

Final Report for Phase I Study to Characterize the Market Potential for Non-Motorized Travel

June 2012

Prepared by
Ho-Ling Hwang
Timothy Reuscher
Daniel Wilson
Richard Schmoyer

DOCUMENT AVAILABILITY

Reports produced after January 1, 1996, are generally available free via the U.S. Department of Energy (DOE) Information Bridge.

Web site <http://www.osti.gov/bridge>

Reports produced before January 1, 1996, may be purchased by members of the public from the following source.

National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
Telephone 703-605-6000 (1-800-553-6847)
TDD 703-487-4639
Fax 703-605-6900
E-mail info@ntis.gov
Web site <http://www.ntis.gov/support/ordernowabout.htm>

Reports are available to DOE employees, DOE contractors, Energy Technology Data Exchange (ETDE) representatives, and International Nuclear Information System (INIS) representatives from the following source.

Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831
Telephone 865-576-8401
Fax 865-576-5728
E-mail reports@osti.gov
Web site <http://www.osti.gov/contact.html>

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

**Developing a Visualization-Based Sketch Planning Tool
for Non-Motorized Travel**

**Final Report for Phase I Study to
Characterize the Market Potential for Non-Motorized Travel**

**March 2012
Revised June 2012**

Prepared for
Federal Highway Administration
U.S. Department of Transportation

Prepared by
Ho-Ling Hwang, Ph.D.
Timothy Reuscher
Daniel Wilson
Richard Schmoyer, Ph.D.

Center for Transportation Analysis
Oak Ridge National Laboratory
Oak Ridge, Tennessee 37831-6283
Managed by
UT-BATTELLE, LLC
for the
U.S. DEPARTMENT OF ENERGY
under contract DE-AC05-00OR22725

TABLE OF CONTENTS

LIST OF FIGURES	ix
LIST OF TABLES	xiii
ACKNOWLEDGMENTS	xv
EXECUTIVE SUMMARY	xvii
1. INTRODUCTION	1
1.1 Background	1
1.2 Objectives	1
1.3 Organization of Report	2
2. SUMMARY OF LITERATURE REVIEW	3
2.1 NCHRP 08-78 Project Report	3
2.1.1 Factors Influencing Walking and Bicycling	3
2.1.2 Modeling and Analytic Procedures	4
2.1.3 Data Resources and Needs	7
2.2 Other References	8
3. DATA SOURCES AND OVERVIEW	13
3.1 2009 National Household Travel Survey (NHTS) Data	13
3.1.1 Demographic Characteristics of Households Taking Walk/Bike Trips	14
3.1.2 Demographics of Persons Taking Walk/Bike Trips	20
3.1.3 Land-use proxy with add-on data	25
3.2 Variables from Nielsen Claritas	26
3.3 American Community Survey (ACS) Data	28
3.4 Walk Score Data	31
3.4.1 Walk Scores for the Study Area	33
3.4.2 Transit Scores in the Study Area	34
3.5 Geospatial Data	35
3.5.1 Roadway Density	36
3.5.2 Intersection Density	37
3.5.3 Population Density	38
3.6 Other Data Sources	39
3.6.1 Center for Disease Control Obesity and Physical Inactivity Data	39
3.6.2 School Location and Enrollment Data	41
3.6.3 Square Footage Occupancy Data for Land-Use Proxy and Land-Use Diversity	42
4. ANALYSIS AND MODELING	45
4.1 Background	45
4.2 Screening at the National and NHTS Add-on Level	46
4.2.1 Review of Walk/Bike Distributions	46
4.2.2 First Stage Models	48
4.2.3 Second Stage Models	50
4.3 Data Limitations and Reduced Geographic Coverage	53
4.4 Modeling Using Data from Selected Regions (Without Nielsen Employment Data)	54
4.4.1 Review of Initial Results – from the Linear Screening Model	56

