplace in the late 1950s and early 1960s, opening the door for a greater insight into travel and the factors influencing it (FHWA 1975). The 1970s marked a shift away from aggregate zonal level regression analysis to household cross-classification procedures.

Cross-classification models estimate an average number of trips as a function of two or more household attributes (Ortuzar and Willumsen 2011). This method has long been the most established model for estimating trips in a travel demand model. However, new model forms are becoming more common in the toolbox of models considered for trip generation. These disaggregate models based on discrete choice analysis are considered by some to be a major innovation in the field (Ben-Akiva and Lerman 1985). Discrete choice analysis is commonly used for modeling mode choice, but recent applications have also considered destination choice, and even more recently generation choice. Generation choice models estimate the frequency of daily person trips. Models that estimate person trips are an improvement over household based models as they allow for a greater use of important variables and are more compatible with other components of the modeling system (Ortuzar and Willumsen 2011). In addition to choice-based models, another form of disaggregate model for trip generation is the cumulative logistic regression model. Cumulative logistic regression models, also known as ordered logistic regression, estimate relationships between an ordered categorical dependent variable and a set of independent variables.

Disaggregate trip generation models offer several advantages over the commonly used cross-classification model, including the flexibility to consider more independent variables, the ability to include continuous variables in addition to classification variables, and statistical measures for evaluating the significance of the independent variables. Also, unlike the cross-classification model, where sample size quickly limits the number of stratifications due to the requirement that any given cell have at least 30 observations, a disaggregate model can capture multiple variables, making it possible to capture relationships that are not possible with the standard cross-classification approach (PB 2007). A study in the Greater Toronto Area successfully used cumulative logistic regression to develop a model for shopping trips (Agyemang-Duah and Hall 1997). This study included explanatory variables for household size, worker status, number of children, vehicle ownership, and accessibility. Family life cycle was not included.

7.2.2 Temporal Stability

In application, trip generation models developed for a base year condition are used to forecast trips for a future year condition. This process assumes the temporal stability of the model and the parameters. Evaluating temporal stability is a process of assessing how well base year models actually reflect a future year condition. Early researchers focused heavily on this topic during the 1970s (see Ashford, 1972; Kannel and Heathington, 1973; Smith and Cleveland, 1976; Yunker, 1976; Doubleday, 1977). These studies focused on regression and cross-classification models, with mixed results regarding temporal stability. Person-level Tobit models were the focus of recent research into the spatial transferability and temporal stability of trip generation models (Cotrus, Prashker et al. 2005). Overall results indicated that the models are not temporally stable, but aggregate measures comparing the total number of trips estimated against observed were temporally stable within acceptable ranges. Unpublished research by Huntsinger and Rouphail investigated the temporal stability of generation choice models (Huntsinger and Rouphail 2011). Using both statistical tests and comparisons of how well estimated trips matched observed trips, these researchers found that home-based work trip models and home-based other worker and non-working adult trip models were temporally stable.

7.3 Methodology

This analysis focuses on the estimation of cumulative logistic regression models using 1995 household survey data and a initial set of widely used explanatory variables such as household size, income, workers per household, and auto ownership. This initial variable set is then expanded to include variables capturing life cycle, area type, and accessibility to evaluate whether these variables improve model fit and model results. The selection of the best-fit 1995 models considers coefficient reasonableness (sign and magnitude), check of logical relationships with respect to the trip choice, chi-squared statistics, pseudo R^2 values, and z-statistics. Model application using socio-economic data for the region results in estimated trips. A comparison of estimated to observed trips provides a measure of model performance.

To evaluate temporal stability, survey data from 2006 is used to estimate new models using the final best-fit specification for the 1995 models. Several measures evaluate the temporal stability of these models, including both statistical tests and measures of how well the models predict 2006 observed trips.

To evaluate backwards temporal stability, best-fit models are also estimated using 2006 survey data and then applied to evaluate how well the 2006 models predict 1995 observed trips. This analysis also provides insight into model estimation using data sets from two different points in time.

7.3.1 Data

Two activity-based household travel surveys administered in the Research Triangle Region of North Carolina form the basis of this analysis. Specifically, these data include the 1994/1995 Triangle Travel Behavior Survey (NuStats 1995), and the 2006 Greater Triangle Travel Survey (NuStats 2006). The Triangle Region describes the geographic area covered by the Capital Area Metropolitan Planning Organization (CAMPO) and the Durham-Chapel

Hill-Carrboro Metropolitan Planning Organization (DCHC MPO). Together the boundaries of the two MPOs cover eight counties with a combined MPO population of 1.4 million.

The estimation process requires an estimation file that includes one record for each person surveyed. An example record might include values or dummy variables for the following: Person ID, Worker, Non-Worker, Child, Income 1, Income 2, Income 3, Income 4, Household Size, Number of Vehicles, etc.

7.3.2 List of Variables

This section defines the list of variables evaluated during model estimation. The first group describes the explanatory variables used in the estimation of the Case 1 models; these variables are widely used in trip generation models. The second set describes supplemental variables defining life cycle, area type, and accessibility. Case 2 models are estimated using both groups of variables.

Variables Tested in Case 1 Models

- Age - Age of respondent
- Child - Child flag (age <= 15)
- hhadlts – Number of household adults
- hhsize - Household size
- hhveh_grp - Household vehicle group
- hhwrk – Number of household workers
- inc_grp - Income group
- inc1 - Low income flag (< 25k)
- inc2 - Low medium income flag (25k-50k)
- inc3 - Medium high income flag (50k-100k)
- inc4 - High income flag (> 100k)
- less100k - HH income less than 100k flag
- less25k - HH income less than 25k flag
- less50k - HH income less than 50k flag
- lessveh - Vehicles less than workers flag

- moreveh - Vehicles more than workers flag
- n_nwa - Number of non-working adults
- noinc - Unreported income flag
- numveh - Number of vehicles
- nwa - Non-working adult flag
- otadult - Number of other adults
- p_nwa - Presence of a non-working adult flag
- p_wrk - Presence of a worker flag
- pchild - Presence of children flag
- totchld - Total children
- veh - Vehicle present flag
- worker - Worker flag

Supplemental Variables Tested in Case 2 Models

- c3049na - couple, age of head of household between 30 and 49, no children
- c5059na - couple, age of head of household between 50 and 59, no children at home
- c60na - couple, age of head of household over 60, no children at home
- caa06 - couple, any age, oldest child between 0 and 6
- caa1115 - couple, any age, oldest child between 11 and 15
- caa1618 - couple, any age, oldest child between 16 and 18
- caa1922 - couple, any age, oldest child between 19 and 22
- caa710 - couple, any age, oldest child between 7 and 10
- e1e_tt - Accessibility exponential function where the measure of attraction is employment and the measure of impedance is travel time
- e1e_1 - Categorical variable for previously defined accessibility function, value is 0 to 15
- e1e_2 - Categorical variable for previously defined accessibility function, value is 16 to 102
- e1e_3 - Categorical variable for previously defined accessibility function, value is 103 to 647

- e1e_4 - Categorical variable for previously defined accessibility function, value is 648 and above
- empd1 - Units are employees per square mile where value is 0 to 29
- empd2 - Units are employees per square mile where value is 30 to 680
- empd3 - Units are employees per square mile where value is 681 and above
- lud1 - Units are households plus employees per square mile where value is 0 to 712
- lud2 - Units are households plus employees per square mile where value is 713 to 3161
- lud3 - Units are households plus employees per square mile where value is 3162 and above
- lum1 - Units are land use mix index ranging from 0 (no mix) to 1 (perfect mix) where value is 0 to 0.16
- lum2 - Units are land use mix index ranging from 0 (no mix) to 1 (perfect mix) where value is 0.16 to 0.59
- lum3 - Units are land use mix index ranging from 0 (no mix) to 1 (perfect mix) where value is 0.59 to 1.00
- nc06 - not coupled, oldest child between 0 and 6
- nc1115 - not coupled, oldest child between 11 and 15
- nc1618 - not coupled, oldest child between 16 and 18
- nc1922 - not coupled, oldest child between 19 and 22
- nc30na - not coupled, head of household 30 or younger, no children
- nc55na - not coupled, head of household between 30 and 55, no children
- nc69na - not coupled, head of household between 56 and 69, no children
- nc710 - not coupled, oldest child between 7 and 10
- op7079na - not coupled, head of household between 70 and 79, no children at home
- op80na - not coupled, head of household 80 or older, no children at home
- p2ep_di - Accessibility power function where the measure of attraction is employment plus population and the measure of impedance is distance
- yc30na – young couple, head of household 30 or younger, no children

7.3.3 Cumulative Logistic Regression Model

While widely used in the social sciences, example applications of the cumulative logistic model are less common in travel modeling literature. The cumulative logistic regression model is a type of discrete choice model that estimates relationships between an ordered dependent variable, for example person trip generation, and a set of independent variables, for example household size, income, and workers.

In the ordered logit, the underlying value is estimated as a linear function of the independent variables (x_{1_j},x_{2_j}, …, x_{k_j}) and a set of cutpoints (κ_i, κ_{i-1}) (StataCorp 2007). The probability of observing outcome i corresponds to the probability that the estimated linear function, ($\beta_1 x_{1_j} + \beta_2 x_{2_j} + \cdots \beta_k x_{k_j}$), plus random error, ($u_j$), is within the range of estimated cutpoints, (κ_i, κ_{i-1}), for the estimated outcome as demonstrated in the equation below (StataCorp 2007).

$$Pr(outcome_j = i) = Pr(\kappa_{i-1} < \beta_1 x_{1_j} + \beta_2 x_{2_j} + \cdots \beta_k x_{k_j} + u_j \leq \kappa_i) \qquad \text{(Equation 25)}$$

The error term, (u_j), is assumed to be logistically distributed, and the coefficients, ($\beta_{1,} \beta_2, \ldots, \beta_k$) , are estimated along with a set of cutpoints, (κ_1, κ_1, …, κ_k), where k is the number of possible outcomes, κ_0 is taken as $-\infty$ and κ_k is taken as $+\infty$ (StataCorp 2007). The coefficients and cutpoints are estimated using maximum likelihood, and no constant appears as it is captured in the cutpoints (StataCorp 2007).

Recognizing the number of trips made by an individual as categorical and ordered makes the cumulative logistic regression model a good choice for estimating the probability for a person (i) to make (n) trips:

$$Prob(T_i = n) = P(\kappa_{n-1} < \beta_1 x_{1_i} + \beta_2 x_{2_i} + \cdots \beta_z x_{z_i} + \varepsilon_i \leq \kappa_n) \qquad \text{(Equation 26)}$$

where:

Prob ($T_i = n$) = probability of person "i" making exactly "n" trips

n = number of trips

T_i = trips for person "i"

$\beta_1, \beta_2, \ldots, \beta_z$ = estimated coefficients for variables $x_1, x_2, \ldots, x_z$

$x_1, x_2, \ldots, x_z$ = model independent variables

ε = random error, assumed to be logistically distributed

κ = estimated cutpoints

z = number of independent variables.

The probability of a given observation is (StataCorp 2007) :

$$Prob(T_i = n) = P(\kappa_{n-1} < \beta x_i + \varepsilon \leq \kappa_n) = \qquad \text{(Equation 27)}$$

$$\frac{1}{1 + exp\,(-\kappa_n + x_i\beta)} - \frac{1}{1 + exp\,(-\kappa_{n-1} + x_i\beta)}$$

Trip based travel demand models consider two primary trip purposes, home-based trips and non-home-based trips. Home-based trips are further disaggregated into home-based work (HBW), home-based shopping (HBSH), home-based school (HBSC), and home-based other (HBO). This research explores home-based work (HBW) trips and home-based other (HBO) trips. Cumulative logistic regression models are used to estimate the probability of an individual making 0, 1, 2, or 3+ HBW trips and the probability of an individual making 0, 1, 2, 3, or 4+ HBO trips.

The estimation of the models considered logical combinations of the variables summarized in the previous section for each trip purpose. The 1995 and 2006 survey data described previously was the primary source for the estimation. Model estimation relied on the statistical software package STATA (StataCorp 2007). The estimated trips are person trips by trip purpose. The final model selection was based on coefficient reasonableness (sign and magnitude), chi-squared statistics, pseudo R^2 values, z-statistics and a check of logical relationships with respect to the trip choice.

7.3.4 Tests of Temporal Stability

The measure of temporal stability considers how well the estimated models predict observed total trips by trip purpose, how well the estimated models predict observed shares by trip purpose and stratification, and the difference between the model coefficients for the 1995 and 2006 estimations looking both at the individual parameters.

The measure of how well the estimated 1995 models predict observed 2006 total trips by trip purpose considers percent errors. Comparisons include percent error summaries of 2006 estimated trips using the 1995 models against the 2006 observed trips:

$$\%E = \frac{|ET_j - OT_j|}{OT_j} * 100 \qquad \text{(Equation 28)}$$

where:

%E = percent error

ET_j = 2006 estimated trips using 1995 model

OT_j = 2006 observed trips.

The measure of how well the estimated 1995 models predict observed 2006 fractions of trips by trip purpose and stratification considers the weighted root mean square error (WRMSE) and the root of sum of residual squared (RSRS) (Agyemang-Duah and Hall 1997). Both are measures of the aggregate prediction accuracy of the transferred model:

$$WRMSE = (\sum_k PS_k * (\frac{PS_k - OS_k}{PS_k})^2)^{.5} \qquad \text{(Equation 29)}$$

and

$$RSRS = (\sum_k (PS_k - OS_k)^2)^{.5} \qquad \text{(Equation 30)}$$

where:

k = number of trip alternatives by trip purpose (0, 1, 2, …, n)

WRMSE = weighted root mean square error

RSRS = root of sum of residual squared

PS_k = predicted fraction for trips k

OS_k = observed fraction for trips k.

The t-test evaluates the difference between the estimated parameters for the 1995 and 2006 models:

$$t = \frac{\alpha_{xi} - \alpha_{xi}}{(var(\alpha_{xi}) + var(\alpha_{xj}))^{0.5}} \quad \text{(Equation 31)}$$

where:

t = t-test statistic

α_{xi} = estimated coefficient for variable x in 1995 context

α_{xj} = estimated coefficient for variable x in 2006 context

var(α_{xi}) = variance of α_{xi}

var(α_{xj}) = variance of α_{xj}.

The measure of stability of the trips by stratification uses the aggregate prediction statistic to test the hypothesis that the observed choice probabilities in 2006 are given by the 1995 model (Agyemang-Duah and Hall 1997):

$$APS = \sum_k \frac{(PS_k - OS_k)^2}{PS_k} \quad \text{(Equation 32)}$$

where:

APS = aggregate prediction statistic.

This statistic is distributed as the chi-square distribution with degrees of freedom equal to the number of alternatives – 1 (Agyemang-Duah and Hall 1997).

7.4 Analysis and Results

7.4.1 Model Specification

The model specification process used 1995 survey data, focusing first on the development of models testing an initial set of variables widely used in trip generation models. These models are referred to as the Case 1 models. HBW trip models estimate the probability of an individual making 0, 1, 2, or 3+ HBW trips. HBO trip models estimate the probability of an individual making 0, 1, 2, 3, or 4+ HBO trips. Secondly, following the estimation exercise, additional models were estimated testing the initial variables supplemented by variables defining life cycle, area type, and accessibility. These models are the Case 2 models. The goal is to evaluate models with and without the supplemental variables to determine whether life cycle, area type, and accessibility improve model performance and temporal stability. Tables 25 and 26 report the results for the Case 1 models, and Tables 27 and 28 the results for the Case 2 models.

The selection of equations for estimation started with the identification of variables that influence the decision process, such as variables related to the individual, the household, the trip type, the land use, and the transportation system. From these variables, specific variable sets were identified for testing based on an evaluation of co-linearity, consideration of binary versus categorical variables, logical groupings of variables a priori, and consideration of previously documented model sets. Models were estimated for each subset of variables and variable selection was refined through trial and error to select best fit.

7.4.2 Model Estimation Results

The outcome measure for the estimated models is trips and the coefficients reflect the relationships that exist between trip making and the variables. The response variable, trips, is ordinal under the assumption that the number of trips has a natural ordering from zero to the maximum number of trips. Considering the statistics for the 1995 models, the probability of getting a likelihood ratio test statistic as extreme as the observed under the null hypothesis (Prob>chi^2) is less than 0.0001 leading us to conclude that at least one of the regression coefficients in the model is not equal to zero. The null hypothesis is that all of the regression coefficients are zero. The Pseudo R^2 provides a measure of how good the model is at predicting the data relative to doing nothing. Interpretation of this statistic should be used with caution as it is often a small value and is not equivalent to the R^2 found in ordinary least squares regression (Acock 2006). The values reported for the 1995 models are similar to statistics reported in previous research (Agyemang-Duah and Hall 1997). The 1995 model coefficients have the correct sign and are significant at the 95% level.

Interpretation of the model coefficients is that for a one-unit increase in the predictor variable (for example, household workers), the response variable (trips) is expected to change by its respective coefficient in the ordered log-odds scale while other variables are held constant (StataCorp 2007). A review of the coefficients in the following tables shows logical relationships with respect to the predictor variable and the response to trip making.

Table 25. Model Estimation Results for HBW Case 1 Models Using 1995 Survey

Coefficients specific to	*1995 Parameter*	*1995 z-statistic*
hhsize	*-0.285*	-6.62
age	*0.026*	10.81
n_nwa	*-0.952*	-10.31
hhwrk	*0.781*	10.49
Cut points specific to		
Trips=1	*1.007*	
Trips=2	*1.758*	
Trips=3+	*4.379*	
Summary Statistics		
No. of Observations	3007	
Chi-square	768.20	
Degrees of freedom	4	
Prob>chi-square	*<0.00001*	
Log likelihood		
At zero	-3429.69	
At convergence	-3045.59	
Pseudo R^2	0.112	

NOTE – All coefficients are significant at the 95% confidence level

Table 26. Model Estimation Results for HBO Case 1 Models Using 1995 Survey

Coefficients specific to	*1995 Parameter*	*1995 z-statistic*
nwa	*1.056*	11.89
hhsize	*0.137*	3.76
hhveh_grp	*0.230*	3.78
hhadlts	*-0.498*	-6.59
inc1	*-0.266*	-2.39
Cut points specific to		
Trips=1	*0.170*	
Trips=2	*0.857*	
Trips=3	*2.120*	
Trips=4+	*2.817*	
Summary Statistics		
No. of Observations	3007	
Chi-square	164.84	
Degrees of freedom	5	
Prob>chi-square	*<0.00001*	
Log likelihood		
At zero	-3797.60	
At convergence	-3715.18	
Pseudo R^2	0.022	

NOTE – All coefficients are significant at the 95% confidence level

The inclusion of the supplemental variables yields mixed results. Accessibility variables are significant for HBO trips, but not for HBW trips. None of the area type variables considered was significant for either trip purpose. Life cycle variables are significant for both trip purposes, with a couple of life cycle variables being significant for both trip purposes. The coefficient for all life cycle variables is positive for the HBW trip models with the exception of the variable for an older person between the ages of 70 to 79.

Table 27. Model Estimation Results for HBW Case 2 Models Using 1995 Survey

Coefficients specific to	*1995 Parameter*	*1995 t-statistic*
c3049na	*0.798*	6.66
caa06	*0.859*	6.42
caa1922	*0.756*	3.27
nc30na	*1.510*	9.38
nc55na	*1.192*	9.07
op7079na	*-1.473*	-3.56
yc30na	*1.504*	8.33
n_nwa	*-1.051*	-12.21
age	*0.042*	16.34
hhwrk	*0.636*	9.16
Cut points specific to		
Trips=1	*2.577*	
Trips=2	*3.371*	
Trips=3+	*6.048*	
Summary Statistics		
No. of Observations	3007	
Chi-square	956.99	
Degrees of freedom	10	
Prob>chi-square	*<0.00001*	
Log likelihood		
At zero	-3429.69	
At convergence	-2951.20	
Pseudo R^2	0.14	

NOTE – All coefficients are significant at the 95% confidence level

Table 28. Model Estimation Results for HBO Case 2 Models Using 1995 Survey

Coefficients specific to	1995 Parameter	1995 t-statistic
nwa	*0.920*	10.59
inc1	*-0.339*	-3.16
e1e1	*0.472*	3.23
e1e3	*0.525*	4.53
e1e4	*0.611*	5.01
nc55na	*-0.262*	-2.14
yc30na	*-0.487*	-2.62
nc1115	*0.560*	2.63
Cut points specific to		
Trips=1	*0.745*	
Trips=2	*1.432*	
Trips=3	*2.695*	
Trips=4+	*3.391*	
Summary Statistics		
No. of Observations	3007	
Chi-square	163.73	
Degrees of freedom	8	
Prob>chi-square	*<0.00001*	
Log likelihood		
At zero	-3797.60	
At convergence	-3715.74	
Pseudo R^2	0.02	

NOTE – All coefficients are significant at the 95% confidence level

7.4.3 Model Verification

Model verification includes an examination of person trips per person, person trips per household, and HBW person trips per worker. Table 29 provides a summary of these measures. The results show a good match between the estimated and observed total trips for both the HBW and HBO trip purposes. Total trip estimations are slightly better for the Case 2 HBO models. The trip rates per person and per household show a good match between estimated and observed. HBW trips rates per worker are also within the typical range of 1.2 to 1.55 cited in the literature (Schiffer and Rossi 2008). The verification measures results between the Case 1 and Case 2 models do not show enough difference to support one over the other. Overall, the results support the cumulative logistic regression model as a good model for trip generation.

Table 29. Model Verification Measures for Models Estimated Using 1995 Survey

Performance Measure	*1995 Case 1 Models*		*1955 Case 2 Models*		*1995 Survey*
HBW Trips/Person		0.90		0.90	0.78
HBO Trips/Person		0.96		0.95	0.78
HBW Trips/Household		2.03		2.03	1.99
HBO Trips/Household		2.15		2.13	2.00
HBW Trips/Worker		1.43		1.42	1.30
			HBW		
Number of Trips	*Estimated Trips*	*%E*	*Estimated Trips*	*%E*	*Observed Trips*
1	90814	14.0%	89439	12.2%	79691
2	360765	7.0%	358093	7.7%	387875
3+	105943	10.9%	108655	13.7%	95526
Total	557522	1.0%	556187	1.2%	563092
			HBO		
Number of Trips	*Estimated Trips*	*%E*	*Estimated Trips*	*%E*	*Observed Trips*
1	93371	0.0%	93059	0.4%	93411
2	229706	3.9%	228074	4.6%	238990
3	95825	15.5%	94700	14.2%	82937
4+	170757	8.8%	168469	7.4%	156930
Total	589659	3.0%	584301	2.1%	572269

7.4.4 Temporal Stability

To evaluate the temporal stability of the 1995 models, the 2006 household survey data was used to estimate 2006 models with the same specification as the 1995 HBW and HBO Case 1 and Case 2 models. Tables 30 and 31 report the results for the Case 1 models, and Tables 32 and 33 the results for the Case 2 models. There are several instances where the sign of the coefficient changes between the 1995 and 2006 models. This difference is likely a consequence of differences in sample size and the demographic breakdown of the survey sample. The statistics of the 2006 models suggest that the models are acceptable.

Table 30. Model Estimation Results for HBW Case 1 Models Using 2006 Survey

Coefficients specific to	2006 Parameter	2006 z-statistic
hhsize	*-0.256*	-9.28
age	*0.025*	19.14
n_nwa	*-0.759*	-14.14
hhwrk	*0.814*	17.9
Cut points specific to		
Trips=1	*1.511*	
Trips=2	*2.269*	
Trips=3+	*5.148*	
Summary Statistics		
No. of Observations	8150	
Chi-square	2037.71	
Degrees of freedom	4	
Prob>chi-square	*<0.00001*	
Log likelihood		
At zero	-8619.79	
At convergence	-7600.93	
Pseudo R^2	0.118	

NOTE – All coefficients are significant at the 95% confidence level

Table 31. Model Estimation Results for HBO Case 1 Models Using 2006 Survey

Coefficients specific to	2006 Parameter	2006 z-statistic
nwa	*0.803*	15.39
hhsize	*0.130*	6.97
hhveh_grp	*0.030*	0.88
hhadlts	*-0.345*	-7.77
inc1	*-0.267*	-3.20
Cut points specific to		
Trips=1	*-0.337*	
Trips=2	*0.335*	
Trips=3	*1.604*	
Trips=4+	*2.081*	
Summary Statistics		
No. of Observations	8150	
Chi-square	273.10	
Degrees of freedom	5	
Prob>chi-square	*<0.00001*	
Log likelihood		
At zero	-11355.05	
At convergence	-11218.50	
Pseudo R^2	0.012	

NOTE – All coefficients are significant at the 95% confidence level

Table 32. Model Estimation Results for HBW Case 2 Models Using 2006 Survey

Coefficients specific to	2006 Parameter	2006 t-statistic
c3049na	*1.238*	13.45
caa06	*0.565*	7.38
caa1922	*0.026*	0.15
nc30na	*1.512*	9.62
nc55na	*1.324*	14.32
op7079na	*-1.226*	-5.47
yc30na	*1.604*	11.31
n_nwa	*-0.839*	-17.41
age	*0.039*	29.4
hhwrk	*0.721*	17.14
Cut points specific to		
Trips=1	*2.940*	
Trips=2	*3.737*	
Trips=3+	*6.673*	
Summary Statistics		
No. of Observations	8150	
Chi-square	2485.72	
Degrees of freedom	10	
Prob>chi-square	*<0.00001*	
Log likelihood		
At zero	-8619.79	
At convergence	-7376.93	
Pseudo R^2	0.14	

NOTE – All coefficients are significant at the 95% confidence level

Table 33. Model Estimation Results for HBO Case 2 Models Using 2006 Survey

Coefficients specific to	2006 Parameter	2006 t-statistic
nwa	*0.689*	13.54
inc1	*-0.220*	-2.77
e1e1	*-0.228*	-2.95
e1e3	*0.027*	0.5
e1e4	*0.068*	1.14
nc55na	*-0.468*	-5.42
yc30na	*-0.290*	-2.08
nc1115	*-0.334*	-2.39
Cut points specific to		
Trips=1	*-0.164*	
Trips=2	*0.506*	
Trips=3	*1.772*	
Trips=4+	*2.248*	
Summary Statistics		
No. of Observations	8150	
Chi-square	246.01	
Degrees of freedom	8	
Prob>chi-square	*<0.00001*	
Log likelihood		
At zero	-11355.05	
At convergence	-11232.05	
Pseudo R^2	0.01	

NOTE – All coefficients are significant at the 95% confidence level

Table 34 summarizes the measures of temporal stability for the 1995 HBW and HBO Case 1 models as applied to 2006 data. Both models have a percent error above ten percent with very little difference between the two. Ideally, this value would be below ten percent. Considering the measures of aggregate prediction power for the 1995 Case 1 models, the WRMSE indicates that HBW model performs best with regard to temporal stability. The RSRS measure leads to the same finding.

Applying the t-test at the 0.05 significance level to the 1995 and 2006 Case 1 models shows equality for all parameters between the two models for the HBW trips suggesting that the 1995 HBW model should perform well for estimating HBW trips in 2006. For the HBO models, two of the five variables do not pass the equality test, non-working adults and number of vehicles per household. The t-test suggests equality between the 1995 and 2006 coefficients for household size, household adults, and income. This finding suggests that the HBO trip model may not be as temporally stable as the HBW trip model, as was noted considering the WRMSE and RSRS measures.

The Aggregate Prediction Statistic (APS) tests the hypothesis that the 2006 observed trip making probabilities are estimated by the 1995 model. At a significance level of 0.05, we fail to reject this hypothesis for both the HBW and HBO models. This finding suggests that the models are temporally stable and acceptable for forecasting trip generation.

Table 34. Measures of Temporal Stability for 1995 Case 1 Models Applied to 2006 Socio-economic Data

Trips	Predicted Fractions	Observed Fractions	Predicted Trips	Observed Trips
HBW				
0	0.5369	0.5847	0	0
1	0.1364	0.1308	124479	119402
2	0.2787	0.2572	508505	469649
3+	0.0480	0.0272	162464	92341
Total Trips			795447	681392
%E	19.86			
WRMSE	0.1231			
RSRS	0.0566			
APS	**0.0151**			
HBO				
0	0.5416	0.4466	0	0
1	0.1529	0.1570	139511	143348
2	0.1904	0.2389	347487	436204
3	0.0534	0.0510	146228	139784
4+	0.0616	0.1064	271679	474488
Total Trips			904905	1193824
%E	18.31			
WRMSE	0.2486			
RSRS	0.1158			
APS	**0.0618**			

For bolded cells, H_0 cannot be rejected at the 95% confidence level

Table 35 summarizes the measures of temporal stability for the 1995 HBW and HBO Case 2 models as applied to 2006 data. When life cycle variables are included in the model specification, the HBW model performs the best in consideration of the percent error, although still above the desired 10 percent. The WRMSE and RSRS values also point to the HBW model as having the best aggregate prediction power for the 1995 Case 2 models.

A comparison of the individual parameters for the 1995 and 2006 Case 2 models using the t-test at the 0.05 significance level shows equality for seven of the ten variables in the model. Not meeting the equality test are the coefficients for the variables representing couples with no kids where the age of the head of household is between 30 and 49, couples of any age with kids between the ages of 19 and 22, and the variable representing the number of non-working adults. This finding suggests temporal stability of the model. As with the Case 1 model, the HBO trip model appears less temporally stable considering only the coefficient equality. In this case, five of the eight variables fail the test of equality, including all of the variables capturing accessibility. This finding suggests that it may be difficult to represent changes in accessibility over time and how those changes influence trip making. The HBO models also fail to show equality for the coefficient on the dummy variable for non-working adults and non-coupled adults with children between the ages of 11 and 15.

At a significance level of 0.05, we fail to reject the hypothesis that the observed trip making probabilities in the 2006 data are given by the 1995 model for both the HBW and HBO models.

Table 35. Measures of Temporal Stability for 1995 Case 2 Models Applied to 2006 Socio-economic Data

Trips	Predicted Fraction of Trips	Observed Fraction of Trips	Predicted Trips	Observed Trips
HBW				
0	0.5522	0.5847	0	0
1	0.1371	0.1308	125094	119402
2	0.2665	0.2572	486339	469649
3+	0.0442	0.0272	149575	92341
Total Trips			761009	681392
%E	14.67			
WRMSE	0.0952			
RSRS	0.0384			
APS	**0.0091**			
HBO				
0	0.5377	0.4466	0	0
1	0.1534	0.1570	139952	143348
2	0.1921	0.2389	350573	436204
3	0.0541	0.0510	148128	139784
4+	0.0627	0.1064	276253	474488
Total Trips			914906	1193824
%E	17.41			
WRMSE	0.2399			
RSRS	0.1115			
APS	**0.0576**			

For bolded cells, H_0 cannot be rejected at the 95% confidence level

Considering the APS test, all Case 1 and Case 2 models pass the test for temporal stability. The most temporally stable model overall considering all measures of temporal stability is the HBW Case 2 model that includes variables related to life cycle. In this case, the inclusion of life cycle variables appears to improve the temporal stability of the model.

7.4.5 Backwards Temporal Stability

To evaluate backwards temporal stability, models were estimated using the 2006 data and then applied to the 1995 data. As before, two cases are evaluated; one considers only the initial set of variables and the other supplemental variables for life cycle, area type, and accessibility. To differentiate from the previous models, these are labeled Case 3 and Case 4, respectively. Table 36 provides a summary of the 2006 model estimation results for the Case 3 and Case 4 models for both trip purposes. As with the models estimated using1995 survey data, results show that household size, the age of the individual, household vehicles, non-working adults, household workers, and income are significant. The presence of children also showed up as significant in these models. In comparison to the 1995 Case 2 models, the supplemental variables found to be significant include many more life cycle variables, area type variables, and accessibility variables. It is important to note that the 2006 survey has a much richer database with 8150 individual records as compared to 3007 for the 1995 survey, which likely influences this result.

Table 36. Model Estimation Results for HBW and HBO Case 3 and Case 4 Models using 2006 Survey

HBW Case 3 Variables and Coefficients		*HBO Case 3 Variables and Coefficients*		*HBW Case 4 Variables and Coefficients*		*HBO Case 4 Variables and Coefficients*	
hhsize	*-0.251*	nwa	*0.914*	age	*0.035*	c3049na	*-0.480*
age	*0.027*	pchild	*0.387*	n_nwa	*-1.487*	c5059na	*-0.307*
n_nwa	*-0.786*	otadult	*-0.274*	c3049na	*0.929*	c60na	*0.361*
hhwrk	*0.769*	less50k	*-0.214*	caa1618	*0.846*	caa06	*0.336*
less25k	*-0.306*	veh	*0.705*	nc1618	*0.253*	caa710	*0.283*
				nc1922	*0.700*	nc55na	*-0.528*
				nc69na	*-0.631*	nc1115	*-0.516*
				op7079na	*-2.082*	op7079na	*0.385*
				op80na	*-3.300*	op80na	*0.494*
				caa1115	*-0.364*	e1e4	*0.143*
				lum3	*0.166*	empd2	*0.118*
						inc4	*0.114*
						otadult	*-0.178*
						moreveh	*0.255*
Cut points specific to		*Cut points specific to*		*Cut points specific to*		*Cut points specific to*	
Trips=1	1.472	Trips=1	0.465	Trips=1	1.047	Trips=1	-0.029
Trips=2	2.234	Trips=2	1.139	Trips=2	1.821	Trips=2	0.643
Trips=3+	5.120	Trips=3	2.416	Trips=3+	4.720	Trips=3	1.907
		Trips=4+	2.893			Trips=4+	2.383

NOTE – All coefficients are significant at the 95% confidence level

Table 37 summarizes the model verification measures for the Case 3 and 4 models estimated using the 2006 survey. All models and trip purposes show a good match between estimated and observed. The inclusion of the supplemental variables does not lead to any improvement of the model verification measures.

Table 37. Model Verification Measures for Models Estimated Using 2006 Survey

Performance Measure		*2006 Case 3 Models*		*2006 Case 4 Models*	*2006 Survey*
HBW Trips/Person		0.83		0.70	0.68
HBO Trips/Person		1.31		1.22	1.14
HBW Trips/Household		2.02		1.69	1.76
HBO Trips/Household		3.18		2.96	2.94
HBW Trips/Worker		1.50		1.26	1.59
HBW					
Number of Trips	*Estimated Trips*	*%E*	*Estimated Trips*	*%E*	*Observed Trips*
1	122572	2.7%	123232	3.2%	119402
2	438951	6.5%	432489	7.9%	469649
3+	89547	3.0%	85747	7.1%	92341
Total	651069	4.5%	641467	5.9%	681392
HBO					
Number of Trips	*Estimated Trips*	*%E*	*Estimated Trips*	*%E*	*Observed Trips*
1	146158	2.0%	145893	1.8%	143348
2	432268	0.9%	410281	5.9%	436204
3	146955	5.1%	135163	3.3%	139784
4+	472869	0.3%	424974	10.4%	474488
Total	1198250	0.4%	1116311	6.5%	1193824

Applying the same temporal tests for the 2006 models yields the results summarized in Table 38. Based on these results, there is evidence of backwards temporal stability for all 2006 models. Of the models estimated, the HBW and HBO Case 4 models appears to provide the best temporal stability with the lowest percent error for the HBW model, and the lowest error considering WRMSE, RSRS, and APS for the HBW model. Additionally, we fail to reject the null hypothesis that the 1995 observed trip-making probabilities are given by the 2006 model. While we fail to reject this hypothesis for all of the models, the percent error is higher than the more desirable 10 percent. Considering all measures, the supplemental variables do improve the temporal stability of both models.

Table 38. Measures of Temporal Stability for 2006 Models Applied to 1995 Socio-economic Data

Trips	Predicted Fraction of Trips	Observed Fraction of Trips	Predicted Trips	Observed Trips
HBW – Case 3				
0	0.5862	0.5110	0	0
1	0.1441	0.1301	91924	79691
2	0.2450	0.3166	383663	387875
3+	0.0248	0.0423	88321	95526
Total Trips			563908	563092
%E	20.83			
WRMSE	0.2103			
RSRS	0.1062			
APS	**0.0442**			
HBO – Case 3				
0	0.4657	0.5530	0	0
1	0.1593	0.1525	98051	93411
2	0.2269	0.1951	281005	238990
3	0.0500	0.0451	93238	82937
4+	0.0982	0.0542	284577	156930
Total Trips			756871	572269
%E	30.84			
WRMSE	0.2031			
RSRS	0.1032			
APS	**0.0413**			
HBW – Case 4				
0	0.5777	0.5110	0	0
1	0.1426	0.1301	89195	79691
2	0.2530	0.3166	331001	387875
3+	0.0267	0.0423	69504	95526
Total Trips			489699	563092
%E	18.48			
WRMSE	0.1842			
RSRS	0.0943			
APS	**0.0339**			

For bolded cells, H_0 cannot be rejected at the 95% confidence level

Table 38 (continued)

Trips	Predicted Fraction of Trips	Observed Fraction of Trips	Predicted Trips	Observed Trips
HBO – Case 4				
0	0.5094	0.5530	0	0
1	0.1558	0.1525	95682	93411
2	0.2072	0.1951	254389	238990
3	0.0438	0.0451	80654	82937
4+	0.0838	0.0542	239375	156930
Total Trips			670100	572269
%E	22.23			
WRMSE	0.1220			
RSRS	0.0542			
APS	**0.0150**			

For bolded cells, H_0 cannot be rejected at the 95% confidence level

7.5 Summary and Conclusions

The purpose of this research is to evaluate the usefulness of cumulative logistic regression models for trip generation and to determine whether the inclusion of supplemental variables related to life cycle, area type, and accessibility improve model fit and performance.

Two activity-based household travel surveys formed the basis of this analysis, one from 1995 and the other from 2006. The selection of the final models considered coefficient reasonableness, check of logical relationships, chi-squared statistics, pseudo R^2, and z-statistics. The evaluation considers home-based work (HBW) and home-based other (HBO) trip purposes. Model development included 1995 and 2006 models with commonly used explanatory variables (Case 1 and 3, respectively), and then a second set of models considering commonly used variables along with variables defining life cycle, area type, and accessibility (Case 2 and 4, respectively).

The 1995 Case 1 and 2 model coefficients have the correct sign and are significant at the 95% level and model statistics are with acceptable ranges. The inclusion of the supplemental variables yields mixed results for the 1995 models. Accessibility variables are significant for HBO trips, but not for HBW trips. None of the area type variables is significant for either trip purpose, while life cycle variables are significant for both trip purposes. Model verification results show a good match between the estimated and observed total trips for both the HBW and HBO trip purposes. The trip rates per person and per household show a good match between estimated and observed and HBW trip rates are with the expected range. The verification results of the Case 2 models are slightly better, but not enough to argue for one model over the other. The 1995 model verification results support the cumulative logistic regression model as a good model for trip generation.

Coefficients for the 2006 models also have the correct sign and are significant at the 95% level with model statistics within acceptable ranges. In comparison to the 1995 Case 2 models, many more life cycle, area type, and accessibility variables show up as significant in the 2006 Case 4 models, perhaps reflecting the difference in the 1995 sample size of 3007 as compared to 8150 for the 2006 survey. Model verification results for the 2006 models show a good match between estimated and observed trips. Trip rate measures also show a good match between the models and the survey.