4.5 Modeling Effort with Nielsen Employment Data on Walk Trips60

4.5.1 Results Based on the Linear Model with Final Selection of Predictors60

4.5.2 Consideration of Dependency and Use of Non-Linear Predictor Variables.....61

4.5.3 Review of NHTS Coverage Issue.....62

4.5.4 Imputation of Missing NHTS Variables for Walk Trips64

5. DISCRIMINATING FACTORS67

5.1 Household Demographics67

5.1.1 Drivers.....67

5.1.2 Household/Person Counts.....67

5.1.3 Income.....67

5.1.4 Education68

5.1.5 Language and Race68

5.1.6 Vehicle Ownership.....68

5.1.7 Gender.....68

5.2 Employment68

5.2.1 Worker and Job Type.....68

5.2.2 ACS Journey to Work.....69

5.2.3 Nielsen Employment Data69

5.3 Activity Measures69

5.3.1 CDC Obesity and Inactivity.....69

5.3.2 Walkability and Transit Accessibility Measures70

5.4 Destination of Travel.....70

5.5 Other Factors70

6. EVALUATION OF THE DISCRIMINANT MODEL AND RESULTING DATABASE71

6.1 Classification Using the Discriminant Function71

6.2 Evaluation of Performance.....71

6.2.1 Selections of Training and Testing Data Sets72

6.2.2 Review of Error Rates.....72

6.3 Resulting Scores/Indices73

7. SUMMARY AND CONCLUSIONS85

7.1 OVERVIEW.....85

7.2 CHALLENGES AND LESSONS LEARNED86

7.3 FUTURE STUDIES OR NEXT STEPS89

APPENDIX A. VARIABLES OF INTERESTS BY DATA SOURCES IDENTIFIED DURING THE LITERATURE RESEARCH A - 1

APPENDIX B. VARIABLES USED OR DISCUSSED IN OTHER REFERENCES COMPILED DURING THE LITERATURE RESEARCH.....B - 1

APPENDIX C. INITIAL NATIONAL SCALED MODELS SUMMARY RESULT TABLESC - 1

APPENDIX D. STATISTICAL SUMMARY TABLES WITH THE RESCALED GEOGRAPHIC REGIONS PRIOR TO NIELSEN EMPLOYMENT DATA.... D - 1

APPENDIX E. STATISTICAL SUMMARY TABLES WITH THE NIELSEN EMPLOYMENT DATA.....E - 1

APPENDIX F. STATISTICAL SUMMARY TABLES WITH INTERACTION
TERMS INCORPORATED F - 1

APPENDIX G. FINAL RESULTS AFTER FINAL DISCRIMINANT
ANALYSIS FACTORS APPLIED G - 1

APPENDIX H. A LIST OF BLOCK GROUPS WITH THE HIGHEST 100
WALK INDICES H - 1

APPENDIX I. NHTS SAMPLE COVERAGE AND WALK INDICES IN
SELECTED STUDY REGIONS I - 1

LIST OF FIGURES

Figure ES-1. Estimated walk indices for block groups in San Diego, CA.	xvii
Figure ES-2. NHTS sample coverage and walk index in the San Diego region.	xviii
Figure ES-3. NHTS sample coverage and walk indices in Virginia near Washington DC.....	xix
Figure 1. Distribution of number of households with/without walk/bike trips by geographic location of the household.	15
Figure 2. Nine Census geographic regions used in the 2009 NHTS.	16
Figure 3. Distribution of number of households with/without walk/bike trips by race.....	17
Figure 4. Distribution of number of households with/without walk/bike trips by size.	17
Figure 5. Distribution of number of households with/without walk/bike trips by type.....	18
Figure 6. Distribution of number of HHs with/without walk/bike trips by urban status.	19
Figure 7. Number of HHs with/without walk/bike trips by workers in the household.....	19
Figure 8. Distribution of number of HHs with/without walk/bike trips by rail availability.....	20
Figure 9. Percent of persons with/without walk/bike trips by birth status.....	21
Figure 10. Percent of persons with/without walk/bike trips by driver status.....	21
Figure 11. Percent of persons with/without walk/bike trips by education level.....	22
Figure 12. Percent of persons with/without walk/bike trips by use of public transit.....	22
Figure 13. Percent of persons with/without walk/bike trips by gender.	23
Figure 14. Percent of persons with/without walk/bike trips by worker status.....	23
Figure 15. Distribution of walk/bike trips by age group.....	24
Figure 16. Percent of trips taken by the age group that is a walk/bike trip.	25
Figure 17. Use of NHTS "WHYTO" variable as a land-use mix proxy, example of a downtown neighborhood in Dallas, TX.....	26
Figure 18. Employment densities based on Claritas variable included in the 2009 NHTS.	27
Figure 19. Employment densities based on ORNL-purchased Claritas data.....	28
Figure 20. Number of commuters that walk to work by block group, with insert showing more details in the greater Washington, DC area. (Source: ACS Journey to Work data, 2005-2009.)	30
Figure 21. Number of commuters that bike to work in Census Tracts with an insert view of the greater Milwaukee, Wisconsin area. (Source: ACS Journey to Work data, 2005-2009.).....	31
Figure 22. Walk scores for the top 2,500 cities in the U.S.	32
Figure 23. Work Scores in the Dallas, TX neighborhood.....	34
Figure 24. Transit score patterns at the downtown neighborhood of Dallas, TX.	35
Figure 25. Example of roadway density in the Dallas, TX downtown neighborhood.....	36
Figure 26. Intersection density at the downtown neighborhood of Dallas, TX.....	37
Figure 27. Population density in a downtown Dallas TX neighborhood.....	38

Figure 28. Population density for children ages 5-17 that reside in the Dallas, TX downtown neighborhood.39

Figure 29. Percentage of adults who are physically inactive in 2008.....40

Figure 30. Percent of adults who are obese in 2008.41

Figure 31. K-12 school student enrollments at the downtown Dallas, TX neighborhood.42

Figure 32. Square footage data for the downtown Dallas, TX area.....43

Figure 33. Block group level land-use mix diversity for the Dallas, TX neighborhood.44

Figure 34. Distribution of block groups by walk trips frequency, based on block groups with positive walk trip rates from 2009 NHTS data.47

Figure 35. Distribution of Census tracts by bicycle trips frequency, based on Census tracts with positive biking activities from 2009 NHTS data.48

Figure 36. Distribution of block groups by frequency of walk trips (including transit access/egress) per capita.55

Figure 37. Distributions showing a hypothetical “ideal” senario with discriminant analysis model.....57

Figure 38. Distribution of the estimated probability on block groups with or without walk trips.....58

Figure 39. Distribution of the estimated probability on block groups with or without expanded walk trips (including transit access/egress trips).59

Figure 40. Distribution of the estimated probability for Census tracts with or without bike trips.59

Figure 41. Distribution of the estimated probability based on linear model with Nielsen employment data.....60

Figure 42. Distribution of the estimated probability based on model with Nielsen employment data and transformations of selected variables.61