A secondary purpose of this research is an evaluation of the temporal stability of cumulative logistic regression models. Comparisons include models with and without life cycle, area type, and accessibility variables to determine whether these variables improve temporal stability. The measures of temporal stability consist of an analysis of how well the estimated models predict observed total trips by trip purpose, how well the estimated models predict observed share by trip purpose and stratification, and a comparison of individual model parameters. The results of the temporal stability tests suggest temporal stability for several of the models tested.

For the 1995 model analysis, all models pass the significance test using the aggregate prediction statistic (APS), failing to reject the null hypothesis that the observed trip making probabilities in the 2006 data are given by the 1995 model at a significance level of 0.05.

This finding suggests that the models are temporally stable and acceptable for forecasting trip generation. However, considering the percent error, the measure of how well the 1995 models predict 2006 total trips by trip purpose is not as expected with values close to twenty percent; a value closer to ten percent would increase our confidence regarding temporal stability. The weighted root mean square error (WRMSE) and the root of sum of residual squared (RSRS) suggest that the HBW key factor model is the most temporally stable; this model also has the lowest percent error at 14.67.

To answer the question of backwards temporal stability, models were also estimated using 2006 data and applied using 1995 data. In this example, the Case 4 models are the most temporally stable with the lowest percent error and the lowest error for WRMSE and RSRS. All models were temporally stable considering the APS, although the percent error was higher than desired all models. The supplemental variables improved the temporal stability for both the HBO and HBW trips.

This research shows that cumulative logistic regression models, with their ability to estimate person trips and accommodate more explanatory variables are a good choice for trip generation models. Additionally, the analysis shows temporal stability for cumulative logistic regression models using various measures of stability. With respect to whether life cycle, area type, and accessibility improve temporal stability, this analysis shows that life cycle variables improve the temporal stability of the models.

Future research in this area should focus on cumulative logistic regression models for improving temporal transferability. Spatial transferability refers to the practice of applying data or models developed in one geographic region to another geographic region. Future research should also compare the performance of the cumulative logistic model to the more commonly used generation choice model. Finally, this study considers only HBW and HBO trips. An exploration of home-based shopping and home-based school trips could answer the question of whether this model type is temporally stable for additional trip purposes.

CHAPTER 8. ASSESSING THE IMPACT OF SAMPLE SIZE, MODEL TYPE, AND EXPLANATORY VARIABLES ON TEMPORAL STABILITY

8.1 Introduction and Motivation

A travel demand model is a series of mathematical equations used to describe travel and travel choices. In its most basic form, this series of models is broken into a 4-step process, trip generation, trip distribution, mode choice, and trip assignment. Planners use these models to forecast travel demand 20 – 30 years into the future for the purposes of evaluating transportation strategies and investment. An implicit assumption of a travel demand model is that its parameters remain stable over time (Ortuzar and Willumsen 2011). A violation of this assumption could lead to transportation analyses and travel forecasts that either over- or underestimate travel demand and associated transportation deficiencies, which could in turn lead to poorly allocated investments in transportation infrastructure.

A peer information exchange held in December 2004 to discuss issues of data transferability identified temporal transferability (stability) as a concept that is regularly assumed by modelers, while the validity of this concept has not be sufficiently studied (TMIP 2004). Without a better understanding of temporal stability, it is difficult to defend the use of model parameters developed in one point in time to forecast behavior many years into the future. This investigation adds to that understanding for trip generation models in particular, and addresses the question of whether survey sample size, model form, and explanatory variables contribute to temporal stability.

Relying on findings reported in the literature, comparisons of temporal stability between generation choice models, cumulative logistic regression models, and cross-classification models are drawn. The key question is whether advanced trip generation models with their consideration of person level trips and greater flexibility in including additional explanatory variables lead to better temporal stability than the more commonly used cross-classification model. If shown to offer advantages regarding temporal stability, this may encourage more agencies to move towards advanced trip generation models, perhaps as a step towards advancing the remaining components of their models.

8.2 Methodology

This section provides an overview of the survey data used for model estimation, the two advanced model types explored, the measures of model performance, and the temporal stability tests.

8.2.1 Survey Data and Variable Description

Activity based household travel surveys from the same geographic region, but at two different points in time form the basis of the analysis. The geographic region includes the North Carolina counties of Wake, Durham, and Orange. The 1995 data set includes travel survey records collected between November 1994 and April 1995 and includes 3007 individual person trip records (NuStats 1995). The 2006 data set covers the period between January and June 2006 and includes 8150 individual person trip records (NuStats 2006).

Both surveys were processed using similar techniques for weighting and expansion, trip linking, and the development of trip purposes. This leaves the main difference between the two surveys the sample size; with the 2006 sample size nearly 3 times the size of the 1995 survey. Model estimation and the subsequent tabulation of model verification and temporal stability measures provides for a comparative analysis between models estimated with the 1995 survey and models estimated with the 2006 survey to determine whether the larger sample size leads to better results.

The evaluation covers two case studies for each model type and survey year. The first case study (Case 1) uses a set of explanatory variables widely used in trip generation models. The second case study (Case 2) supplements the initial list of variables with variables describing life cycle, area type, and accessibility.

Variables Tested in Case 1 Models

- age - Age of respondent
- child - Child flag (age <= 15)
- hhadlts – Number of household adults
- hhsize - Household size
- hhveh_grp - Household vehicle group
- hhwrk – Number of household workers
- inc_grp - Income group
- inc1 - Low income flag (< 25k)
- inc2 - Low medium income flag (25k-50k)
- inc3 - Medium high income flag (50k-100k)
- inc4 - High income flag (> 100k)
- less100k - HH income less than 100k flag
- less25k - HH income less than 25k flag
- less50k - HH income less than 50k flag
- lessveh - Vehicles less than workers flag
- moreveh - Vehicles more than workers flag
- n_nwa - Number of non-working adults
- noinc - Unreported income flag
- numveh - Number of vehicles
- nwa - Non-working adult flag
- otadult - Number of other adults
- p_nwa - Presence of a non-working adult flag
- p_wrk - Presence of a worker flag
- pchild - Presence of children flag
- totchld - Total children
- veh - Vehicle present flag
- worker - Worker flag

Supplemental Explanatory Variables Tested in Case 2 Models

- c3049na - couple, age of head of household between 30 and 49, no children
- c5059na - couple, age of head of household between 50 and 59, no children at home
- c60na - couple, age of head of household over 60, no children at home
- caa06 - couple, any age, oldest child between 0 and 6
- caa1115 - couple, any age, oldest child between 11 and 15
- caa1618 - couple, any age, oldest child between 16 and 18
- caa1922 - couple, any age, oldest child between 19 and 22
- caa710 - couple, any age, oldest child between 7 and 10
- e1e_tt - Accessibility exponential function where the measure of attraction is employment and the measure of impedance is travel time
- e1e_1 - Categorical variable for previously defined accessibility function, value is 0 to 15
- e1e_2 - Categorical variable for previously defined accessibility function, value is 16 to 102
- e1e_3 - Categorical variable for previously defined accessibility function, value is 103 to 647
- e1e_4 - Categorical variable for previously defined accessibility function, value is 648 and above
- empd1 - Units are employees per square mile where value is 0 to 29
- empd2 - Units are employees per square mile where value is 30 to 680
- empd3 - Units are employees per square mile where value is 681 and above
- lud1 - Units are households plus employees per square mile where value is 0 to 712
- lud2 - Units are households plus employees per square mile where value is 713 to 3161
- lud3 - Units are households plus employees per square mile where value is 3162 and above
- lum1 - Units are land use mix index ranging from 0 (no mix) to 1 (perfect mix) where value is 0 to 0.16

- lum2 - Units are land use mix index ranging from 0 (no mix) to 1 (perfect mix) where value is 0.16 to 0.59
- lum3 - Units are land use mix index ranging from 0 (no mix) to 1 (perfect mix) where value is 0.59 to 1.00
- nc06 - not coupled, oldest child between 0 and 6
- nc1115 - not coupled, oldest child between 11 and 15
- nc1618 - not coupled, oldest child between 16 and 18
- nc1922 - not coupled, oldest child between 19 and 22
- nc30na - not coupled, head of household 30 or younger, no children
- nc55na - not coupled, head of household between 30 and 55, no children
- nc69na - not coupled, head of household between 56 and 69, no children
- nc710 - not coupled, oldest child between 7 and 10
- op7079na - not coupled, head of household between 70 and 79, no children at home
- op80na - not coupled, head of household 80 or older, no children at home
- p2ep_di - Accessibility power function where the measure of attraction is employment plus population and the measure of impedance is distance
- yc30na – young couple, head of household 30 or younger, no children

8.2.2 Model Estimation and Verification

Model estimation includes generation choice and cumulative logistic regression models for both periods. The selection of the best-fit models started with an evaluation of logical combinations of the variables described in the previous section, with the final selection based on coefficient reasonableness, goodness of fit statistics, and a check of logical relationships with respect to trip making. This research explores home-based work (HBW) and home-based other (HBO) trip generation models. BIOGEME 1.8, a free software package available for the estimation of discrete choice models allows for the estimation of the generation choice models (Bierlaire 2008). Model estimation for the cumulative logistic regression models relied on the statistical software package STATA (StataCorp 2007). The estimated trips for both models are person trips by trip purpose.

Generation Choice Models

The generation choice model is a logit model formulation that estimates the daily trip frequency (n) by trip purpose (y) that a person (x) will make:

$$P_{x,y}(n) = \frac{e^{U_{x,y}(n)}}{\sum e^{U_{x,y}(n)}} \quad \text{(Equation 33)}$$

where:

$P_{x,y}(n)$ = probability of a person of type "x" making "n" daily trips of purpose "y"

$U_{x,y}(n)$ = utility of "n" trips for person type "x" and trip purpose "y".

Cumulative Logistic Regression Models

The cumulative logistic regression model is a type of discrete choice model that estimates relationships between an ordered dependent variable, for example person trip generation, and a set of independent variables, for example household size, income, and workers. Recognizing the number of trips made by an individual as categorical and ordered makes the cumulative logistic regression model a good choice for estimating the probability for a person (i) to make (n) trips of purpose (y):

$$Prob(T_{i,y} = n) = P(\kappa_{n-1} < \beta_1 x_{1_{i,y}} + \beta_2 x_{2_{i,y}} + \cdots \beta_z x_{z_{i,y}} + \varepsilon_{i,y} \leq \kappa_n) \quad \text{(Equation 34)}$$

where:

Prob ($T_{i,y}$ = n) = probability of exactly "n" trips for person "i" and trip purpose "y"

n = number of trips

$T_{i,y}$ = trips for person "i" and purpose "y"

$\beta_1, \beta_2, \ldots, \beta_z$ = estimated coefficients for variables $x_1, x_2, \ldots, x_z$

$x_1, x_2, \ldots, x_z$ = model independent variables

ε = random error, assumed to be logistically distributed

κ = estimated cutpoints

z = number of independent variables.

The probability of a given observation is then (StataCorp 2007) :

$$Prob(T_{i,y} = n) = P(\kappa_{n-1} < \beta x_{i,y} + \varepsilon \leq \kappa_n) = \qquad \text{(Equation 35)}$$

$$\frac{1}{1 + exp\,(-\kappa_n + x_{i,y}\beta)} - \frac{1}{1 + exp\,(-\kappa_{n-1} + x_{i,y}\beta)}$$

Model verification is the process of comparing estimated model results to survey observations. This study includes a comparison of model estimated person trips per person, person trips per household, and work trips per worker against observed surveyed values.

8.2.3 Temporal Stability

Temporal stability addresses the issue of how well models estimated in one period match behavior observed in another. There are several measures for capturing temporal stability in trip generation models. The measures discussed in this paper include how well the estimated models predict observed total trips by trip purpose, and how well the estimated models predict observed shares by trip purpose and stratification.

The measure of how well the estimated models predict observed total trips by trip purpose considers the percent error. Comparisons include percent error summaries of 2006 estimated trips using the 1995 models against the 2006 observed trips, and 1995 estimated trips using the 2006 models against the 1995 observed trips:

$$\%E = \frac{|ET_j - OT_j|}{OT_j} * 100 \qquad \text{(Equation 36)}$$

where:

j = label for 2006 trips

%E = percent error

ET_j = 2006 estimated trips using 1995 model

OT_j = 2006 observed trips.

The measure of how well the estimated models predict observed shares by trip purpose and stratification considers the weighted root mean square error (WRMSE) and the root of sum of residual squared (RSRS) (Agyemang-Duah and Hall 1997). Both are measures of the aggregate prediction accuracy of the transferred model:

$$WRMSE = (\sum_k PS_k * (\frac{PS_k - OS_k}{PS_k})^2)^{.5} \quad \text{(Equation 37)}$$

and

$$RSRS = (\sum_k (PS_k - OS_k)^2)^{.5} \quad \text{(Equation 38)}$$

where:

k = number of trips by trip purpose (0, 1, 2, …, n_j)

j = trip purpose index

WRMSE = weighted root mean square error

RSRS = root of sum of residual squared

PS_k = predicted fraction for trips k

OS_k = observed fraction for trips k.

The t-test evaluates the difference between the estimated parameters between the two models:

$$t = \frac{\alpha_{xi} - \alpha_{xi}}{(var(\alpha_{xi}) + var(\alpha_{xj}))^{0.5}} \quad \text{(Equation 39)}$$

where:

t = t-test statistic

i = label for 1995 context

j = label for 2006 context

α_{xi} = estimated coefficient for variable x in 1995 context

α_{xj} = estimated coefficient for variable x in 2006 context

var(α_{xi}) = variance of α_{xi}

var(α_{xj}) = variance of α_{xj}.

The measure of stability of the trips by stratification uses the aggregate prediction statistic (APS). This statistic is equivalent to the chi-square distribution with degrees of freedom equal to the number of alternatives – 1 and tests the hypothesis that the observed trip fractions are given by the estimated model (Agyemang-Duah and Hall 1997):

$$APS = \sum_k \frac{(PS_k - OS_k)^2}{PS_k} \qquad \text{(Equation 40)}$$

where:

APS = aggregate prediction statistic.

The null hypothesis of no significant difference between the estimated and observed trip fractions is rejected when the calculated APS (i.e. χ^2) is greater than or equal to $\chi^2_{\alpha,df}$. The test is applied at a 95% confidence level, or α = 0.05 and degrees of freedom equal to the number of trip alternatives minus one.

8.2.4 Experimental Design

Figure 8 graphically displays the experimental design for this analysis. The analysis includes multiple comparisons evaluating both model performance and temporal stability. First, generation choice and cumulative logistic regression models are estimated using 1995 survey data and 2006 survey data. As noted previously, two cases are evaluated for each. Case 1 uses a set of explanatory variables widely used in trip generation models, and Case 2 supplements the initial list of variables with variables defining life cycle, area type, and accessibility. Letters a – d are used to denote the difference between survey data and model type as shown in Figure 8. Section 8.3 presents the results of the individual models. Section 8.4 summarizes the results of the comparative analysis of the models. The comparative analysis includes and evaluation of model performance and temporal stability considering survey size, model form, and explanatory variables.

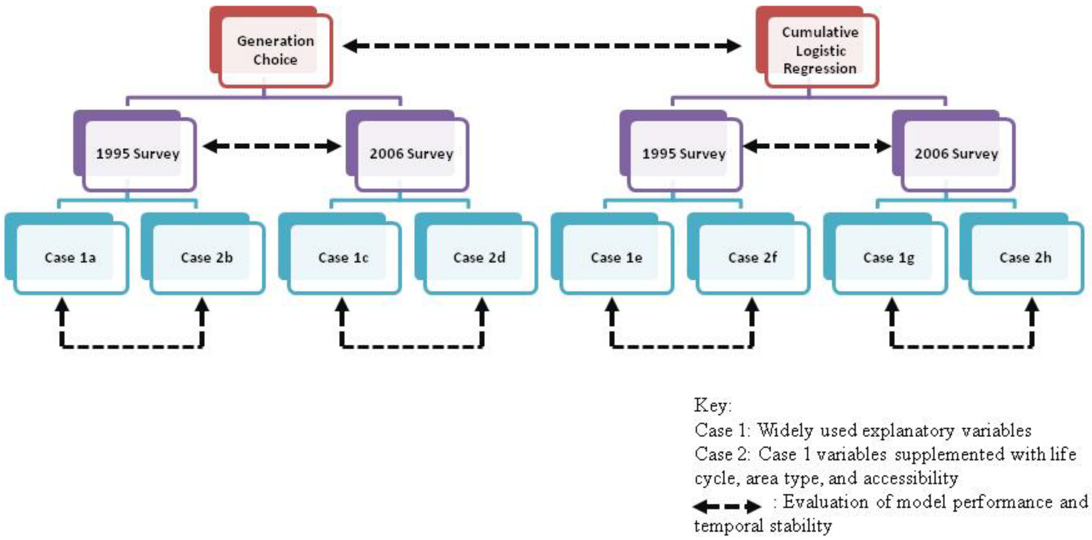


Figure 8. Experimental Design

8.3 Results

This section presents the results of the model estimation, verification, and temporal stability tests for the 1995 and 2006 generation choice and cumulative logistic models. The summary of results is by estimation year and the case considered. The variables evaluated for the Case 1a, 1b, 1c and 1d models are from the list of widely used explanatory variables, and Case 2a, 2b, 2c, and 2d models supplemented these variables with variables defining life cycle, area type, and accessibility.

8.3.1 Generation Choice

Generation choice models estimate the probability of a worker making 0, 1, 2, or 3+ HBW trips; a worker making 0, 1, 2, 3, or 4+ HBO trips; a non-working adult making 0, 1, 2, 3 or 4+ HBO trips, and a child making 0, 1, or 2+ HBO trips.

Model Results

Tables 39 through 42 provide a summary of the generation choice coefficients and model statistics. The rho-squared values are within the typical range for models of this type and coefficients have the correct sign with most being significant at the 95% level (significant coefficients in bold). Retention of coefficients below the specified significance is due to their influence on other significant variables. Many more variables are significant for the 2006 worker models, though this does not necessarily improve the model fit.

Table 39. Generation Choice HBW Worker Model Estimation Results

1995 Case 1a Variables and Coefficients		*1995 Case 2a Variables and Coefficients*		*2006 Case 1b Variables and Coefficients*		*2006 Case 2b Variables and Coefficients*	
ASC_1	***0.422***	ASC_1	*0.094*	ASC_1	*0.068*	ASC_1	***-0.317***
ASC_2	***0.681***	ASC_2	***0.831***	ASC_2	***0.542***	ASC_2	***0.196***
ASC_3	***-1.080***	ASC_3	***-0.915***	ASC_3	***-1.730***	ASC_3	***-1.880***
inc2_3	*0.339*	inc1_3	***-0.637***	inc2_3	***0.569***	inc2_3	***0.551***
lessveh_1	***-0.670***	lessveh_1	***-0.605***	inc3_1	***0.220***	inc3_1	***0.326***
moreveh_2	*0.219*	c3049na_1	***0.485***	inc3_2	***0.204***	inc3_2	***0.190***
otadult_1	*-0.134*	c3049na_2	***0.405***	lessveh_1	*-0.527*	inc4_1	***0.198***
pchild_1	***-0.305***			otadult_1	*-0.143*	lessveh_1	***-0.516***
						c3049na_2	***0.388***
						empd3_3	***0.519***
						hhsize_2	***0.084***
						nc1115_1	***0.991***
						nc1115_2	***0.724***
						nc55na_1	***0.879***
						nc55na_2	***0.821***
						nc710_1	***1.380***
						op7079na_2	***-0.620***
						yc30na_1	***0.684***
						yc30na_2	***0.709***
Log likelihood		*Log likelihood*		*Log likelihood*		*Log likelihood*	
Zero	-2656.14	Zero	-2656.14	Zero	-6627.87	Zero	-6627.87
Constant	-2312.53	Constant	-2312.53	Constant	-5711.00	Constant	-5711.00
Converge	-2300.24	Converge	-2303.07	Converge	-5685.87	Converge	-5637.93
Rho2	0.134	*Rho2*	0.133	*Rho2*	0.14	*Rho2*	0.149

Note: Bold coefficients are significant at the 95% confidence level; _x is used to capture the number of trips, where x=1, 2,…, n depending on the trip purpose

Table 40. Generation Choice HBO Worker Model Estimation Results

1995 Case 1a Variables and Coefficients		*1995 Case 2a Variables and Coefficients*		*2006 Case 1b Variables and Coefficients*		*2006 Case 2b Variables and Coefficients*	
ASC_1	***-1.270***	ASC_1	***-1.310***	ASC1	***-0.859***	ASC1	***-0.838***
ASC_2	***-1.240***	ASC_2	***-1.340***	ASC2	***-1.300***	ASC2	***-0.795***
ASC_3	***-3.090***	ASC_3	***-2.580***	ASC3	***-1.930***	ASC3	***-3.010***
ASC_4	***-2.990***	ASC_4	***-2.760***	ASC4	***-2.360***	ASC4	***-2.180***
nchild_4	***0.412***	c3049na_3	*0.61*	veh_2	***0.847***	nchild_3	***0.150***
otAdult_1	*-0.005*	caa710_4	*1.27*	hhadlt_2	***-0.269***	c3049na_4	***-1.090***
otAdult_2	*-0.071*	nc1115_2	*1.14*	hhsize_4	***0.230***	c5059na_4	***-1.040***
vehgrp_3	***0.292***	nc30na_1	*0.46*	otadult_1	***-0.156***	c60na_3	***0.689***
				otadult_3	***-0.240***	caa1115_4	***0.539***
						e1e_3	***0.205***
						e1e_4	***0.221***
						hhsize_1	***-0.072***
						nc06_4	***1.370***
						nc55na_2	***-0.421***
						nc55na_4	***-0.607***
						nc710_3	***1.330***
						otadult_2	***-0.137***
Log likelihood		*Log likelihood*		*Log likelihood*		*Log likelihood*	
Zero	-3083.68	Zero	-3083.68	Zero	-7694.72	Zero	-7694.72
Constant	-2272.18	Constant	-2272.18	Constant	-6424.84	Constant	-6424.84
Converge	-2261.16	Converge	-2255.20	Converge	-6389.84	Converge	-6345.56
Rho2	0.267	*Rho2*	0.269	*Rho2*	0.17	*Rho2*	0.18

Note: Bold coefficients are significant at the 95% confidence level; _x is used to capture the number of trips, where x=1, 2,…, n depending on the trip purpose

Table 41. Generation Choice HBO Non-Working Adult Model Estimation Results

1995 Case 1a Variables and Coefficients		*1995 Case 2a Variables and Coefficients*		*2006 Case 1b Variables and Coefficients*		*2006 Case 2b Variables and Coefficients*	
ASC_1	***-0.504***	ASC_1	***-0.790***	ASC_1	***-0.398***	ASC_1	***-0.462***
ASC_2	*0.116*	ASC_2	*-0.020*	ASC_2	***0.315***	ASC_2	***0.241***
ASC_3	***-1.850***	ASC_3	***-1.350***	ASC_3	***-1.070***	ASC_3	***-1.150***
ASC_4	***-0.511***	ASC_4	***-1.330***	ASC_4	***-0.525***	ASC_4	***-0.858***
inc3_2	*0.400*	inc2_3	***-1.010***	inc3_4	***0.460***	inc4_3	***0.664***
inc3_3	***0.642***	caa06_4	***1.530***	inc4_3	***0.697***	caa06_4	***1.050***
nchild_4	***0.465***	caa1115_4	***1.050***	inc4_4	***0.573***	caa1115_4	***0.992***
otAdult_1	***-0.598***	caa710_4	***1.520***	nchild_4	***0.307***	caa710_4	***1.790***
otAdult_2	***-0.533***	lum3_1	***0.883***	hhwrk_1	***-0.546***	hhwrk_1	***-0.466***
otAdult_4	***-0.756***	lum3_2	***0.532***	hhwrk_4	***-0.289***	nc1618_2	***-1.130***
		otAdult_1	***-0.477***	otadult_2	***-0.319***	otadult_2	***-0.245***
		otAdult_2	***-0.384***	otadult_3	***-0.517***	otadult_3	***-0.443***
		otAdult_3	*-0.103*	otadult_4	***-0.325***		
Log likelihood		*Log likelihood*		*Log likelihood*		*Log likelihood*	
Zero	-975.32	Zero	-975.32	Zero	-2697.42	Zero	-2697.42
Constant	-877.64	Constant	-877.64	Constant	-2506.14	Constant	-2506.14
Converge	-854.21	Converge	-849.99	Converge	-2461.70	Converge	-2450.81
Rho2	0.124	*Rho2*	0.129	*Rho2*	0.09	*Rho2*	0.09

Note: Bold coefficients are significant at the 95% confidence level; _x is used to capture the number of trips, where x=1, 2,…, n depending on the trip purpose

Table 42. Generation Choice HBO Child Model Estimation Results

1995 Case 1a Variables and Coefficients		*1995 Case 2a Variables and Coefficients*		*2006 Case 1b Variables and Coefficients*		*2006 Case 2b Variables and Coefficients*	
ASC_1	***-1.110***	ASC_1	***-1.450***	ASC_1	***-2.260***	ASC_1	***-1.160***
ASC_2	***-0.532***	ASC_2	***-0.540***	ASC_2	***0.126***	ASC_2	***0.122***
inc1_2	***-1.590***	inc1_2	***-1.530***	inc1_2	***-0.665***	inc1_2	***-0.602***
n-nwa_1	***-0.704***	nc06_1	***2.060***	moreveh_1	***1.170***	nc710_1	***0.885***
		nc710_1	***1.470***				
Log likelihood		Log likelihood		Log likelihood		Log likelihood	
Zero	-532.83	Zero	-532.83	Zero	-1859.95	Zero	-1859.95
Constant	-470.13	Constant	-470.13	Constant	-1681.83	Constant	-1681.83
Converge	-459.88	Converge	-456.99	Converge	-1674.11	Converge	-1674.87
Rho2	0.137	Rho2	0.142	Rho2	0.10	Rho2	0.10

Note: Bold coefficients are significant at the 95% confidence level; _x is used to capture the number of trips, where x=1, 2,…, n depending on the trip purpose

Model Verification

Table 43 summarizes the generation choice verification measures. Results show a good match between the estimated and observed total trips for all models. The results of the HBO-WRK and HBO-CLD models estimated with 2006 survey data are slightly better the models estimated with the 1995 survey data. The introduction of lifecycle improves estimation results for HBO trips by children, the 2006 worker models, and the 1995 HBO-NWA model. In comparison to the observed 1995 values, modeled 1995 HBW trips per person and HBO trips per person are high. The HBW trips per worker is within the acceptable high and low benchmarks for all models (Schiffer and Rossi 2008).

Table 43. Model Verification Measures for Generation Choice Models

	1995 Case 1a	1995 Case 2a	1995 Survey	2006 Case 1b	2006 Case 2b	2006 Survey
HBW Trips/Person	0.92	0.94	0.78	0.73	0.74	0.68
HBO Trips/Person	0.93	0.96	0.78	1.33	1.34	1.14
HBW Trips/Household	2.07	2.11	1.99	1.76	1.78	1.76
HBO Trips/Household	2.10	2.16	2.00	3.23	3.26	2.94
HBW Trips/Worker	1.45	1.48	1.30	1.31	1.33	1.59

	1995 Case 1a		1995 Case 2a		2006 Case 1b		2006 Case 2b	
HBW Worker								
Number of Trips	*Estimated*	*%E*	*Estimated*	*%E*	*Estimated*	*%E*	*Estimated*	*%E*
1	94055	18.0%	87708	10.1%	119042	0.3%	116448	2.5%
2	353537	8.9%	385296	0.7%	450625	4.1%	459436	2.2%
3+	119749	25.4%	106406	11.4%	94954	2.8%	96189	4.2%
Total	567341	0.8%	579410	2.9%	664621	2.5%	672072	1.4%
HBO Worker								
Number of Trips	*Estimated*	*%E*	*Estimated*	*%E*	*Estimated*	*%E*	*Estimated*	*%E*
1	63766	7.8%	63061	6.6%	87791	0.4%	85974	2.4%
2	122622	2.4%	122438	2.2%	175116	4.8%	181876	1.1%
3	56741	26.6%	59721	33.3%	82734	0.3%	83621	1.3%
4+	67417	3.8%	77929	11.2%	229571	1.3%	239830	3.1%
Total	310547	5.7%	323148	10.0%	575212	2.0%	591300	0.7%
HBO Non-Working Adult								
Number of Trips	*Estimated*	*%E*	*Estimated*	*%E*	*Estimated*	*%E*	*Estimated*	*%E*
1	13663	5.0%	13770	4.3%	29424	4.5%	29265	4.0%
2	61528	10.3%	60752	11.4%	117595	3.1%	115696	1.4%
3	27436	9.0%	25204	0.2%	41089	12.2%	41290	12.7%
4+	67350	2.6%	74498	7.7%	160427	2.3%	166735	6.3%
Total	169978	4.1%	174223	1.7%	348535	3.8%	352985	5.1%
HBO Child								
Number of Trips	*Estimated*	*%E*	*Estimated*	*%E*	*Estimated*	*%E*	*Estimated*	*%E*
1	17442	12.2%	17992	9.4%	17532	35.3%	29929	10.5%
2+	77567	4.6%	77271	5.0%	269391	10.5%	253503	4.0%
Total	95009	6.1%	95263	5.8%	286924	5.9%	283432	4.6%

Temporal Stability

Evaluation of the temporal stability of the models involves applying the model estimated in one year (i.e. 1995) to socio-economic data from another year (i.e. 2006), and then comparing the results to observed data from the application year (i.e. 2006). Table 44 summarizes the results of the temporal stability tests for the generation choice models. The first three tests are a measure of how well the models estimated in one year predict observed total trips in another year (%E), and how well the models estimated in one year predict observed fractions of trips in another year (WRMSE and RSRS). These are measures of prediction accuracy of the transferred model. The APS test is a statistical test to test the hypothesis that the observed trip fractions in one year (i.e. 2006) are given by the model from the other year (i.e. 1995). The other tests are useful for comparison of prediction accuracy only.

The percent error suggests that the HBW trip models for workers (HBW-WRK) and HBO trip models for non-working adults (HBO-NWA) are the most temporally stable. Considering the measures of aggregate prediction power, the WRMSE and RSRS indicate that the HBW-WRK model has the lowest prediction error, while the HBO trip models for children (HBO-CLD) has the highest error. The aggregate error is not dramatically different between one case versus the other with respect to survey data or explanatory variables (comparing Case 1a, 2a, 1b, and 2b).

The Aggregate Prediction Statistic (APS) tests the hypothesis that the observed trip making probabilities in the 2006 data are given by the 1995 model. At a significance level of 0.05, we fail to reject this hypothesis for all but the HBO-CLD models, indicating temporal stability for all but HBO-CLD model. In the case of the HBO-CLD model, we reject this hypothesis for all but the Case 2b model, suggesting that including life cycle as a variable in the HBO-CLD model leads to temporal stability.

Table 44. Measures of Temporal Stability for Estimated Generation Choice Models Applied to Socio-economic Data

Temporal Stability Measures	1995 Models Applied to 2006 SE Data		2006 Models Applied to 1995 SE Data	
	Case 1a	Case 2a	Case 1b	Case 2b
HBW Worker				
%E	9%	12%	4%	1%
WRMSE	0.1275	0.1474	0.1479	0.1112
RSRS	0.0429	0.0587	0.0787	0.0440
APS	**0.0162**	**0.0217**	**0.0219**	**0.0124**
HBO Worker				
%E	23%	24%	48%	44%
WRMSE	0.2563	0.2783	0.2496	0.2332
RSRS	0.1043	0.1077	0.1249	0.1183
APS	**0.0657**	**0.0775**	**0.0623**	**0.0544**
HBO Non-Working Adult				
%E	3%	2%	16%	23%
WRMSE	0.1487	0.1231	0.2212	0.2613
RSRS	0.0735	0.0629	0.1147	0.1324
APS	**0.0221**	**0.0152**	**0.0489**	**0.0683**
HBO Child				
%E	24%	24%	38%	37%
WRMSE	0.3274	0.3343	0.4235	0.2715
RSRS	0.1981	0.1963	0.1997	0.1700
APS	0.1072	0.1118	0.1793	**0.0737**

For bolded cells, H_0 cannot be rejected at the 95% confidence level

8.3.2 Cumulative Logistic Regression

Unlike the generation choice models where the trip purposes are stratified by worker, non-working adult, and children, the cumulative logistic regression model is estimated by trip purpose for all trip alternatives and all individuals in the trip set. The generation choice model is more flexible allowing the user to specify a separate utility for each trip alternative. For example, there is a utility equation for a worker making 0 trips, another for a worker making 1 trip, another for a worker making 2 trips, and yet another for a worker making 3 or more trips. This flexibility facilitates the stratification of the data into worker, non-worker, and child. In this same example, the cumulative logistic regression model estimates the probability of an individual making 0, 1, 2, or 3+ HBW trips with one equation. Because of this limitation, stratification of the data by worker, non-working adult, and child results in poorly fit models. As such, only two cumulative logistic regression models are estimated, a model that estimates the probability of an individual making 0, 1, 2, or 3+ HBW trips and a model that estimates the probability of an individual making 0, 1, 2, 3, or 4+ HBO trips.

Model Results

Tables 45 and 46 provide a summary of the model results for the cumulative logistic regression models. The statistics reported for these models are similar to statistics reported in earlier research (Agyemang-Duah and Hall 1997). Model coefficients have the expected sign and all are significant at the 95% level.

Table 45. Model Estimation Results for HBW Models using 1995 and 2006 Survey

1995 Case 1c Variables and Coefficients		*1995 Case 2c Variables and Coefficients*		*2006 Case 1d Variables and Coefficients*		*2006 Case 2d Variables and Coefficients*	
hhsize	*-0.285*	c3049na	*0.798*	hhsize	*-0.251*	age	*0.035*
age	*0.026*	caa06	*0.859*	age	*0.027*	n_nwa	*-1.487*
n_nwa	*-0.952*	caa1922	*0.756*	n_nwa	*-0.786*	c3049na	*0.929*
hhwrk	*0.781*	nc30na	*1.510*	hhwrk	*0.769*	caa1618	*0.846*
		nc55na	*1.192*	less25k	*-0.306*	nc1618	*0.253*
		op7079na	*-1.473*			nc1922	*0.700*
		yc30na	*1.504*			nc69na	*-0.631*
		n_nwa	*-1.051*			op7079na	*-2.082*
		age	*0.042*			op80na	*-3.300*
		hhwrk	*0.636*			caa1115	*-0.364*
						lum3	*0.166*
Cut points specific to		*Cut points specific to*		*Cut points specific to*		*Cut points specific to*	
Trips=1	*1.007*	Trips=1	*2.577*	Trips=1	1.472	Trips=1	1.047
Trips=2	*1.758*	Trips=2	*3.371*	Trips=2	2.234	Trips=2	1.821
Trips=3+	*4.379*	Trips=3+	*6.048*	Trips=3+	5.120	Trips=3+	4.720

NOTE – All coefficients are significant at the 95% confidence level

Table 46. Model Estimation Results for HBO Models using 1995 and 2006 Survey

1995 Case 1c Variables and Coefficients		*1995 Case 2c Variables and Coefficients*		*2006 Case 1d Variables and Coefficients*		*2006 Case 2d Variables and Coefficients*	
nwa	*1.056*	nwa	*0.920*	nwa	*0.914*	c3049na	*-0.480*
hhsize	*0.137*	inc1	*-0.339*	pchild	*0.387*	c5059na	*-0.307*
hhveh_grp	*0.230*	e1e1	*0.472*	otadult	*-0.274*	c60na	*0.361*
hhadlts	*-0.498*	e1e3	*0.525*	less50k	*-0.214*	caa06	*0.336*
inc1	*-0.266*	e1e4	*0.611*	veh	*0.705*	caa710	*0.283*
		nc55na	*-0.262*			nc55na	*-0.528*
		yc30na	*-0.487*			nc1115	*-0.516*
		nc1115	*0.560*			op7079na	*0.385*
						op80na	*0.494*
						e1e4	*0.143*
						empd2	*0.118*
						inc4	*0.114*
						otadult	*-0.178*
						moreveh	*0.255*
Cut points specific to		*Cut points specific to*		*Cut points specific to*		*Cut points specific to*	
Trips=1	*0.170*	Trips=1	*0.745*	Trips=1	0.465	Trips=1	-0.029
Trips=2	*0.857*	Trips=2	*1.432*	Trips=2	1.139	Trips=2	0.643
Trips=3	*2.120*	Trips=3	*2.695*	Trips=3	2.416	Trips=3	1.907
Trips=4+	*2.817*	Trips=4+	*3.391*	Trips=4+	2.893	Trips=4+	2.383

NOTE – All coefficients are significant at the 95% confidence level

Model Verification

Table 47 summarizes the model verification measures for the cumulative logistic regression models. Results show a good match between the estimated and observed total trips for all models. The introduction of lifecycle, area type, and accessibility improves estimated trips as compared to observed trips for the 1995 HBO model only. The estimated trip rates for HBW trips per person and HBO trips per person are high as in comparison to observed values in all but one case. HBW trips by workers are higher than observed for the 1995 models, but lower than observed for the 2006 models, but in all cases the rates are within the accepted benchmarks for this measure (Schiffer and Rossi 2008).

Table 47. Model Verification Measures for Cumulative Logistic Regression Models

	1995 Case 1c	1995 Case 2c	1995 Survey	2006 Case 1d	2006 Case 2d	2006 Survey
HBW Trips/Person	0.90	0.90	0.78	0.83	0.70	0.68
HBO Trips/Person	0.96	0.95	0.78	1.31	1.22	1.14
HBW Trips/Household	2.03	2.03	1.99	2.02	1.69	1.76
HBO Trips/Household	2.15	2.13	2.00	3.18	2.96	2.94
HBW Trips/Worker	1.43	1.42	1.30	1.50	1.26	1.59

	1995 Case 1c		1995 Case 2c		2006 Case 1d		2006 Case 2d	
HBW								
Number of Trips	*Estimated*	*%E*	*Estimated*	*%E*	*Estimated*	*%E*	*Estimated*	*%E*
1	90814	14.0%	89439	12.2%	122572	2.7%	123232	3.2%
2	360765	7.0%	358093	7.7%	438951	6.5%	432489	7.9%
3+	105943	10.9%	108655	13.7%	89547	3.0%	85747	7.1%
Total	557522	1.0%	556187	1.2%	651069	4.5%	641467	5.9%
HBO								
Number of Trips	*Estimated*	*%E*	*Estimated*	*%E*	*Estimated*	*%E*	*Estimated*	*%E*
1	93371	0.0%	93059	0.4%	146158	2.0%	145893	1.8%
2	229706	3.9%	228074	4.6%	432268	0.9%	410281	5.9%
3	95825	15.5%	94700	14.2%	146955	5.1%	135163	3.3%
4+	170757	8.8%	168469	7.4%	472869	0.3%	424974	10.4%
Total	589659	3.0%	584301	2.1%	1198250	0.4%	1116311	6.5%

Temporal Stability

Table 48 summarizes the measures of temporal stability for the cumulative logistic regression models. The percent error for all models is above 10%, though values below 20% may be considered acceptable for a long-range forecast, this analysis measures change over an 11-year period and lower values are desirable. The aggregate prediction measures tell a better story regarding temporal stability. Regarding the statistical test for whether the models estimated the observed trip making probabilities, the APS tests that at a significance level of 0.05 we fail to reject the hypothesis that the models predict the observed probabilities for all models.