Figure 43. Distribution of estimated probability from model with transformation and interaction terms.62

Figure 44. NHTS sample coverage for block groups within Research Triangle Region, NC.63

Figure 45. NHTS sample coverage for block groups within San Diego, CA.64

Figure 46. Imputation methods used for NHTS variables on average vehicle counts and percent drivers.65

Figure 47. Imputation method used for NHTS variables related to share of trip purpose at destinations.66

Figure 48. Classification table using the discriminant function developed from the NMT study and applied with data from test set of block groups (*observed based on NHTS data).73

Figure 49. Distribution of block group walk index categoriees by each of the nine study regions, and with the combined study region as a whole (i.e., naitonal)74

Figure 50. Estimated walk indices for block groups in Austin, TX.75

Figure 51. Estimated walk indices for block groups in Cedar Rapids, IA.....76

Figure 52. Estimated walk indices for block groups in Dallas, TX.....77

Figure 53. Estimated walk indices for block groups in Virginia portion of Washington, DC.....78

Figure 54. Estimated walk indices for block groups in Los Angeles, CA.....79

Figure 55. Estimated walk indices for block group in Madison, WI.....80

Figure 56. Estimated walk indices for block groups in Miami, FL.....81

Figure 57. Estimated walk indices for block groups in Research Triangles, NC.....82

Figure 58. Estimated walk indices for block groups in San Diego, CA.....83

Figure 59. NHTS sample coverage and walk index in the San Diego region; beige-shaded block groups contain one or more sampled households in the 2009 NHTS, those not shaded are non-sampled block groups.....86

Figure 60. NHTS sample coverage and walk indices in Virginia near Washington DC; beige-shaded areas are block groups with NHTS sampled households and non-shaded areas are non-sampled block groups.....87

Figure 61. NHTS coverage, walk indices, and rail transit stations within the San Diego, CA area.....88

Figure 62. NHTS data coverage of block groups, walk indices, and accessibility for rail transit stations in the Alexandria, VA area.....89

Figure 63. Potential region for selection of Phase 2 geography in DC area.....92

LIST OF TABLES

Table 1. NHTS Add-on Areas That Provided Data Access.....	14
Table 2. Geographic Coverage of the 5-Year 2005-2009 ACS Data	29
Table 3. Walk Score Categories.....	32
Table 4. Transit Score Categories.....	33
Table 5. Estimated Walk Trips Per Person Stepwise-Selected Logistic Regression Model	49
Table 6. Estimated Walk Trips Per Person (When Positive) Stepwise-Selected Least Squares Regression Model.....	51
Table 7. Selected Study Regions	54
Table 8. Definition of Walk Indices Category	74
Table 9. Comparisons of Total Number of Samples Captured in 2009 NHTS and 2007-2008 MWCOG Survey by Jurisdiction	91

ACKNOWLEDGMENTS

This project has benefited from the contributions of many individuals and organizations. The authors would like to specifically thank the following organizations for providing access to their NHTS Add-on data:

Maricopa Association of Governments – Arizona,
Pima Association of Governments – Arizona,
California Department of Transportation (Caltrans),
Cedar Rapids – Iowa,
Iowa Department of Transportation,
Florida Department of Transportation,
Georgia Department of Transportation,
Indiana Department of Transportation,
North Carolina Department of Transportation – Transportation Planning Branch,
Piedmont North Carolina Authority for Regional Transportation (PART),
Omaha Metropolitan Planning Agency (MAPA) – Nebraska,
Rapid City – South Dakota,
Sioux Falls – South Dakota,
South Carolina Department of Transportation,
Wisconsin Department of Transportation,
Tennessee Department of Transportation,
Texas Department of Transportation – Planning and Programming Division, and
Virginia Department of Transportation – Transportation and Mobility Planning
Division.

We wish to thank Dr. Kouros Mohammadian of the University of Illinois at Chicago for providing roadway and intersection density data at the Census tract level, as well as J. Richard Kuzmyak, Transportation Consultant, LLC and Richard H. Pratt, Consultant Inc. for allowing access to their National Cooperative Highway Research Program report for Project 08-78. We also wish to thank the National Capital Regional Transportation Planning Board at the Metropolitan Washington Council of Governments (MWCOG) for providing access to their household travel survey data.

And finally, the authors would like to thank Brian Gardner from the Office of Planning, Environment & Realty (HEP) in the Federal Highway Administration (FHWA) for funding this research. And most importantly, the authors are extremely grateful to Elaine Murakami (HEP-FHWA) for her untiring assistance and exceptional guidance for this project.

EXECUTIVE SUMMARY

BACKGROUND AND OVERVIEW

The idea of livable communities suggests that people should have the option to utilize non-motorized travel (NMT), specifically walking and bicycling, to conduct their daily tasks. Forecasting personal travel by walk and bike is necessary as part of regional transportation planning, and requires fine detail not only about individual travel, but also on transportation and neighborhood infrastructure. In an attempt to characterize the “market” potential for NMT, the Office of Planning, Federal Highway Administration (FHWA) funded the Center for Transportation Analysis (CTA) of the Oak Ridge National Laboratory (ORNL) to conduct a study. The objectives of this effort were to identify factors that influence communities to walk and bike and to examine why, or why not, travelers walk and bike in their communities.