Table 48. Measures of Temporal Stability for Estimated Cumulative Logistic Regression Models Applied to Socio-economic Data

Temporal Stability Measures	1995 Models Applied to 2006 SE Data		2006 Models Applied to 1995 SE Data	
	Case 1c	Case 2c	Case 1d	Case 2d
HBW				
%E	19.86	14.67	20.83	18.48
WRMSE	0.1231	0.0952	0.2103	0.1842
RSRS	0.0566	0.0384	0.1062	0.0943
APS	**0.0151**	**0.0091**	**0.0442**	**0.0339**
HBO				
%E	18.31	17.41	30.84	22.23
WRMSE	0.2486	0.2399	0.2031	0.1220
RSRS	0.1158	0.1115	0.1032	0.0542
APS	**0.0618**	**0.0576**	**0.0413**	**0.0150**

For bolded cells, H_0 cannot be rejected at the 95% confidence level

8.4 Comparative Analysis

The comparative analysis considers differences between surveys, model form, and the influence of supplemental variables for lifecycle, area type, and accessibility. The evaluation considers model performance and temporal stability.

8.4.1 Model Performance

This analysis shows that generation choice and cumulative logistic regression models are both good choices for estimating trip generation. Both are capable of considering multiple explanatory variables and estimating person trips, rather than household trips.

Survey Data

The results suggest that a larger sample size increases the number of variables that are statistically significant, but does not noticeably improve model performance.

Model Form

Both the generation choice and cumulative logistic regression models are acceptable models based on the reasonableness of the coefficients and model statistics. Simple model statistics do not suggest that one model is appreciably better than the other is for any of the trip purposes.

Life Cycle, Area Type, and Accessibility Variables

Model verification measures show a slight improvement with the inclusion of life cycle, area type, and accessibility variables.

8.4.2 Temporal Stability

This analysis shows evidence of temporal stability for both the generation choice models and the cumulative logistic regression models. The following sections summarize findings related to whether temporal stability is better for a larger sample size, better for one model form over the other, and is better when life cycle, area type, and accessibility are included in the models.

Survey Data

In general, the generation choice models is better with models based on the 1995 survey data based on the percent error and RSRS results. Considering only the percent error, the cumulative logistic regression model is also more temporally stable for models based on the 1995 survey data. This finding suggests that a larger sample size does not improve temporal stability.

Model Form

The reported percent error for the HBW-WRK generation choice models is much lower than the percent error reported for the HBW cumulative logistic regression model. Direct comparisons for the HBO trip purpose are more difficult as estimation for the generation choice models considers workers, non-working adults, and children separately. It does appear that this approach improves temporal stability because the reported percent error for the HBO-NWA model is considerably lower than the reported error for the cumulative logistic regression HBO model, while error reported for the HBO-WRK and HBO-CLD models is similar to the cumulative logistic regression HBO model. This finding suggests that generation choice models are more temporally stable.

Life Cycle, Area Type, and Accessibility Variables

The inclusion of life cycle, area type, and accessibility generally improves the temporally stability of the generation choice models estimated with the 2006 survey data, but not the models estimated with the 1995 survey data. The temporal stability of the cumulative logistic regression models improves with the inclusion of life cycle, area type, and accessibility for HBW and HBO models estimated using 1995 and 2006 survey data. This finding suggests improved temporal stability with the inclusion of explanatory variables for life cycle, area type, and accessibility.

Discrete Models versus Cross-Classification Models

The generation choice and cumulative logistic regression models are both discrete models. Model estimation requires individual observations of trip choice. The cross-classification model uses categorical analysis where distinct categories form the basis of the analysis. The literature provides insight into the temporal stability of cross-classification models.

Trip generation cross-classification models were the subject of investigation for Walker and Peng in their 1991 comparison of the 1960 and 1987-88 survey for the Delaware Valley Region (Walker and Peng 1991). The focus of this study was not simply the temporal stability of the models, but on whether or not the stratification scheme selected improved the temporal stability of the models. An initial comparison of household trip rates stratified only by trip purpose showed a statistically significant difference between the 1960 trip rates and the 1987-88 trip rates. To test the effect of household size, income, automobile ownership, and area type additional cross-classification models were estimated using various combinations of stratification. The home-based work trip purpose was most stable when stratified by household size. The stratification of home-based non-work trips by auto ownership and area type displayed reasonable stability, while stratification by household size had the opposite effect, even when further stratified by income, automobile ownership, or area type. They concluded from these results that the selection of variables used in the cross-classification model makes a difference with regard to temporal stability. The authors further hypothesized that three-way classification of household size, automobile ownership, and area type may improve the temporal stability of non-work trips; however, a limited sample size prevented the testing of this hypothesis. The results of this research suggest that the stratification of cross-classification models by key variables can improve temporal stability, though sample size may restrict the number of stratifications. While results varied by trip purpose, in general, their findings showed that appropriately specified cross classification models are sufficiently stable and can be valuable tools for testing long-range transportation plans (Walker and Peng 1991).

Other studies have shown less than favorable results. (Ashford and Holloway 1972) investigated a cross-classification model for work trips. The analysis considered data from a 1958 and 1967 home interview survey provided by the Southwestern Pennsylvania Regional Planning Commission. The model was shown to be unstable over time. Another study conducted in 1976 considered multiple trip purposes for cross-classification models developed using a 1953 and 1965 home interview survey from Detroit, Michigan (Smith and Cleveland 1976). In this instance, results showed differences in the trip generation rates of up to 20%.

(Doubleday 1977) also considered the issue of temporal stability using the Reading, United Kingdom travel surveys from 1962 and 1971. Of particular interest in this analysis was the consideration of auto availability. Testing included cross-classification models stratified across car availability and several variations using employment status, household structure, and socio-economic group. The results indicated that in general trips rates are not stable over time although certain classification schemes appear to have more stability than others do, for example employed males stratified by auto availability and socioeconomic groups stratified by auto availability.

The literature presents more examples than not of unfavorable temporal stability tests for cross-classification models. Temporal stability for cross-classification models does appear to improve with a greater number of stratifications, but sample size quickly limits the number of stratifications considered. This is not a limitation of either the generation choice or cumulative logistic regression models, suggesting that they may be more temporally stable, though an evaluation of this hypothesis is not included in this analysis.

8.5 Findings and Recommendations

This paper reports on a comparison of generation choice models and cumulative logistic regression models. Included is an assessment of model performance and temporal stability. The comparative analysis considers differences between survey data, model form, and the influence of explanatory variables defining life cycle, area type, and accessibility. The paper includes an assessment of the temporal stability of cross-classification models based on findings documented in the literature.

8.5.1 Model Performance Findings

With regard to survey data, the findings in this study suggest that a larger sample size increases the number of variables that are statistically significant, but does not significantly improve model performance as was noted for models estimated using the 2006 survey with 8150 person trip records compared to the 1995 survey with 3007 person trip records. Because model estimation for the generation choice and cumulative logistic regression models relies on individual trip records, models with multiple variables can be estimated with acceptable results, even when the sample size is small. The performance measures reported support the use of generation choice and cumulative logistic regression models for trip generation, while simple model statistics do not suggest that one is appreciably better than the other is. With regard to life cycle, area type, and accessibility, the introduction of these variables improves slightly the model verification measures for the models.

8.5.2 Temporal Stability Findings

In this analysis, sample size does not appear to have a direct impact on the temporal stability for either model. The analysis shows evidence of temporal stability for both the generation choice models and the cumulative logistic regression models. The temporal stability of the generation choice models is best, especially considering the segmentation of the home-based other trips. The comparison to the cross-classification model suggests that generation choice models and cumulative logistic regression models are more temporally stable based on consistent results across all models estimated as compared to the literature cited for the cross-classification models. This finding is inconclusive without the full estimation and testing of cross-classification models with the same data set. The results from this analysis show that explanatory variables defining life cycle, area type, and accessibility do have a positive effect on the temporal stability of discrete trip generation models.

8.5.3 Recommendations

Based on the findings documented in the study, a move towards advanced trip generation models with their consideration of person level trips and the greater flexibility for including more independent variables improves the temporal stability of model parameters, minimizing forecast error related to temporal change. Given the greater flexibility for including more independent variables, these findings suggest that including life cycle, area type, and accessibility in the model specification improves trip model verification and temporal stability. If budget allows, agencies would benefit from a larger survey sample size, but even with a small sample survey, there are benefits of moving to advanced models

CHAPTER 9. EVALUATING THE IMPLICATIONS OF TEMPORAL INSTABILITY

9.1 Introduction and Motivation

Trip generation models are used in traditional 4-step travel models to estimate the number of trips generated by a person or a household. An initial goal of these models is to represent observed travel behavior, but that is only part of the challenge. Second, these trip generation models must also forecast future travel behavior based on a set of input demographics and the behavioral relationships captured in the model specification. The second goal is as important as the first since travel demand models play a significant role in the transportation planning process, providing transportation planners, highway designers, transit operators, and decision makers with critical travel forecasts needed for the development and implementation of transportation plans, projects and policies.

Poor travel forecasts can be the result of many things including errors in model specification, model calibration, model verification, and input parameters like trip generation rates. Forecasting errors can also result from model parameters that change over time, or in other words, when the behavior captured by the base year trip rates does not hold true in the future. Experienced model developers know that well specified and calibrated models can minimize the impact of errors resulting from model specification, calibration, and verification issues, while better forecasting techniques can improve the quality of the input data. Improving our understanding of the factors that influence travel behavior and how model parameters including trip generation rates change over time can lead to better temporal stability of travel models, increasing our confidence in model results.

The purpose of this analysis is to evaluate the implication of trip rates that do not remain stable over time. The impacts of such changes are evaluated from the standpoint of traditional trip generation performance measures, system level performance measures, and project level performance measures.

9.2 Literature Review

One can readily find references in the literature regarding changes in travel demand over time, and the implications of those changes. Perhaps the best source of data for understanding long term trends in travel behavior is the National Household Travel Survey (NHTS). The NHTS, previously called the National Personal Travel Survey (NPTS), has tracked the nation's personal travel and travel trends since 1969, providing information related to the number of trips, trip purpose, travel mode, and trip duration (Hu and Young 1999). NHTS data shows that between 1977 and 2001 person trips increased from 2.92 person trips per day to 4.09 person trips per day, and trips per household increased from 6.36 to 10.39 (Hu and Reuscher 2004). The biggest change was in person trips per day related to family or personal business. These trips increased from 0.91 in 1977 to 1.79 in 2001, a near doubling of the 1977 rate.

The time of day distribution of trips varies by trip purpose, but the overall time of day distribution of personal trips remained stable in the decade between the 1990 and 2001 NHTS (Hu and Reuscher 2004). One change noted in the time of day distribution of trips by trip purpose is the increasing number of non-work related vehicle trips that occur during the peak periods (USDOT 2007). After work trips and drop off trips, the next largest reason for travel during the peak period is for shopping. In the 2001 NHTS data 20 percent of all trips made during the peak period were solely to shop, excluding shopping trips made during the commute (USDOT 2007).

The literature also offers theories regarding the consequences of addressing those changes, in particular through improved modeling tools. One group of researchers used panel data from the Puget Sound Transportation Panel survey to understand the travel behavior patterns and to extrapolate trends for baby boomers (Goulias, Blain et al. 2007). The findings suggest that baby boomers will travel much as they do today, but for different reasons, displaying more behavioral diversity that will challenge planners and policy makers to provide a wider variety of services (Goulias, Blain et al. 2007). These researchers cite the need for travel models that can handle this diversity and capture with-in household interactions. In another study researchers relied on household travel surveys from Ottawa, Canada from 1986, 1995, and 2005 to study trends in travel behavior (Kriger, Wolff et al. 2011). In this study, the researchers noted population and employment shifts to the suburbs, changes in the work force, changes in mode shares, and decreases in trip rates (Kriger, Wolff et al. 2011). The purpose of the study was noted as providing information to help modelers and planners develop better analytical tools that anticipate change. Finally, a paper published by the Victoria Transport Policy Institute provides further evidence of changing demographic, economic, and market trends that will affect travel demand (Litman 2006). As with the earlier study, this study suggests that future travel demand will be increasingly diverse and good planning will require that we not simply extrapolate past trends, but seek to understand the factors that lead to change.

More difficult to find in the literature is an evaluation of how these trends might affect model parameters over time, and in turn, the affect on model outputs and performance measures. This paper attempts to address that gap with a focus on trip generation rates.

9.3 Study Methodology

Previous research has documented significant changes in trip rates over time. This research analyzed travel survey data from two different metropolitan regions covering five different points in time, three data sets from the Baltimore, Maryland region taken in 1977, 1993, and 2001, and the other two from the Triangle region of North Carolina taken in 1995 and 2006 (Huntsinger and Rouphail 2011). Datasets from both regions provided for an evaluation of changes in trip rates over the short and long-term horizon. Using the t-statistic at a 95% confidence level, findings show significant changes in total household and person trip rates between all years of the Baltimore data, and significant changes in total household trip rates for the Triangle region (Huntsinger and Rouphail 2011). Findings from this research inform the application of trip rate changes as documented below in the Trip Rate Development section, and summarized in Table 49 Percent Change in Trip Rates by Trip Purpose.

In order to assess the impact of trip rates that change over time, this analysis considers two separate case studies. The first is an advanced practice regional model and the second a traditional small community model. The analysis involves a series of model runs. The first model run is the baseline model and reflects the original trip rates for the validated model. Subsequent model runs include one model run for each of three trip rate sets. Trip Rate Set 1 (TR Set 1) reflects a change in the baseline trip rates by the percent observed between the 1995 and 2006 Triangle surveys. Trip Rate Set 2 (TR Set 2) reflects a change in the baseline trip rates by the percent observed between the 1977 and 1993 Baltimore surveys. Trip Rate Set 3 (TR Set 3) reflects a change in the baseline trip rates by the percent observed between 1993 and 2001 Baltimore surveys. Finally, Trip Rate Set 4 (TR Set 4) reflects the change observed between the 1977 and 2001 reported NHTS trip rates. All other model parameters and input data are held constant to isolate the implications of both increasing and decreasing trip rates. In this case, parameters refer to rates and coefficients in the final validated models, and input data refers to zonal demographic and transportation network data. Each trip rate set is evaluated using performance measures focused on the trip generation output, overall system performance, and specific measures for a key project in the study area.

9.3.1 Case Study 1 - Research Triangle Region, North Carolina

The Triangle Region describes the geographic area covered by the Capital Area Metropolitan Planning Organization (CAMPO) and the Durham-Chapel Hill-Carrboro Metropolitan Planning Organization (DCHC MPO), and surrounding counties. The modeled area includes 2,579 traffic analysis zones covering 3,380 square miles, and nine counties with a combined population of approximately 1.7 million. The Triangle Regional Model (TRM) is an advanced trip based model with four home-based trip purposes in the trip generation model: home-based work (HBW), home-based shopping (HBSH), home-based other (HBO), and home-based school (HBSC). The trip generation model is a generation choice model that estimates trip frequency for working adults, non-working adults, and children.

9.3.2 Case Study 2 - Jackson County, North Carolina

Jackson County, North Carolina is located in the western part of the state adjoining the Cherokee Indian Reservation and containing approximately 45 miles of the Blue Ridge Parkway (Hotaling 2011). Jackson County is a part of the Southwestern Rural Planning Organization and includes the municipalities of Sylva, Dillsboro, Webster, and Forest Hills. Jackson County is also home to Western Carolina University. The travel demand model for the region is a traditional trip based model with 51 traffic analysis zones covering 187 square miles and a population of approximately 27,300. The trip generation model is a cross-classification model of households stratified by household size and auto ownership. The Jackson model includes three home-based trip purposes: home-based work (HBW), home-based other (HBO), and home-based school (HBSC).

9.3.4 Trip Rate Development

Previous analysis on urban trends and changes in travel behavior documented trip rate changes for two urban areas between the years of 1977 and 2006 (Huntsinger and Rouphail 2011). Table 49 summarizes the rate changes documented in that study. The study analyzed survey data from the Research Triangle Region, North Carolina, to calculate rate changes between 1994 and 2006; and survey data from the Baltimore Metropolitan Commission (BMC), Maryland, to calculate rate changes between 1977 and 1993, and 1993 to 2001. Rather than applying an arbitrary percent change to the baseline rates, applying the rate changes documented in previous research allows for a comparison of what might happen based on real observations in survey data. Application of these changes to the original trip rates for the Jackson County model and the Triangle Regional model (Baseline) yield new trip rates for evaluating the impact of change on model outputs and performance measures (TR Set 1, TR Set 2, and TR Set3).

Table 49. Percent Change in Trip Rates by Trip Purpose

Trip Purpose	TR Set 1 (1994-2006)	TR Set 2 (1977-1993)	TR Set 3 (1993-2001)
HBW	-12%	20%	-18%
HBSH	11%	-11%	54%
HBSC	-16%	68%	-58%
HBO	43%	-1%	-1%

Source: (Huntsinger and Rouphail 2011)

Figure 9 demonstrates the variability of change by trip purpose, with some rates increasing over the base rate (represented as zero on the chart) and some rates decreasing from the base. This variation may dampen the impact of rate changes depending on the trip purpose demonstrating the greatest deviation from the base.

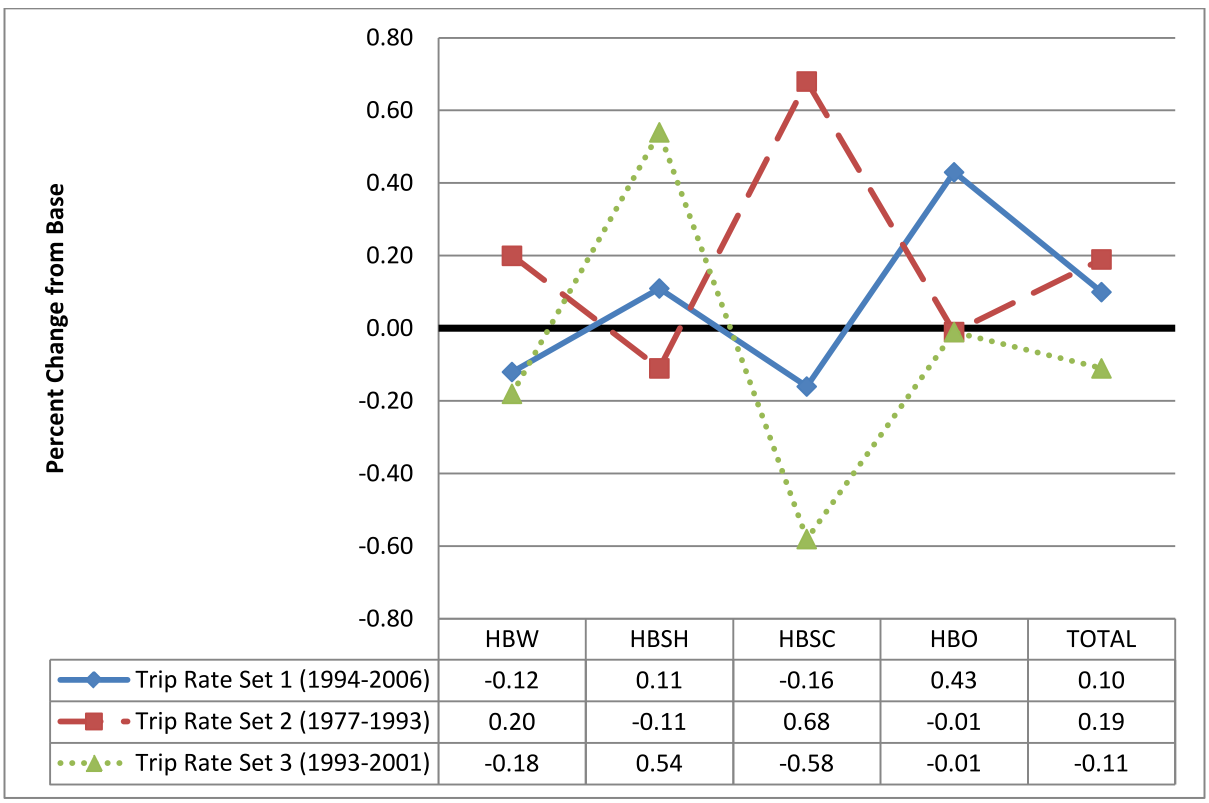


Figure 9. Percent Change in Trip Rates by Trip Purpose and Rate Set

9.3.5 Trip Generation Performance Measures

The performance measures for the trip generation reflect how changes in trip rates might affect the trip generation output. These measures include a comparison of trip productions by trip purpose, the percent of trips by trip purpose, work trips per household, work trips per worker, and trips per person. Table 50 documents typical ranges for these values cited in the literature (Schiffer and Rossi 2008).

Table 50. Typical Ranges for Trip Generation Performance Measures

Statistic	Benchmarks	
	Low	High
Percent Trips by Purpose		
HBW	12	24
HBSH	10	20
HBSC	5	8
HBO	14	28
Trips per person	3.3	4.0
Work trips per worker	1.20	1.55

Source: (Schiffer and Rossi 2008)

9.3.6 System Level Performance Measures

Since trip generation is the first step in the model process, changes in trip generation affect the remaining steps including trip distribution, mode choice, and network assignment. Comparing the overall model performance measures, such as system wide daily vehicle miles traveled (VMT), VMT by facility type, and total system demand provides an indication of the magnitude of the overall system impacts of trip generation rate changes. Table 51 documents typical ranges for VMT cited in the literature.

Table 51. Typical Ranges for VMT Performance Measures

Statistic	Large Urban Area		Small Urban Area	
	Low	High	Low	High
VMT per household	60	75	30	40
VMT per person	24	32	10	16

Source: (PB 2008; Schiffer and Rossi 2008)

9.3.7 Key Project Performance Measures

In addition to systems level impacts, travel demand models evaluate the performance of specific projects. Initially the model determines the need for the project in terms of future year deficiencies in link (project) capacity, speed, etc. Subsequent analysis using project performance measures identifies measures useful for design features such as roadway cross-section, pavement design, transit operations, and user benefits. From this perspective, project performance measures are important. The measures considered in this study include project daily demand, project congested speed, and volume over capacity ratio (VOC). Table 52 provides generalized annual average daily volumes by level of service (LOS) for the project facility types evaluated in this research.

Table 52. Generalized Annual Average Daily Volumes by Level of Service (LOS)

Facility Type	Lanes	LOS C	LOS D	LOS E
Arterial Class III/IV	4	12,600	28,200	31,900
Arterial Class III/IV	8	27,000	59,500	64,700

Source: FDOT (FDOT 2009)

9.4 Analysis and Results

9.4.1 Research Triangle Region, NC

Table 53 provides the performance measures for the Triangle Regional Model (TRM), including both system wide measures and project level measures. The project selected for analysis is Capital Boulevard/US 401 between I-440 and I-540 in Raleigh, North Carolina, an 8-lane urban arterial. This facility is a significant corridor for the state, as well as the region. Designated as a Strategic Highway by the North Carolina Department of Transportation, this route serves multiple travel purposes including local travel, commuter travel, interstate travel, and freight movement (NCDOT 2006).

The TRM system wide performance measures indicate that the impacts of the rate changes are minimal because the survey data showed rates that both increased and decreased over time for the same survey data see Figure 9. Increases for some trip purposes are offset by decreases in others; this is especially true for TR Set 2 and 3 where HBSC trips increased by 164% in TR Set 2, but HBSH and HBO trips decreased. HBSC trips make up only 11% of the original distribution of trips by trip purpose so the impact of this increase on the total trip productions is not as high as it might otherwise have been. While having only a small impact on total trips, this increase is likely to have a big impact on localized travel, especially within the vicinity of large residential neighborhoods or school locations. This could in turn have an impact on decisions made about investments in those areas. For TR Set 3 there is a 105% increase in shopping trips, but reductions in the other trip purposes. These offsetting effects have a net result of increasing overall trips by only 15% and system VMT by only 3%. The biggest change in total trips is a 41% increase for TR Set 1, due in large part to a 96% increase in HBO trips. HBO trips are the highest portion of overall trips in the baseline at 43%. The increase in HBO trip rates for TR Set 1 makes them 60% of the total trips and results in a 21% increase in system VMT.

The change in trip rates affects the distribution of trips by trip purpose and changes the estimated trip rates. Figure 10 demonstrates how the percent of trips by trip purpose changes for each TR Set and how this change measures up against the upper benchmark for each trip purpose. The percent of HBO trips for TR Set 1 far exceeds the maximum benchmark of 28%, while the percent of HBSC trips for TR Set 2 far exceeds the maximum benchmark of 8%. The percent of HBSH trips in TR Set 3 is equal to the percent of HBO trips at 28%, this value is 8% higher than the maximum benchmark for HBSH trips. As expected, given the trip production changes, the biggest impact to system VMT is 21% for TR Set 1, with a 14% increase in freeway VMT and 17-26% percent for arterials. For an area with air quality issues, this increase could mean the difference between acceptable and unacceptable levels of emissions. For TR Set 2 and 3, the VMT increases are primarily at the local street level.

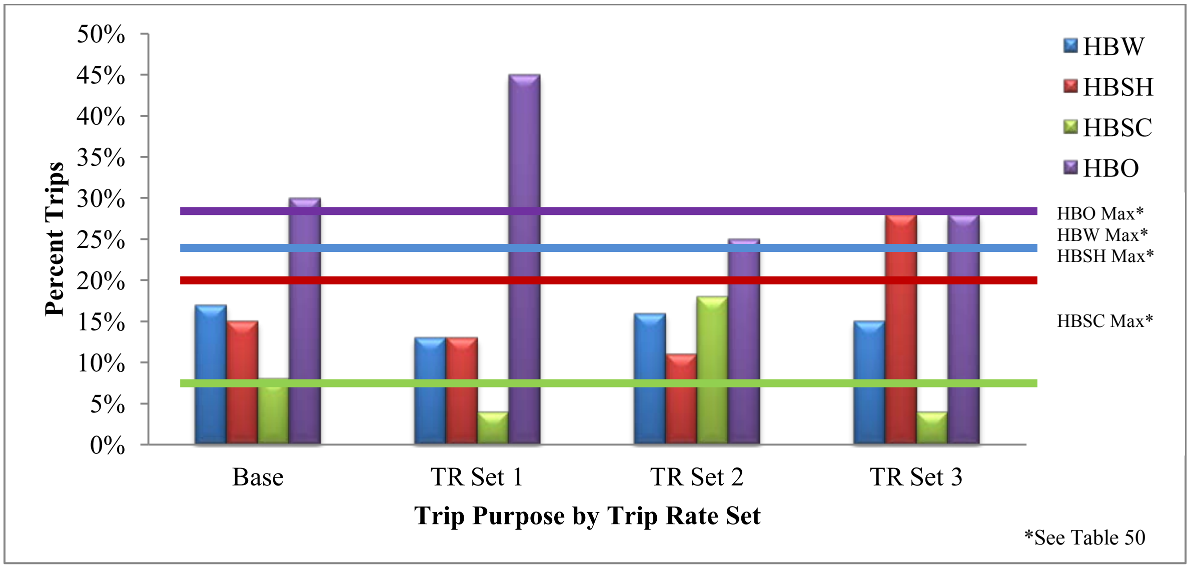


Figure 10. TRM Percent Trips by Trip Purpose by Trip Rate Set

The total trips per person and work trips per worker stay within the ranges recommended in Table 50. The VMT per person and per household is on the high end for the baseline. The rate changes in TR Set 1 move this value well above the recommended upper benchmark for the household measures.

Changes at the project level are also greatest for TR Set 1 where the forecast increases by 24% and the maximum volume over capacity (VOC) by 24%. The magnitude of the absolute difference for TR Set 1 is such that it might affect project scoping and design. Figure 11 shows the effect of these changes on the level of service (LOS). LOS D is maintained for the facility for all but the TR Set 1 that exceeds LOS E.

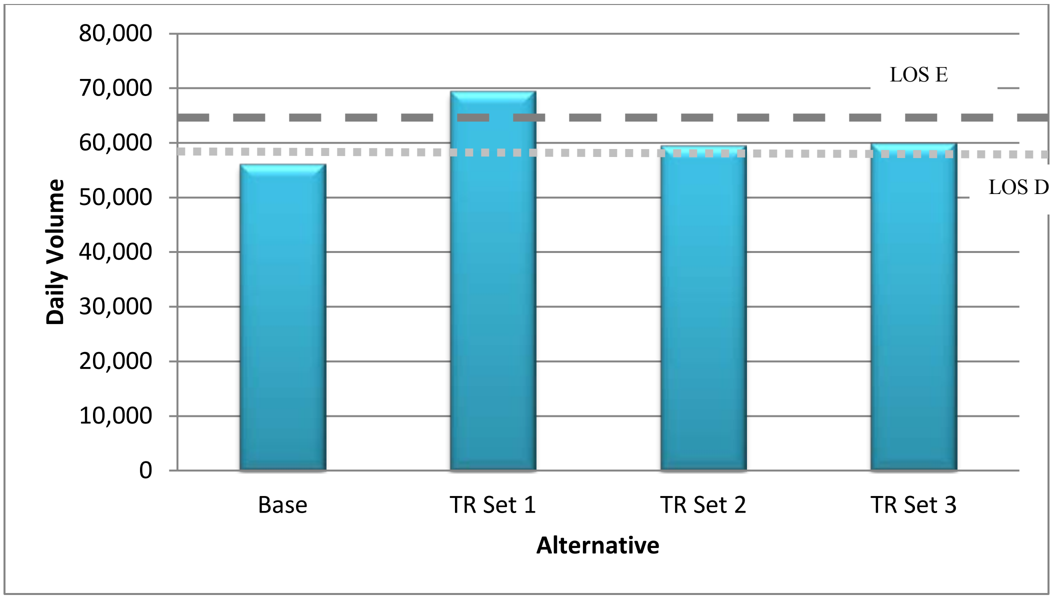


Figure 11. TRM Project Demand by TR Set and Level of Service

Trip rate changes also have an impact on transit ridership, which may influence transit investment. The rate changes in TR Set 1 result in a 56% increase in peak transit ridership and 12% increase in off peak ridership. The rate changes in TR Set 3 have a small impact on highway travel, but result in a 19% increase in peak transit ridership.

Table 53. TRM Performance Measures

Measure		Trip Rate Sets				% Change from Baseline		
		Baseline	TR Set 1	TR Set 2	TR Set 3	TR Set 1	TR Set 2	TR Set 3
Home-Based (HB) Daily Trip Productions	HBW	972,368	942,140	1,029,205	928,458	-3%	6%	-5%
	HBSH	830,752	972,827	707,599	1,704,323	17%	-15%	105%
	HBSC	448,415	325,642	1,182,126	242,455	-27%	164%	-46%
	HBO	1,700,716	3,337,193	1,674,288	1,674,489	96%	-2%	-2%
	Total HB	3,952,251	5,577,802	4,593,218	4,549,725	41%	16%	15%
% Trips by Trip Purpose	HBW	17%	13%	16%	15%	NA		
	HBSH	15%	13%	11%	28%			
	HBSC	8%	4%	18%	4%			
	HBO	30%	45%	25%	28%			
	Total HB	70%	75%	69%	75%			
Work Trips / Household		1.54	1.49	1.63	1.47			
Work Trips / Worker		1.37	1.33	1.45	1.31			
Trips / Person		2.41	3.40	2.80	2.78			
VMT / Household		78	95	83	81			
VMT / Person		30	36	32	31			
Daily VMT by Facility Type	Freeway	17,906,632	20,421,752	18,519,298	18,202,285	14%	3%	2%
	Principal Arterial	12,251,504	14,315,724	12,847,411	12,664,977	17%	5%	3%
	Minor Arterial	8,616,537	10,857,425	9,292,982	8,983,773	26%	8%	4%
	Collector	5,728,895	7,376,849	6,210,774	5,986,127	29%	8%	4%
	Local	4,832,311	6,683,206	5,356,278	5,090,407	38%	11%	5%
System Daily VMT		49,335,880	59,654,955	52,226,744	52,226,744	21%	6%	3%
System Daily Demand		177 mil	218 mil	188 mil	183 mil	23%	7%	4%
Project Daily Demand		55,986	69,395	59,436	59,889	24%	6%	7%
Project Congested Speed		35.1	28.6	34.2	33.9	-19%	-3%	-3%
Project Max VOC (hourly)		1.7	2.1	1.7	1.8	24%	0%	6%
Transit Daily Ridership	Peak	49,042	76,572	52,966	58,384	56%	8%	19%
	Off Peak	45,812	51,161	50,403	45,554	12%	10%	-1%

NOTE – Shaded values are outside typical norms

9.4.2 Jackson County, NC

Table 54 provides the reported performance measures for the Jackson County model. The project selected for analysis is NC 107 between Sylva and Western Carolina University in Jackson County. NC 107, a 4-lane urban arterial, is the only major north-south transportation corridor in Jackson County. The NCDOT Strategic Highway Corridors Vision Plan identifies this facility as an important route for mobility and connectivity to activity centers and interstates (NCDOT 2011).

Figure 12 shows that for the Jackson County model, the impact of the rate changes is not as great as observed in the Triangle model. HBO trips increase by 32% and total trips by 17% as compared to 96% and 41% respectively in the Triangle model.

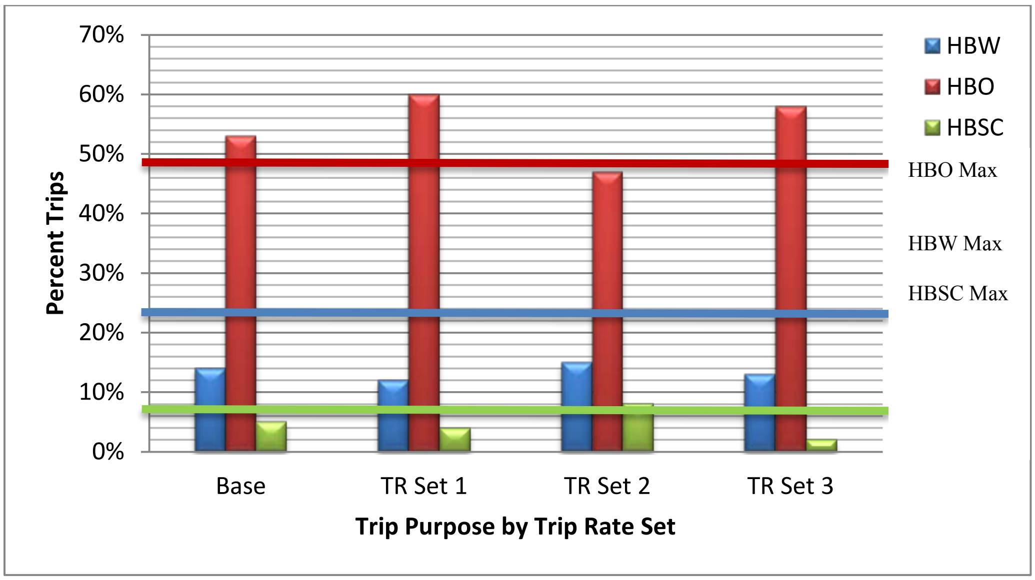


Figure 12. Jackson Percent Trips by Trip Purpose by Trip Rate Set

The trip rate changes for the Jackson County model are likely to have the biggest impact at a localized level with the 68% increase in HBSC trips of TR Set 2, and the 58% decrease in school trips for TR Set 3. This is supported by the large reported changes in VMT on local streets for all TR Sets, while the change in system VMT is less than 10%. The trips per person and work trips per worker are slightly high, but within acceptable ranges for all TR Sets, including the baseline.

The change in project demand is less than 10%, and the absolute difference in demand is unlikely to affect project scoping and design as demonstrated in Figure 13 where demand is above LOS E for all TR Sets.

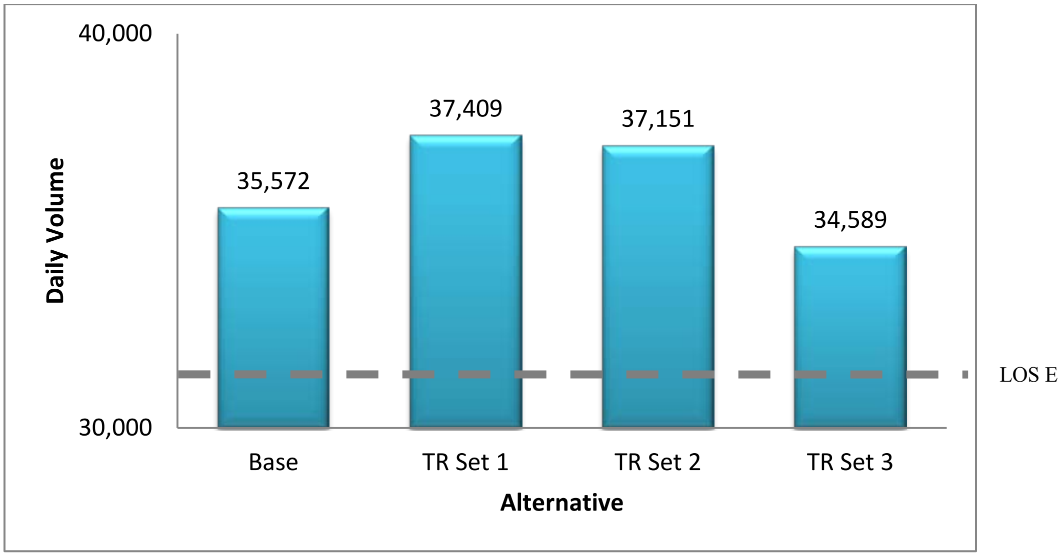


Figure 13. Jackson Project Demand by TR Set and Level of Service

Table 54. Jackson County Performance Measures

Measure		Trip Rate Sets				% Change from Baseline		
		Baseline	TR Set 1	TR Set 2	TR Set 3	TR Set 1	TR Set 2	TR Set 3
Home-Based (HB) Daily Trip Productions	HBW	21,467	18,891	25,764	17,605	-12%	20%	-18%
	HBO	81,752	107,602	81,152	81,152	32%	-1%	-1%
	HBSC	7,792	6,544	13,089	3,273	-16%	68%	-58%
	Total HB	111,011	133,037	120,005	102,030	20%	8%	-8%
% Trips by Trip Purpose	HBW	14%	10%	15%	13%	NA		
	HBO	53%	60%	47%	58%			
	HBSC	5%	4%	8%	2%			
	Total HB	72%	74%	70%	73%			
Work Trips / Household		1.49	1.31	1.79	1.22			
Work Trips / Worker		1.25	1.10	1.49	1.02			
Trips / Person		4.23	4.95	4.70	3.85			
VMT/Household		97	93	101	103			
VMT/Person		38	37	40	41			
Daily VMT by Facility Type	Multilane	910,125	945,934	939,259	885,710	4%	3%	-3%
	Arterial	147,900	155,877	154,732	142,859	5%	5%	-3%
	2-lane	139,055	148,933	144,718	134,362	7%	4%	-3%
	Collector	174,981	195,815	190,767	161,059	12%	9%	-8%
	Local	22,703	29,716	28,463	19,956	31%	25%	-12%
System Daily VMT		1,394,764	1,476,274	1,457,939	1,343,946	6%	5%	-4%
System Daily Demand		3,175,334	3,372,975	3,334,145	3,047,488	6%	5%	-4%
Project Daily Demand		35,572	37,409	37,151	34,589	5%	4%	-3%
Project Congested Speed		13.28	12.07	11.87	15.50	-10%	-12%	14%
Project Max VOC (hourly)		1.18	1.22	1.22	1.15	3%	3%	-3%

9.4.3 Implications of Temporal Instability over the Longer Term

Thus far, the analysis has focused on rate changes measured over a period of 11 to 16 years, depending on the surveys evaluated. The changes are applied to the modeled trip rates as if that were the only change that may occur over the planning horizon. Transportation planning analysis and long-range transportation plan development is typically based on a 30-year planning horizon. The next step of the analysis is to determine the extent of the longer-term impacts using trip rate changes measured over a 32-year period. To address this question, the reported change in NHTS trip rates for work trips, school trips, and personal business trips between 1977 and 2001 (see Table 4) forms the basis of adjustment to the base trip rates for the Jackson model. The Jackson model is selected due to the relative ease of updating the trip rates (rates are stored in a table, not hardcoded into the model code which requires recompiling) and quick model run time (several minutes compared to over 20 hours). The final trip rates are TR Set 4. Table 55 summarizes the results of the model run considering this change.