This study relied on information collected under the 2009 National Household Travel Survey (NHTS) as the major source of data, and was supplemented with data from the American Community Survey (ACS), educational survey, health, employment, and others. Initial statistical screening methods were applied to sort through over 400 potential predictor variables, and examined with various measures (e.g., walk trip per person, walk mileage per person, bike trip per person, bike mileage per person) as the dependent variables. The best geographic level of detail used in the modeling for this study was determined to be the Census block group level for walking and Census tract level for biking.

The need for additional supplemental private data (i.e., Walk Scores and Nielsen employment data), and geospatial information that reflects land use and physical environments, became evident after an examination of findings from the initial screening models. To be feasible, in terms of costs and time, the geographic scale of the study region was scaled down to nine selected NHTS add-on regions. These regions were chosen based on various criteria including transit availability, population size, and a mix of geographic locations across the nation. Given the similarities in modeling results from walk trips and walk mileages, additional modeling efforts conducted under the later part of this study were focused on walk trips per person.

Bike models were limited only with the stepwise logistic models using Census tracts in the selected regions. Due to NHTS sampling limitations, only about 12% of these tracts have bike trips recorded from NHTS sampled households. The modeling with NHTS bike data proved to be more challenging and time consuming than what was anticipated. Along with the late arrival of Nielsen employment data, the project team had to limit the modeling effort to focus on walking. Therefore, the final modeling and discriminant analysis was conducted only for walking trips.

FINDINGS

Major factors impacting walking include:

Household demographics: drivers, household/person counts, income, education, language and race, vehicle ownership, and gender;

Employment variables: worker and job type, means of transportation to work, time leaving from home to work, trip length (in time) to work, total number of employment, as well as total retail employment;

Activity measures: obesity and inactivity, as well as walkability and transit accessibility;

Destination of travel: trip purpose in destination block group; and

Other factors: land use and urban indicator.

Examination of the performance of this discriminant function suggested promising results:

The final discriminant function with the final set of variables was developed using 75% of the records from the study data set (i.e., training data set). This function was then evaluated with the remaining 25% of the records from the study data (i.e., testing data set). Based on this result, the error rate on misclassifications was found to be at about 30%.

Walk indices:

The walk indices were computed for block groups within the study region, using the discriminant function formulation as developed under this study. An example of the walk indices for block groups in the San Diego area is presented in Figure ES-1.

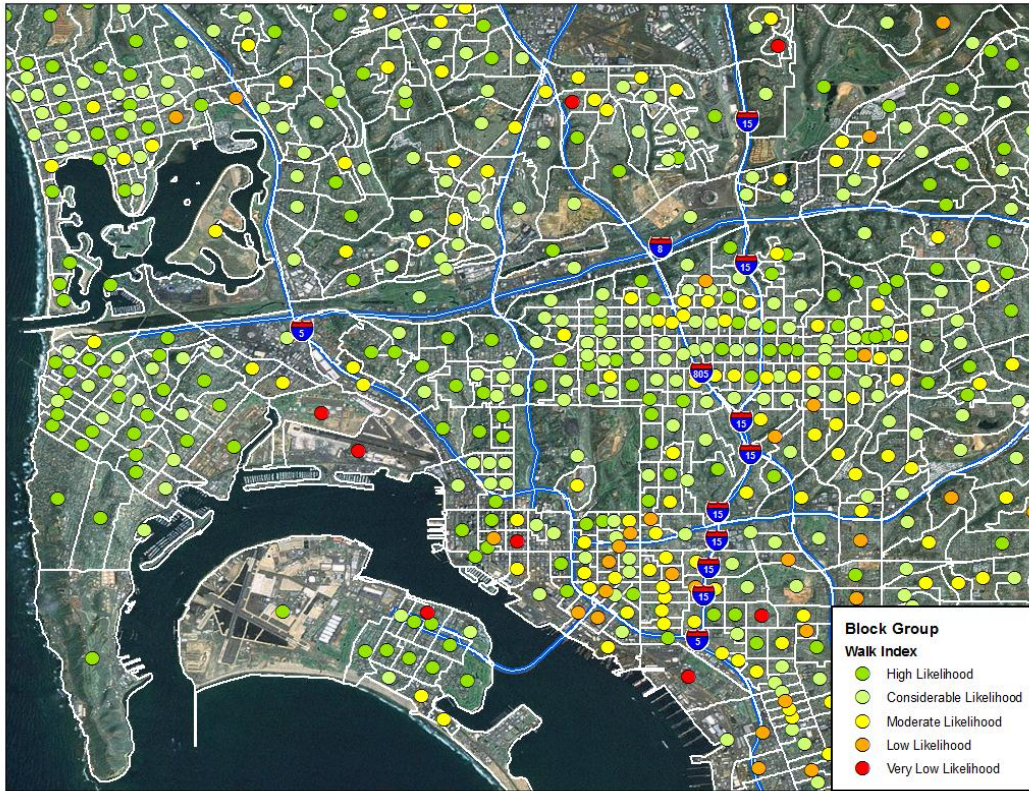


Figure ES-1: Estimated walk indices for block groups in San Diego, CA.

CHALLENGES AND LESSONS LEARNED

Developing estimates on the propensity to travel non-motorized proved to be too large a task on a national scale. The reduction of geography to nine add-on areas with better NHTS sampling made outside data acquisition more affordable and walking models possible, although the scarcity of biking data was too difficult to overcome.