As expected, the magnitude of change for TR Set 4 is much greater for 32 years than that observed for the 11 to 16 year period. What is interesting is to focus on the details of that change and the possible implication for project and systems planning. The results for TR Set 4 show a 14% increase in system VMT and 84% change in VMT for local roadways; this amount of change would likely result in increased emissions and may lead to transportation project needs not anticipated based on stable trip rates. Trips per person increases to 5.88, well outside the range typically seen in observed data (Schiffer and Rossi 2008). The trip rate for work trips per worker increases slightly, but is still well within the range typically seen. Given the minimal impact of the short-term rate changes in the Jackson model as compared to those in the Triangle model, we can infer that a 32-year rate changes in the Triangle model would have a major impact on the performance measures. This change would likely result in significant increases in VMT, transit ridership, and project forecasts. Changes of this magnitude would most certainly affect planning and programming decisions and may result in misdirected investments, either away from needed highway and transit projects, or into highway and transit projects not needed.

Table 55. Jackson County Performance Measures – 32 Year Horizon

Measure		Baseline	TR Set 4	% Change from Baseline
Home-Based (HB) Daily Trip Productions	HBW	21,467	22,328	4%
	HBO	81,752	128,049	57%
	HBSC	7,792	8,024	3%
	Total HB	111,011	158,401	43%
% Trips by Trip Purpose	HBW	14%	10%	NA
	HBO	53%	60%	
	HBSC	5%	4%	
	Total HB	72%	74%	
Work Trips / Household		1.49	1.55	
Work Trips / Worker		1.25	1.29	
Trips / Person		4.23	5.88	
VMT/Household		97	111	
VMT/Person		38	44	
Daily VMT by Facility Type	Multilane	910,125	994,843	9%
	Arterial	147,900	167,484	13%
	2-lane	139,055	163,217	17%
	Collector	174,981	226,827	30%
	Local	22,703	41,692	84%
System Daily VMT		1,394,764	1,594,064	14%
Project Daily Demand		35,572	38,948	9%
Project Congested Speed		13.28	10.27	-23%
Project Max VOC (hourly)		1.13	1.60	42%

9.5 Summary of Findings

The purpose of this analysis has been to evaluate the implications of changes in trip rates over time with respect to performance measures for trip generation, system level analysis, and project level analysis. Trip rate changes were applied to two case studies: a regional travel demand model and a small community model. Three trip rate sets (TR Set) were evaluated for both models. The trip rate sets for each case study captured changes observed between five different household surveys over a period ranging from 11 to 16 years. To evaluate the effect on long term changes a fourth trip rate set was developed. This trip rate set captured documented changes in NHTS trip rates over a 32-year period.

Of the three TR Sets evaluated for the advanced practice model, TR Set 1 showed the biggest impact to the performance measures. TR Set 1 reflects a 43% increase in home-based other (HBO) trip rates, 11% increase in home-based shopping (HBSH) trip rates, 12% reduction in home-based work (HBW) trip rates, and 16% reduction in home-based school (HBSC) trip rates. The impact of these changes results in 96% increase in HBO trips, 41% increase in total trips, and a 21% increase in system wide vehicle miles traveled (VMT). For an area experiencing air quality issues, a 21% increase in VMT could result in unacceptable levels of emissions. At the project level the forecast increased by 24% given the rate changes, an increase that may likely impact project scoping and design. The system wide changes for the other two TR Sets were not as great. The reported results for TR Set 2 show a 16% increase in total trips and 6% increase in VMT, and for TR Set 3 a 15% increase in total trips and 3% increase in VMT. The localized impacts of these TR Sets will however create planning challenges at the small scale, this due to a 164% increase in school related trips for TR Set 2 and 105% increase in shopping related trips for TR Set 3. Results show that the mode choice results are more sensitive to rate changes than the VMT. Transit ridership increases 56% in the peak and 12% in the off peak for TR Set 1, and 19% in the peak for TR Set 3. Changes of this magnitude could result in over or under-funding future transit investments, and may present particular challenges for an area pursuing New Starts funding.

The reported changes in the performance measures for the traditional trip based small community model were not as great as in the regional model. In comparison, TR Set 1 resulted in a 32% increase in HBO trips and 17% increase in total trips as compared to increases of 96% and 41% in the regional model. System wide VMT changed less than 10% for all TR Sets, and 5% or less at the project level. This finding seems to suggest that the system wide travel in smaller urban area is less impacted by trip rate changes than those in large urban areas. However, localized impacts can still be considerable with increases of 68% in HBSC trips and 31% in local roadway VMT.

As expected, the longer-term affect of unstable trip rates as reported for TR Set 4 show a much greater impact for the Jackson model than those reported with the 10-year change. The impact of rate changes over a 32-year period leads to a 43% increase in total home-based trips, 14% increase in system wide VMT, 84% increase in local roadway VMT, as compared to 20%, 6%, and 31% respectively for the worst case scenario for 10-year rate changes. Based on this finding, it is likely that the impact of this change would be even greater for a regional model. Changes of this nature would likely affect planning and programming decisions and may result in misdirected investments.

9.6 Conclusions and Recommendations

Survey data shows that trip rates do change over time. How these rates change and the trip purposes most affected by that change can have a notable impact on the performance measures that transportation planners and decision makers use to guide project development and implementation. The developers and users of travel demand models cannot overcome all factors that might influence change, but a better understanding these factors may lead to the development of models that are more temporally stable, and as such, serve as better decision-making tools.

Current practice suggests collecting household travel survey every 10 years. This analysis suggests that impact of rate changes is not as great for smaller communities and may be able to collected data over longer horizon. On the other hand, this analysis suggests that larger communities will benefit from collecting new data every ten years to understand changes in travel demand such that this understanding can inform the development of travel models. This study focused on two communities in North Carolina. Future research should expand this analysis to other areas, ideally using rate changes found in historical survey data from that specific region. If those data are not available, National Household Travel Survey data could be used to inform both short and long-term rate changes.

CHAPTER 10. SUMMARY, FINDINGS AND RECOMMENDATIONS

The purpose of this research is to evaluate the temporal stability of trip generation models with a particular focus on model form and explanatory variables defining life cycle, area type, and accessibility. The key research hypothesis tested is that advanced trip generation models are temporally stable and that capturing variables defining life cycle, area type, and accessibility improves this temporal stability. In this context, advanced trip generation models refer to discrete choice models where trip generation is the selection of one trip alternative from among a set of mutually exclusive trip alternatives.

To address these questions, analysis is conducted on survey data from two different metropolitan regions covering five different points in time. Three datasets are from the Baltimore, Maryland region taken in 1977, 1993, and 2001. The other two datasets are from the Triangle region of North Carolina taken in 1995 and 2006. Datasets from both regions are used to evaluate changes in trip rates over the short and long-term horizon. The long-term analysis on the Baltimore data provides an understanding of how key indicators have changed over the longer horizon, providing insights into whether changes in the first 10 years follow a similar trajectory in the subsequent decade, or whether change increases as time goes on. While the Baltimore data allow for analysis of change over the longer-term, the Triangle data allows for a more in-depth analysis of change due primarily to the availability of supporting data useful for model development. These data include highway networks and socio-economic data for both survey years.

10.1 Summary

The research is conducted in four phases 1) data analysis, 2) model development, 3) comparative analysis, and 4) case studies.

10.1.1 Phase 1: Data Analysis

Data analysis focuses on techniques designed understand changes in trip making and the factors that influence trip making over time. Of particular interest in this phase is the investigation of explanatory variables defining life cycle, area type, and accessibility. The hypothesis tested is whether these variables influence trip making and temporal stability.

10.1.2 Phase 2: Model Development

Results from Phase 1 inform the model specification and development in Phase 2. Model investigation includes two advanced model forms: generation choice and cumulative logistic regression. Model development considers two cases, the first using explanatory variables widely used in trip generation models, and the second supplementing these variables with variables defining life cycle, area type, and accessibility. This phase seeks to test the following hypotheses:

1. Generation choice and cumulative logistic regression models have value as trip generation models;
2. Generation choice and cumulative logistic regression models are temporally stable; and
3. Temporal stability improves with the consideration of life cycle, area type, and accessibility.

10.1.3 Phase 3: Comparative Analysis

Phase 3 involves a comprehensive evaluation of the impact of survey size, model type, and explanatory variables on model performance and temporal stability. This phase seeks to answer three key questions with respect to model performance and temporal stability:

1. Does a larger survey sample size improve model performance and temporal stability;
2. Is one model form better than the other with respect to model performance and temporal stability; and
3. Does model performance and temporal stability improve with the consideration of life cycle, area type, and accessibility?

While not estimated in this research, a review of the literature provides insight into the temporal stability of cross-classification models providing an understanding into the question of whether generation choice and cumulative logistic regression models are more temporally stable.

10.1.4 Phase 4: Case Studies

The final phase seeks to address the question of relevance. Two case studies assess the impacts of changes in trip generation rates over time. The critical issue is to understand the implications of temporal instability on trip generation performance measures, system wide performance measures, and project level performance measures. The key hypothesis tested is that unstable trip rates can affect transportation investment decisions.

10.2 Findings and Conclusions

10.2.1 Phase 1: Data Analysis

Overall, the analysis shows trip rates that change over time, but there is evidence of stability depending on how the data are stratified, suggesting that the choice of variables may improve temporal stability. Temporal stability improves through the consideration of life cycle, area type, and accessibility. Results based solely on data analysis, not modeling, suggest that all three variables are useful as a means of understanding trip making behavior and improving temporal stability. The life cycle variable is best for improving temporal stability of person trips, while area type and accessibility are best for improving the temporal stability of household trips. These results inform the selection of explanatory variables for improving trip generation model temporal stability depending on whether the models are person trip models or household trip models.

Key findings from Phase 1:

1. Person trip rates are more temporally stable than household trip rates.
2. Life cycle has a strong influence on trip making behavior while also improving the stability of trip rates over time.
3. Certain definitions of area type and accessibility capture differences in trip making while also improving temporal stability of household trips only.
4. Advanced trip generation models that accommodate more independent variables may lead to improved models that better capture the dynamics that influence trip making leading to models that are more temporally stable.

10.2.2 Phase 2: Model Development

One of the advantages of generation choice and cumulative logistic regression models in comparison to cross-classification models is the ability to consider additional variables like life cycle, area type, and accessibility. The statistics reported in this research support the use of the generation choice and cumulative logistic regression models for trip generation. The results of this research also suggest that the temporal stability of generation choice and cumulative logistic regression models improves with the inclusion of life cycle, area type, and accessibility.

Key findings from Phase 2:

1. Generation choice and cumulative logistic regression models are valuable tools for trip generation.
2. Generation choice and cumulative logistic regression models are temporally stable.
3. The temporal stability of these two models improves with the consideration of life cycle, area type, and accessibility.

10.2.3 Phase 3: Comparative Analysis

Regarding model performance, the comparative analysis shows that the larger sample size is better for capturing a greater range of variables with respect to statistical significance in the model, but does not noticeably improve model verification results. Both the generation choice and cumulative logistic regression models are acceptable models based on the reasonableness of the coefficients and model statistics, while simple model statistics do not suggest that one model is appreciably better than the other is for any of the trip purposes. The introduction of life cycle, area type, and accessibility variables into the models does not noticeably improve model fit statistics, but there is evidence of improved model verification measures.

With respect to temporal stability, the detailed comparison of models shows that a larger survey sample size does not improve temporal stability. The comparative analysis also shows evidence of temporal stability for both models, but the generation choice model appears to be more temporally stable. The inclusion of life cycle, area type, and accessibility generally improves the temporally stability of the generation choice models estimated with the 2006 survey data, but not the models estimated with the 1995 survey data. The temporal stability of the cumulative logistic regression models improves with the inclusion of life cycle, area type, and accessibility for HBW and HBO models estimated using 1995 and 2006 survey data. This finding suggests improved temporal stability with the inclusion of explanatory variables for life cycle, area type, and accessibility. In comparison to findings reported in the literature on the temporal stability of cross-classification models, generation choice and cumulative logistic regression models appear to be more temporally stable, though this finding is inconclusive without the full estimation and testing of cross-classification models with the same data set.

Key findings from Phase 3:

1. A larger survey sample size does not markedly improve model performance or temporal stability.
2. The generation choice is not appreciably better than the cumulative logistic regression model for estimating home-based work and home-based other trips.
3. The generation choice model is more temporally stable than the cumulative logistic regression model, in large part due to the stratification of trip choice by workers, non-working adults, and children.
4. Model verification measures show a slight improvement with the inclusion of explanatory variables defining life cycle, area type, and accessibility.
5. The inclusion of life cycle, area type, and accessibility improves the temporal stability of generation choice and cumulative logistic regression models.

10.2.4 Phase 4: Case Studies

The results of analysis from the two case studies show that trip rates that do not remain stable over time can have implications for systems and project level planning, resulting in unexpected changes in vehicle miles traveled and associated emissions, transit ridership, project traffic forecasts, and localized travel. The analysis shows that changes in trip rates over time can lead to differences between acceptable and unacceptable levels of emissions, highway projects that either are over or under funded, and transit investments that fall either short of demand or over promise.

10.3 Recommendations for Professional Practice

The developers and users of travel demand models cannot overcome all factors that might influence change, but better understanding these factors and capturing them in the development of trip generation models can lead to the development of models that are more temporally stable, and as such, serve as better decision-making tools. The findings from this research offer several approaches for addressing this challenge and contributing to improved trip generation models through temporal stability.

This analysis informs the debate on advanced travel demand models and provides useful insights into the benefits these models provide. In general, agencies have been slow to move towards advanced travel modeling tools. A recent report documenting the state of the practice in metropolitan area travel forecasting notes that most agencies continue to use traditional four-step models that are essentially the same as they were 40 years ago (TRB 2007). When asked to document the barriers to moving towards advanced models, agencies cite resource limitations, practitioners' uncertainty as to whether new practices will be better, lack of coordination among stakeholders, and inadequate investment in the development and transfer of new techniques (TRB 2007).

The move towards advanced models does not need to encompass the full model frame; it can happen in stages, advancing one or more components at the time. Given enough evidence to support the incorporation of advanced tools into modeling practice, agencies may begin to move in that direction. One first step is advancing the trip generation model. The cross-classification model is the mostly widely used form of trip generation models, regardless of the fact that cross-classification models greatly limit the number of variables in comparison to more advanced models. Generation choice and cumulative logistic regression models accommodate more independent variables in the model specification, providing an opportunity to capture more factors that influence travel behavior and improve temporal stability.

The findings documented in this research support a move towards advanced trip generation models for improving temporal stability and minimizing forecast error related to temporal change. This research also recommends evaluating life cycle, area type, and accessibility in model specification to improve model verification and temporal stability. Finally, if budget allows, agencies will benefit from a larger sample size, but even with a small sample survey, there are benefits of moving to advanced trip generation models.

10.4 Recommendations for Future Research

This research focused on trip based models, future research in this area should explore the effect that life cycle, area type, and accessibility have on the temporal stability of tour rates as tour based modeling is a growing area of interest in the advanced modeling framework. Additionally, the focus of this research was on the temporal stability of advanced trip generation models. The promising results of this research suggest that generation choice and cumulative logistic regression models may also be good candidates for spatial transferability. An evaluation of this hypothesis would further inform the discussion on advanced models as well as offer potential benefits for spatial transferability.

Several limitations of this research effort could benefit from additional research. Data analysis included survey datasets spanning 1977 to 2006, but model estimation was limited to datasets from 1995 and 2006. The tradeoff was that a rich set of supporting data was available for these two years that allowed not only for model estimation, but also for model application. The Triangle region anticipates administering another household travel survey in 2015. Data analysis and model estimation using this data would provide insights into changes in trip making behavior over the longer term and would allow for the estimating and evaluation of models over a 20 year period. Future research should also explore the expansion of this analysis to other geographic regions to determine whether the findings will hold.

Model development focused on generation choice and cumulative logistic regression models. There may be other mathematical models useful for estimating trip generation that offer equal or better temporal stability than the ones explored in this research. Another limitation of this research was the focus on HBW and HBO trips. Future research should explore home-based shopping and home-based school trips to answer the question of whether generation choice and cumulative logistic regression models are temporally stable for these trip purposes as well. This analysis would also provide insight into these particular markets further advancing our understanding of survey data and individual travel behavior. Finally, this research focused on trip generation models. Analysis of the temporal stability of distribution and mode choice models would further improve modeling tools and increasing our confidence in model results.

REFERENCES

Acock, A. C. (2006). A Gentle Introduction to Stata. College Station, Stata Press.

Agyemang-Duah, K. and F. L. Hall (1997). "Spatial transferability of an ordered response model of trip generation." Transportation Research Part A: Policy and Practice **31**(5): 389-402.

Ashford, N. and F. Holloway (1972). "Time Stability of Zonal Trip Production Models." Transportation Engineering Journal **98**(4): 799-806.

Beige, S. and K. W. Axhausen (2008). "Long-term and Mid-term Mobility Decisions During the Life Course - Experiences with a Retrospective Survey." IATSS Research **32**(2): 16-33.

Ben-Akiva, M. E. and S. Lerman (1985). Discrete Choice Analysis: Theory and Application to Travel Demand. Cambridge, Massachusetts, MIT Press.

Bierlaire, M. (2008). BIOGEME: A free package for the estimation of discrete choice models. Ascona, Switzerland, Proceedings of the 3rd Swiss Transportation Research Conference.

BMC (2005). 2001 Household Travel Survey: Baltimore Region Analysis. Baltimore, Baltimore Metropolitan Council.

BMC. (2011). "Baltimore Metropolitan Council." Retrieved 5/17/2011, 2011, from http://www.baltometro.org/about-the-region/about-the-region.

Cambridge Systematics, I. (2007). NCHRP 20-24(55) National Summit for State and Metropolitan Agencies on Future Transportation Funding and Finance Strategies: Implications ofr Planning, Public Policy, and Institutional arrangements, AASHTO, AMPO, FHWA.

Census. (2011). "U.S. Census Bureau." Retrieved May 2011, 2011.

Cotrus, A. V., J. N. Prashker, et al. (2005). "Spatial and Temporal Transferability of Trip Generation Demand Models in Israel." Journal of Transportation and Statistics **8**(1).

Doubleday, C. (1977). "Some studies of the temporal stability of person trip generation models." Transportation Research **11**(4): 255-264.

Downes, J. D. and L. Gyenes (1976). "Temporal Stability and Forecasting Ability of Trip Generation Models in Reading." Transport and Road Research Laboratory.

Everett, J. D. (2009). An Investigation of the Transferability of Trip Generation Models adn the Utilization of a Spatial Context Variable. Civil Engineering. Knoxville, University of Tennessee. **PhD:** 122.

FDOT (2009). 2009 FDOT Quality/Level of Service Handbook. Planning. Tallahassee, FDOT.

FHWA (1975). Trip Generation Analysis. Washington, DC, USDOT.

FHWA (2007). The Transportation Planning Process: Key Issues. USDOT. Washington, DC.

Froehlich, P. (2008). "Changes in Travel Behavior of Commuters Between 1970 and 2000." Transportation Research Record: Journal of the Transportation Research Board **2082**: 35-42.