As mentioned previously, sampling limitations resulted in a lack of NHTS data (i.e., sampled-households) within many block groups in the study region. This data coverage issue impacts the lower income households living in inner-city/downtown areas especially. Without a proper representation in the sampled data, the model and the associated discriminant function developed based on data obtained from sampled block groups may not necessarily capture certain unique characteristics that exist among non-sampled block groups. For example, Figure ES-2 shows the NHTS coverage in the San Diego region, along with the estimated block group walk indices. Note that, San Diego was oversampled under the 2009 NHTS, thus its coverage is much better than most other NHTS-sampled areas.



Figure ES-2 NHTS sample coverage and walk index in the San Diego region. (Beige-shaded block groups contain one or more sampled households in the 2009 NHTS, those not shaded are non-sampled block groups.)

As a comparison, sample coverage of block groups within Virginia, near Washington, DC, was much spottier (see Figure ES-3). Clearly, block groups with red dots (i.e., very low likelihood of walking) are more likely to be in unshaded areas, signifying data limitation in NHTS samples and, therefore, impacting results from the model. This also provides an indication that with better sample coverage it is possible that the developed model would perform better.

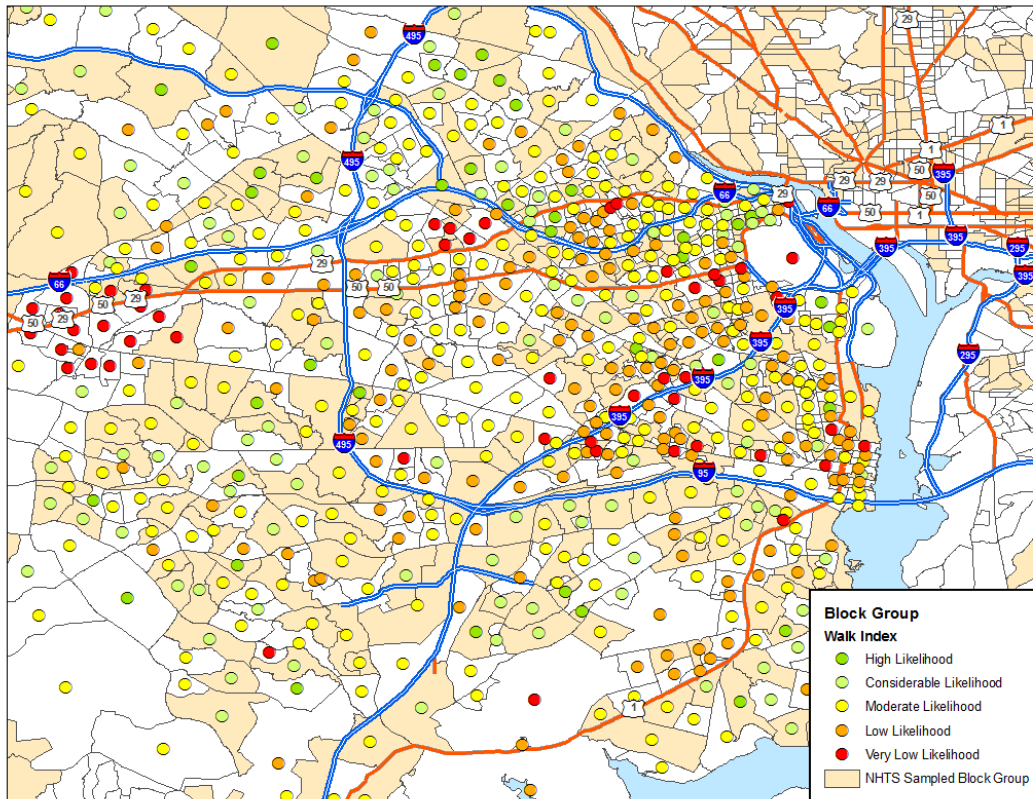


Figure ES-3 NHTS sample coverage and walk indices in Virginia near Washington DC. (Beige-shaded areas are block groups with NHTS sampled households and non-shaded areas are non-sampled block groups.)

The sampling limitation also impacted the model's ability to accurately estimate characteristics in areas known for high walking activities, such as the City of Alexandria, Virginia. In fact, there are only 42 NHTS-sampled households in Alexandria, with 20 of these households reporting walk trips which resulted in a total of 57 walk trips. Using data from such a small sample, the model was unfortunately not able to reinforce the walking activities that should have been in this specific region. Furthermore, the majority of the missing block groups were imputed by using values from their corresponding Census tracts, which is much coarser in terms of representing local characteristics. As a result, a highly walked region, such as the City of Alexandria, Virginia, was not accurately captured by the current model due to the use of largely imputed block group level data. Actually, only 34 of the total 99 block groups in Alexandria had NHTS-sampled households. That is, values of NHTS variables in about two-thirds of the block groups within Alexandria were imputed. Because of this limitation, the model was not able to produce an index to accurately reflect the high walkability within this specific area. Greater sample coverage of the region, i.e., more data, could allow better calibration of the model and an improvement of accuracy would be expected.

A national scale analysis such as that conducted in this study had the advantage of pulling all available data together to allow sufficient information in examining and identifying major factors that impact walking and biking. This data set, however, lacks the necessary

small geographic detail that is needed for a model to echo localized activities such as walking and biking. Walk and bike activities are more likely to be associated with regional characteristics and local environment. Moreover, as a national survey, the NHTS was not designed to provide coverage for tract level or block group level analyses. Thus, the determination of whether such a small geographic area is more likely to walk or not is difficult to make using this data set. The modeling framework used in this study, nevertheless, appeared to have reasonable performance (with about 30% error rate) when there is data. Although imputation methods were applied to produce estimates for missing variables so that discriminant functions could be applied to generate estimated walk indices for non-sampled block groups, these estimates (i.e., based on imputed variables) appeared to have some weaknesses.

NEXT STEPS

Future studies include the upcoming Phase 2 of this research, which would involve developing a prototype of a visualization-based sketch planning tool for non-motorized travel. Specifically, to further examine how factors such as those identified from this current study, and the modeling framework developed under this study can be applied to local/regional level planning activities, a small geographic area with more detailed local data would be necessary. Although Washington DC was not one of the 2009 NHTS additions, it conducted a household travel survey of 11,000 households in 2007-2008. The National Capital Region Planning Board at the Metropolitan Washington Council of Governments (MWCOG) conducted this household travel survey, which collected data from February 2007 through April 2008. The data coverage under the MWCOG survey, especially in the DC area, is much higher than those in the NHTS.

In addition, FHWA has information on DC's bus routes and timetables. With geo-coded information on household locations, the origin and destination of individual trips (where available), DC's bus routes, rail transit network, and walk and bike routes, such a database would provide detailed local information for conducting focused analysis that was not possible at the national level. Furthermore, the sidewalk inventory that FHWA has for the DC area, as well as specific data obtained by the ORNL research team on schools, shopping malls, crime statistics, and pedestrian and bike fatalities for the region can all be integrated into a Geospatial Information System (GIS). With this GIS database, a visualization-based decision support tool that allows local planners to examine possible impacts in walking/biking behaviors under various scenarios of change in specific demographic, environmental, or other factors can be developed. The combination of travel data and transit service profile would also facilitate the future development of a "pedestrian-friendliness" index between origin and destination pairs.

Such a visualization-based sketch planning tool could also be built with enhancements, or capabilities, to provide policy makers and planners a means to conduct informed investment decisions by predicting relative costs and benefits of potential investment scenarios. With a more focused region and increased data coverage, an improved model that takes into consideration all walk trips, including transit access/egress by walking, will also be more achievable.

1. INTRODUCTION

1.1 BACKGROUND

The idea of livable communities suggests that people should have the option to utilize non-motorized travel (NMT), specifically walking and bicycling, to conduct their daily tasks. Forecasting personal travel by walk and bike is necessary as part of regional transportation planning, and requires fine detail not only about individual travel, but also on transportation and neighborhood infrastructure.

Current transportation models lack rigor and reliability either because sufficient data may not be available, or because data integration may not have been tested to combine national data sources with local data sources. The 2009 National Household Travel Survey (NHTS) provides a robust personal travel behavior dataset with more than 87,000 walk trips (persons 16 and over), and over 6,000 bike trips (persons 16 and over) in the sample that can be used to improve estimates for these modes. The 2009 NHTS data, combined with 2005-2009 American Community Survey (ACS) data which is limited to commuting, and other data sources can benefit regional planning agencies.

This project is funded by the Office of Planning, Federal Highway Administration (FHWA), and will be carried out in two phases. Phase I of this project is to characterize the “market” potential for NMT. Based on results from Phase I, FHWA may pursue Phase II which will develop a prototype of the Geographic Information System (GIS)-based sketch planning visualization tool using a neighborhood in the Washington DC region as a case study. This report is to document the efforts conducted under the Phase I study.

1.2 OBJECTIVES

The objectives for Phase I were: (1) to identify factors that influence communities to walk and bike; and (2) to understand why, or why not, travelers walk and bike in their communities. To accomplish these objectives, ORNL conducted the project through the following tasks:

- Task 1 Conducted a literature review;
- Task 2 Identified discriminating factors of walk-and-bike friendly communities; and
- Task 3 Developed a database for estimates of personal travel by motorized and non-motorized trips.

The purpose of this report is to document findings from the Phase I study. This includes discussions of data sources, data limitations, data analysis summaries, as well as descriptions of modeling approaches taken under this NMT study. To capture lessons learned from the modeling efforts, factors considered but failed to be included in the final models are also documented in this report.

1.3 ORGANIZATION OF REPORT

The literature review is discussed in Section 2 of this technical memorandum. Because a very comprehensive literature review was recently performed as part of the Transportation Research Board (TRB) sponsored National Cooperative Highway Research Program (NCHRP) Project 08-78,¹ the discussion in Section 2 is a synopsis of that review. However, several other reports, including a second report of the NCHRP 08-78 project, were also considered. Comments on how these reports apply to this ORNL NMT project are included when applicable. In Section 3 of this report, data sources considered in this present study are discussed in more detail. Overviews of these data sources are also provided in this section. The modeling approaches considered for this study are described in Section 4 of this report. Section 4 also includes a discussion on the rescaled study region and the associated modeling efforts. A brief discussion on all discriminating factors identified based on results from the models is then provided in Section 5. Performance measure of the final model was evaluated and reported in Section 6; followed by a general summary and conclusions in Section 7 of this report.

Several appendices are also included with this report. Appendix A is a list of variables of interests for this study by data sources; and Appendix B contains a list of variables used or discussed in other references. Both lists are prepared as a part of the literature research effort under this project. Appendix C provides the summary result tables produced from the initial national scaled models. Similar summary statistics tables with the rescaled geographic regions, and produced prior to the use of additional data (i.e., Nielsen employment data), are included in Appendix D. Appendix E presents the summary result tables generated using the model that includes the Nielsen employment data. The final modeling results, adding interaction terms to the previous model, are then reported in Appendix F. Results from the final discriminant analysis based on factors identified in Section 5 are provided in Appendix G. A list of block groups with the highest 100 walk indices is found in Appendix H.