Goulias, K., L. Blain, et al. (2007). "Catching the Next Big Wave Do Observed Behavioral Dynamics fo Baby Boomers Force Rethinking of Regional Travel Demand Models?" Transportation Research Record **2014**: 67-75.

Harrison, I. (2005). An Analysis of Preliminary NHTS Results Before and After September 11. Washington, DC, Bureau of Transportation Statistics.

Harvey, G. (1980). Exploratory Analysis of the Baltimore Travel Data Set. Berkely, Institute of Transportation Studies U.C. Berkeley.

Hill, D. and N. Dodd (1966). "Studies of Trends of Travel Between 1954 and 1964 in a Large Metropolitan Area." Highway Research Record **141**: 1-23.

Hotaling, L. (2011). "Jackson County, North Carolina." Retrieved 12/26/2011, 2011, from http://www.jacksonnc.org/.

Hu, P. and T. Reuscher (2004). Summary of Travel Trends 2001 National Household Travel Survey. Washington, DC.

Hu, P., T. Reuscher, et al. (2007). Transferring 2001 National Household Travel Survey. USDOT. Washington, DC, U.S. Department of Energy.

Hu, P. and J. Young (1999). Summary of Travel Trends 1995 Nationwide Personal Transportation Survey. Washington, DC, U.S. Department of Transportation Federal Highway Administration.

Huntsinger, L., F. Nwoko, et al. (1999). Triangle Regional Model: Model Development Part 1. Raleigh, NC, North Carolina Department of Transportation.

Huntsinger, L. and N. Rouphail (2011). Temporal Stability of Generation Choice Models. Raleigh, North Carolina State University.

Huntsinger, L. and N. Rouphail (2011). Urban Trends and Changes in Travel Behavior: Implications for Travel Models. Raleigh, North Carolina State University**:** 17.

Kannel, E. J. and K. W. Heathington (1973). "Temporal Stability of Trip Generation Relations." Highway Research Record **472**: 17-27.

Kermanshah, M. (1997). "Life cycle concept: Application to trip generation procedures." Iranian Journal of Science and Technology **21**(1): 47-67.

Kitamura, R., T. Akiyama, et al. (2001). "Accessibility in a Metropolis: Toward a Better Understanding of Land Use and Travel." Transportation Research Record **1780**: 64-75.

Kitamura, R. and L. P. Kostyniuk (1986). "Maturing motorization and household travel: The case of nuclear-family households." Transportation Research Part A: General **20**(3): 245-260.

Kockelman, K. (1997). "Travel Behavior as Function of Accessibility, Land Use Mixing, and Land Use Balance: Evidence From San Francisco Bay Area." Transportation Research Record **1607**: 116-125.

Koppelman, F. S. and C. G. Wilmot (1982). "Transferability Analysis of Disaggregate Choice Models." Transportation Research Record **No. 895**: 18-24.

Kriger, D., R. Wolff, et al. (2011). Trends in Urban Travel: Implications for Planning and Forecasting. 2011 Annual Conference of the Transportation Association of Canada. Edmonton, Alberta.

Liss, S., N. McGuckin, et al. Our Nations Travel: Current Issues 2001 National Household Travel Survey. U. D. o. Transportation. Washington, DC, USDOT.

Litman, T. (2006). "Changing Travel Demad: Implications for Transport Planning." ITE Journal **76**(9): 27-33.

Litman, T. (*2011*). Land Use Impacts on Transport: How Land Use Factors Affect Travel Behavior. V. T. P. Institute. Victoria, BC, Victoria Transport Policy Institute.

Madre, J.-L. and T. Yamamoto (2009). "Long-Term Dynamics of Mobility - Introduction to the Theme of the Special Issue." IATSS Research **33**(2).

Maritz Marketing Research, I. (1993). Baltimore Metropolitan Council 1993 Household Travel Survey Data Users Guide. B. M. Council. Baltimore.

Marker, J. (2000). "Period Effects and Cohort Effects in Life Cycles." Transportation Research Record **1718**: 90-96.

McGuckin, N. and N. Srinivasan (2004). "A Walk Through Time - Changes in the Landscape of U.S. Commuting." ITE Journal.

McGurrin, M., David Greczner (2011). Performance Metrics: Calculating Accessibility Using Open Source Software and Open Data. Transportation Research Board Annual Meeting 2011. TRB. Washington, DC, Transportation Research Board.

Meurs, H. (1990). "Trip generation models with permanent unobserved effects." Transportation Research Part B: Methodological **24**(2): 145-158.

Minnesota, U. o. (2010). "Metropolitan Travel Survey Archive." Retrieved 11/13/2010, from http://www.surveyarchive.org/index.html.

Mohammadian, A. and Y. Zhang (2007). "Investigating Transferability of National Household Travel Survey Data." Transportation Research Record(1993): 67-79.

MTC. (1998, 8/26/98). "Historical and Projected Auto Operating Costs, 1990-2020." 2011, from http://www.mtc.ca.gov/maps_and_data/datamart/forecast/ass98_tab3.htm.

NCDOT (2006). US 1 Corridor Study, No. CAMPO 2005-02, Final Report. Raleigh. **CAMPO 2005-02**.

NCDOT (2011). Feasibility Study NC 107 Improvements, FS-0814A. Raleigh.

NuStats (1995). 1994/1995 Triangle Travel Behavior Survey. Austin, Texas.

NuStats (2006). Greater Triangle Travel Study: Household Travel Survey Final Report. Austin, Texas.

Ortuzar, J. d. D. and L. Willumsen (2011). Modelling Transport. Chichester, John Wiley & Sons Ltd.

PB (2007). Triangle Regional Model 2006 Documentation. Raleigh, North Carolina, Triangle Regional Model Service Bureau.

PB (2008). Small Area Travel Demand Model Guidelines. Raleigh, NC, NCDOT.

Purvis, C. (1996). "Incorporating Work Trip Accessibility in Non-Work Trip Generation Models in the San Francisco Bay Area." Transportation Research Record **1556**: 37-45.

Purvis, C. (1998). Incorporating Land Use and Accessibility Variables in Travel Demand Models. ASCE Specialty Conference on Transportation, Land Use, and Air Quality. Portland, Oregon.

Richardson, A. J. (2006). "An Alternative Measure of Household Structure and Stage in Life Cycle for Transport Modeling." Transportation Research Record **1985**: 1-11.

Ross, C. and A. Dunning (1997). Land Use Transportation Interaction: An Examination of the 1995 NPTS Data. U. S. D. o. Transportation. Atlanta, Georgia, FHWA.

Santos, A., N. McGuckin, et al. (2011). Summary of Travel Trends: 2009 National Household Travel Survey. NHTS. Washington, DC, USDOT.

Schiffer, R. and T. Rossi (2008). New Calibration and Validation Standards for Travel Demand Modeling. TRB Conference Proceedings, Washington, DC.

Smith, R. L. and D. E. Cleveland (1976). "Time Stability Analysis of Trip Generation adn Predistribution Modal-Choice Models." Transportation Research Record **569**: 76-86.

StataCorp (2007). Stata Statistical Software Release 10. College Station, TX, StataCorp LP.

Sun, Y., E. Waygood, et al. (2009). "Built Environment or Household Life-Cycle Stages—Which Explains Sustainable Travel More?" Transportation Research Record: Journal of the Transportation Research Board **2135**(-1): 123-129.

Thill, J.-C. and M. Kim (2005). "Trip making, induced travel demand, and accessibility." Journal of Geographical Systems **7**(2): 229-248.

TMIP (2004). Summary Report for the Peer Exchange on Data Transferability. TMIP Data Transferability Peer Exchange. V. N. T. S. Center. Washington, DC**:** 21.

TRB (2006). Critical Issues in Transportation. Washington, DC, National Academies.

TRB (2007). Special Report 288 Metropolitan Travel Forecasting: Current Practice and Future Direction. Washington, DC, Transportation Research Board of the National Academies.

TRB. (2010). "The On-Line Travel Survey Manual: A Dynamic Document for Transportation Professionals." Retrieved 11/13/2010, 2010, from http://www.travelsurveymanual.org/HomePage.html.

USDOT. (2007). "The Other Side of Congestion." NHTS Brief National Household Travel Survey Retrieved August 12, 2010, 2010, from http://nhts.ornl.gov.

USDOT (2009). 2009 NHTS Household File - Public Use Codebood, Version 2.1. F. H. A. U.S. Department of Transportation. DC.

Vadarevu, R. V. and P. Stopher (1996). "Household activity, life cycle and role allocation." Transportation Research Record **1556**: 15-26.

Walker, W. T. and H. Peng (1991). "Long Range Temporal Stability of Trip Generation Rates Based on Selected Cross-Classification Models in the Delaware Valley Region." Transportation Research Record **1305**: 61-71.

Wilmot, C. G. (1995). "Evidence of Transferability of Trip Generation Models." Journal of Transportation Engineering **9**: 405-410.

Yamamoto, T. (2008). "The Impact of Life-Course Events on Vehicle Ownership Dynamics - The Cases of France and Japan." IATSS Research **32**(2): 34-43.

Yee, J. L. and D. A. Niemeier (2000). "Analysis of activity duration using the Puget sound transportation panel." Transportation Research Part A: Policy and Practice **34**(8): 607-624.

Yunker, K. R. (1976). "Tests of the Temporal Stability of Travel Simulation Models in Southeastern Wisconsin." Transportation Research Record **610**: 1-5.

Zimmerman, C. (1981). A Life Cycle of Travel by the American Family. F. H. A. U.S. Department of Transportation. Washington, DC, Federal Highway Administration.

APPENDIX

APPENDIX A: EXAMPLE BIOGEME OUTPUT

BIOGEME Version 1.8 [Sat Mar 7 14:36:56 CEST 2009]
Michel Bierlaire, EPFL
This file has automatically been generated.
09/03/11 13:05:24

Multinomial Model specification file for trip production model HBW Model 4

Note that ASC1 is constrained to 0.0 and will not be estimated

Model:	Multinomial Logit
Number of estimated parameters:	8
Number of observations:	1916
Number of individuals:	1916
Null log-likelihood:	-2656.140
Cte log-likelihood:	-2312.525
Init log-likelihood:	-2656.140
Final log-likelihood:	-2298.921
Likelihood ratio test:	714.438
Rho-square:	0.134
Adjusted rho-square:	0.131
Final gradient norm:	+1.287e-002
Diagnostic:	Convergence reached...
Iterations:	4
Run time:	00:01
Variance-covariance:	from analytical hessian
Sample file:	95WRK.dat

Utility parameters

Name	Value	Std err	t-test	p-value		Robust Std err	Robust t-test	p-value	
ASC0	0.00	fixed							
ASC1	0.360	0.0855	4.21	0.00		0.0854	4.22	0.00	
ASC2	0.657	0.192	3.42	0.00		0.193	3.41	0.00	
ASC3	-0.921	0.100	-9.18	0.00		0.100	-9.18	0.00	
bInc13	-0.593	0.322	-1.84	0.07	*	0.324	-1.83	0.07	*
bLessVeh1	-0.787	0.282	-2.79	0.01		0.281	-2.80	0.01	
bMoreVeh2	0.246	0.199	1.24	0.22	*	0.199	1.24	0.22	*
bpChild1	-0.326	0.114	-2.86	0.00		0.114	-2.87	0.00	
bpnwa1	-0.310	0.147	-2.11	0.04		0.147	-2.11	0.03	

Utility functions

0	Alt0	av1	ASC0 * one
1	Alt1	av1	ASC1 * one + bpnwa1 * p_nwa + bpChild1 * pchild + bLessVeh1 * lessveh
2	Alt2	av1	ASC2 * one + bMoreVeh2 * moreveh
3	Alt3	av1	ASC3 * one + bInc13 * inc1

Correlation of coefficients

Coefficient1	Coefficient2	Covariance	Correlation	t-test	p-value		Rob. cov.	Rob. corr.	Rob. t-test	p-value	
bpChild1	bpnwa1	-0.00217	-0.129	-0.08	0.93	*	-0.00205	-0.123	-0.08	0.93	*
ASC3	bLessVeh1	-0.000345	-0.0122	-0.45	0.66	*	-0.000364	-0.0129	-0.45	0.65	*
bInc13	bLessVeh1	0.00269	0.0297	0.46	0.65	*	0.00700	0.0770	0.47	0.64	*
ASC1	bMoreVeh2	0.00209	0.123	0.55	0.58	*	0.00208	0.123	0.55	0.58	*
bInc13	bpChild1	-0.000758	-0.0207	-0.78	0.44	*	0.000139	0.00377	-0.78	0.44	*
bInc13	bpnwa1	0.000175	0.00368	-0.80	0.42	*	6.95e-005	0.00146	-0.80	0.43	*
ASC3	bInc13	-0.00810	-0.251	-0.91	0.36	*	-0.00815	-0.251	-0.90	0.37	*
ASC2	bMoreVeh2	-0.0363	-0.949	1.06	0.29	*	-0.0364	-0.950	1.06	0.29	*
ASC1	ASC2	0.000670	0.0407	-1.43	0.15	*	0.000669	0.0407	-1.43	0.15	*
bLessVeh1	bpChild1	-0.00150	-0.0467	-1.49	0.14	*	-0.00100	-0.0314	-1.50	0.13	*
bLessVeh1	bpnwa1	0.00309	0.0745	-1.55	0.12	*	0.00208	0.0503	-1.54	0.12	*
bInc13	bMoreVeh2	-0.00261	-0.0409	-2.18	0.03		-0.00677	-0.105	-2.11	0.03	
bMoreVeh2	bpnwa1	-5.26e-006	-0.000180	2.25	0.02		0.000122	0.00418	2.26	0.02	
bMoreVeh2	bpChild1	1.37e-005	0.000603	2.50	0.01		-0.000204	-0.00902	2.49	0.01	
bLessVeh1	bMoreVeh2	-0.0234	-0.419	-2.54	0.01		-0.0234	-0.419	-2.54	0.01	
ASC1	bInc13	6.00e-005	0.00218	2.86	0.00		-0.000637	-0.0230	2.83	0.00	
ASC2	bInc13	0.00238	0.0385	3.39	0.00		0.00621	0.0996	3.47	0.00	
ASC3	bpnwa1	-2.24e-005	-0.00152	-3.42	0.00		0.000172	0.0117	-3.45	0.00	
ASC1	bpnwa1	-0.00301	-0.239	3.58	0.00		-0.00293	-0.234	3.60	0.00	
ASC1	bLessVeh1	-0.00448	-0.186	3.71	0.00		-0.00441	-0.184	3.72	0.00	

ASC3	bpChild1	9.74e-005	0.00852	-3.93	0.00		0.000121	0.0106	-3.94	0.00	
ASC2	bpnwa1	4.79e-006	0.000169	3.99	0.00		-5.73e-005	-0.00202	3.99	0.00	
ASC1	bpChild1	-0.00441	-0.453	4.02	0.00		-0.00445	-0.458	4.02	0.00	
ASC2	bpChild1	-1.24e-005	-0.000568	4.40	0.00		0.000156	0.00712	4.41	0.00	
ASC3	bMoreVeh2	0.000336	0.0169	-5.28	0.00		0.000346	0.0174	-5.28	0.00	
ASC2	bLessVeh1	0.0213	0.394	5.32	0.00		0.0214	0.395	5.33	0.00	
ASC2	ASC3	0.00226	0.117	7.66	0.00		0.00225	0.116	7.64	0.00	
ASC1	ASC3	0.00256	0.299	11.57	0.00		0.00251	0.293	11.53	0.00	

Smallest singular value of the hessian: 8.96745

APPENDIX B: EXAMPLE STATA OUTPUT

```
-----------------------------------------------------------------------------------
--------------------------------------------------------
       log:  F:\Research\Data\Estimation Files\Final\Stata\1995 Best Fit KeyFac
Models on 2006 Data.smcl
  log type:  smcl
 opened on:   5 Oct 2011, 09:17:00

. ******** First Estimate 1995 HBW Best Fit KeyFac Models

. ologit hbw_t c3049na c5059na caa06 caa1922 nc30na nc55na op7079na yc30na  n_nwa
age hhwrk

Iteration 0:   log likelihood = -8619.7863
Iteration 1:   log likelihood = -7408.8327
Iteration 2:   log likelihood = -7358.1484
Iteration 3:   log likelihood =  -7357.445
Iteration 4:   log likelihood = -7357.4447

Ordered logistic regression                       Number of obs   =       8150
                                                  LR chi2(11)     =    2524.68
                                                  Prob > chi2     =     0.0000
Log likelihood = -7357.4447                       Pseudo R2       =     0.1464

------------------------------------------------------------------------------
       hbw_t |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     c3049na |   1.325995   .0930665    14.25   0.000     1.143588    1.508402
     c5059na |   .4950492   .0791372     6.26   0.000      .339943    .6501553
       caa06 |   .6354018    .077489     8.20   0.000     .4835261    .7872775
     caa1922 |   .1228547   .1732397     0.71   0.478    -.2166888    .4623983
      nc30na |   1.579416   .1573566    10.04   0.000     1.271003    1.887829
      nc55na |   1.405944   .0934063    15.05   0.000     1.222871    1.589017
    op7079na |  -1.128687   .2240261    -5.04   0.000     -1.56777   -.6896041
      yc30na |    1.67856   .1422825    11.80   0.000     1.399692    1.957429
       n_nwa |   -.809625   .0483895   -16.73   0.000    -.9044666   -.7147834
         age |   .0376784   .0013335    28.26   0.000     .0350648     .040292
       hhwrk |   .7041614   .0421701    16.70   0.000     .6215095    .7868132
-------------+----------------------------------------------------------------
       /cut1 |   2.956049   .1163824                      2.727944    3.184155
       /cut2 |   3.758005   .1191816                      3.524414    3.991597
       /cut3 |   6.696352   .1400332                      6.421892    6.970813
------------------------------------------------------------------------------

. ******** Now Estimate 1995 HBO Best Fit KeyFac Models

. ologit hbo_t nwa inc1 e1e1 e1e3 e1e4  nc55na yc30na  nc1115

Iteration 0:   log likelihood = -11355.049
Iteration 1:   log likelihood = -11232.304
Iteration 2:   log likelihood = -11232.045
Iteration 3:   log likelihood = -11232.045

Ordered logistic regression                       Number of obs   =       8150
                                                  LR chi2(8)      =     246.01
                                                  Prob > chi2     =     0.0000
Log likelihood = -11232.045                       Pseudo R2       =     0.0108
```

```
------------------------------------------------------------------------------
       hbo_t |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         nwa |   .6890235   .0508816    13.54   0.000     .5892974    .7887495
        inc1 |  -.2199505   .0793781    -2.77   0.006    -.3755288   -.0643722
        e1e1 |  -.2283376   .0774257    -2.95   0.003    -.3800891    -.076586
        e1e3 |   .0265742   .0535884     0.50   0.620    -.0784571    .1316055
        e1e4 |    .068077   .0597308     1.14   0.254    -.0489933    .1851473
      nc55na |  -.4682199   .0863856    -5.42   0.000    -.6375326   -.2989072
      yc30na |  -.2899104   .1393106    -2.08   0.037    -.5629542   -.0168667
      nc1115 |  -.3340623   .1397828    -2.39   0.017    -.6080315    -.060093
-------------+----------------------------------------------------------------
       /cut1 |  -.1635671   .0456375                     -.2530149   -.0741193
       /cut2 |   .5063173   .0459924                      .4161738    .5964608
       /cut3 |   1.772197   .0505084                      1.673202    1.871192
       /cut4 |   2.248121   .0541722                      2.141946    2.354297
------------------------------------------------------------------------------

. *********** End

. log close
       log:  F:\Research\Data\Estimation Files\Final\Stata\1995 Best Fit KeyFac
Models on 2006 Data.smcl
  log type:  smcl
 closed on:   5 Oct 2011, 09:19:59
-------------------------------------------------------------------------------------
---------------------------------------------------------
```