¹ See "Estimating Bicycling and Walking for Planning and Project Development, Task 1 Report, State of the Practice Review," J. Richard Kuzmyak and Richard H. Pratt, August 2010.

2. SUMMARY OF LITERATURE REVIEW

2.1 NCHRP 08-78 PROJECT REPORT

An extensive review of literature on forecasting pedestrian and bicycle travel is given by Richard Kuzmyak and Richard H. Pratt in a main report and a separate special addendum written for NCHRP Project No. 08-78 “Estimating Bicycling and Walking for Planning and Project Development.”^{2, 3} In addition to a general introduction of their project, the main report covers three major sections: (1) factors influencing walking and bicycling, (2) modeling and analytic procedures, and (3) data resources and needs. The following is a brief summary of the discussions from these three major sections.

2.1.1 Factors Influencing Walking and Bicycling

This section explores some basic characteristics of walking and bicycling, which include trip frequency, length, time, and purpose.⁴ Factors affecting walking and bicycling include:

- Demographics of pedestrians and cyclists (e.g., age, sex, automobile ownership, etc.);
- Environment (e.g., climate, weather, topography, and light condition, etc.);
- Facilities and support programs (e.g., sidewalks, street crossings, bicycle lanes, etc.); and
- Land use characteristics (e.g., density of occupants, households, motor traffic, neighborhood design, etc.).

Kuzmyak and Pratt (K&P) include discussions on walkability scores as well as safety and the perception of safety in the last category shown above. To the extent that walkability scores correlate with actual walking, these scores – or at least the underlying theory – might be useful in modeling changes in walking or bicycling behavior that occur in response to changes to the physical or social environment. Manuagh and El-Geneidy⁵ discuss this correlation between walkability indices and actual walking.

Pratt further discusses all four classes of factors affecting walking and bicycling in the special addendum report, with tables showing results calculated from various sources, including the 2007 report “Bicycling and Walking in the U.S.”⁶ which, like the special addendum, is itself a compendium of results. However, K&P do not discuss the more recent January 2010 edition of “Bicycling and Walking in the U.S.”⁷

² Ibid.

³ “Special Addendum to Task 1 Report: State of the Practice Review, Factors Associated with Bicycle and Walking Activity and Facility Usage,” prepared by Richard H. Pratt, September 2010.

⁴ K&P do not discuss consequential characteristics such as fuel or carbon savings or numbers of accidents.

⁵ Manuagh, K. and El-Geneidy, A., “Validating walkability indices: How do different households respond to the walkability of their neighbourhood?” October 2010, tram.mcgill.ca/Research/Publications/walkability.pdf

⁶ Thunderhead Alliance, “Bicycling and Walking in the U.S.: Benchmarking Report 2007,” Washington, DC, August 2007. www.peoplepoweredmovement.org/site/index.php/site/memberservices/C529.

⁷ Alliance for Biking & Walking, “Bicycling and Walking in the U.S.: 2010 Benchmarking Report,” Washington, DC, January 2010. www.peoplepoweredmovement.org/site/index.php/site/memberservices/C529.

2.1.2 Modeling and Analytic Procedures

In this second section, K&P classify models and procedural techniques that could be used to address bicycling and walking demand into four categories:

- Traditional four-step travel forecasting models,
- Sketch planning methods,
- Micro-simulation methods, and
- Direct-demand approaches.

These categories, which can overlap, differ in specificity and in their basis in either simulation or statistical estimation.

Four-step Forecasting Models

Generally, the four steps in the traditional four-step forecasting models are: trip generation (and attraction), trip distribution, modal split, and traffic assignment. K&P point out that an increasing number of Metropolitan Planning Organizations (MPOs) are developing separate models of “auto ownership” that are more sensitive to their regional characteristics, and then using those outputs to better inform the trip generation model. Clearly, auto ownership is not applicable to pedestrian travel, and at least from a cost standpoint, probably not very important for bicycle travel either, although “No access to a bicycle” is listed as the primary reason for not bicycling in an NHTSA report.⁸ As K&P point out,

Unfortunately, almost no four step models in current practice have been designed to incorporate non-motorized travel in any credible detail relative to motorized modes...Perhaps the ultimate limiting characteristic of traditional four-step models for incorporating non-motorized travel is the geographic aggregation of traffic analysis zones...Virtually all walking trips – and probably the majority of bike trips – are likely to begin and end within the same traffic analysis zone. This essentially makes them invisible to the four-step process...

Of course, the four-step model logic might be applied at a finer geographic level to walking and bicycling. Nevertheless, the suggestion here is that the four-step approach is probably not appropriate for the NMT study in general, or the current ORNL NMT project in particular.

Sketch Planning Methods

Sketch planning models make use of GIS-based⁹ tools to relate land use (including density, diversity, and design), demographics, environmental, transportation-system and other characteristics to household travel. Most conditions are determined through certain buffering processes. For example, an area between 0.25 to 0.5 miles was typically assumed to be the

⁸ “National Survey of Bicyclist and Pedestrian Attitudes and Behavior,” NHTSA, August 2008, available at: www.nhtsa.gov/DOT/NHTSA/Traffic%20Injury%20Control/Articles/Associated%20Files/810971.pdf.

⁹ Although the title of this ORNL project is “Developing a Visualization-Based Sketch Planning Tool...” which suggests a sketch planning approach, the term as applied actually refers to Phase II of the project.

reasonable walking distance for the household. Specific tools (i.e., models) discussed in the NCHRP08-78 report are INDEX¹⁰, I-PLACE3S¹¹, and ENVISION Tomorrow¹².

Note that, although GIS is pervasive in both the data and analysis, it typically needs to work in tandem with other regional travel forecasting or demand models (or other models if non-travel related) so that clear and meaningful impacts can be reflected, and proper interpretation of relationships among the data/information being visualized can be conveyed. A GIS tool is commonly used as a pre-processor for examining potential relationships among factors (i.e., variables or data elements) or as a post-processor to identify necessary modifications needed in the forecasting or demand models (which could be developed based on standard least squares regression, for example).

Microsimulation Models

Microsimulation models, according to K&P, represent “essentially a melding of the best characteristics of the traditional trip-based network models and the highly detailed parcel or grid cell level land use models.” As the name suggests, microsimulation models tend to be specific and local, as opposed to general and national. National-scope generic models would typically be based on statistical analysis (e.g., a regression model), whereas more specific local-scale models might be based on microsimulation. Because of this, the microsimulation modeling approach is more suitable for the NCHRP08-78 research and will not be appropriate for our current NMT research project.

In the current ORNL NMT project, specifically in the Phase I study, our focus is on general statistical estimates of pedestrian and bicycle volumes (or shares) as a function of model parameters (i.e., factors). The geography detail will begin with a national level scope, moving down to subsets of the nation, and later to more specific, selected, geographies (e.g., NHTS additions).

Direct Demand Approaches

Direct demand models are used to forecast trips or volumes directly, “rather than determining what walk or bike demand might be in a particular situation as a result of a sequence of steps that winnows out non-motorized travel from a sequential set of ‘choices’ between destinations, modes, by type of traveler, trip purpose, etc.,” as stated in the NCHRP08-78 Task 1 report. In the direct demand approach, travel (trips, miles, travel time, etc.) is modeled directly as a function of its demographic, environmental, and other predictors. Linear or loglinear regression is often the approach of choice for fitting these models. According to K&P, “what these tools may lack in formal sophistication, they make up for in convenience, realism, and response time.”

¹⁰ Smart Growth INDEX, which is a GIS-based tool developed for, and is distributed by, the U.S. Environmental Protection Agency.

¹¹ A GIS-based land use sketch planning model, originally developed under the sponsorship of U.S. Department of Energy, the California Energy Commission, and Caltrans. Currently, the model is managed by the Sacramento Area Council of Governments (SACOG).

¹² The Envision Tomorrow is another GIS-based tool developed by Fregonese Association for the Southern California association of Governments (SCAG) in Los Angeles.

Two types of direct demand models are discussed in K&P's NCHRP08-78 Task 1 report, the Latent Demand Model and the Enhanced – Validated Demand Model. The latent demand models produce an index (for walking or bicycling) that is a function of several geographic variables (e.g., population density, intersection density). The enhanced – validated demand approach utilizes regression models to estimate volumes of bicycle or pedestrian travel. Data on bicycle and pedestrian counts are utilized along with many exploratory variables in the linear regression model. Thus, in the K&P taxonomy of models, direct demand models are the most likely candidates for the modeling approaches that could be used in the Phase I study of the ORNL NMT project.

Specific examples of direct demand model formulations referenced in the NCHRP 08-78 report¹³ that might have some relevance to the present ORNL project include:

- **Total PM peak pedestrian intersection crossings** = $3.217 \times 10^{-3} * \text{Employment Density} + 3.675 * \text{PM Bus Frequency} + 82.695 * \text{Neighborhood Shopping District Proximity} + -6.855 \times 10^{-3} * \text{Distance from Ocean} + -5.699 * \text{Average Speed Limit of Approach Street} + 222.18$

This model is an enhanced-validated demand model developed for the City of Santa Monica, California by Fehr & Peers Transportation Consultants. Pedestrians were counted each time they crossed a leg of the intersection within the study area during the period from October 2007 to June 2009. The model establishes a relationship between the number of walkers (pedestrian) and area activities (e.g., employment, bus, shopping, etc.).

- **Total pedestrian intersection crossings per week** = $0.928 * \text{Total population within 0.5 miles} + 2.19 * \text{Total employment within 0.25 miles} + 98.4 * \text{Number of commercial retail properties within 0.25 miles} + 54,600 * \text{Number of regional transit stations within 0.10 miles} - 4910$

This is a demand model developed for Alameda County, California by researchers of the Safe Transportation Research & Education Center (SafeTREC) at UC-Berkeley. Similar to the previous model, pedestrians were counted each time they crossed a leg of the intersection. This study, however, also considered the distance to crosswalk, thus pedestrians crossing the intersection within 50 feet of the crosswalk are counted. The data collection period for this specific study was from April through June of 2008. In addition to variables similar to those included in the Santa Monica model (e.g., employment, shopping, and transit), this model also takes into account the nearby total population (i.e., residents within 0.5 miles).

- **Weekday pedestrian intersection crossings between 2:30 p.m. and 6:30 p.m.** = $0.792 * \text{Presence of bike lane at intersection} - 0.230 * \text{Employment density within 1/4 mile} - 1.107 * \text{Residential land use within 1/16 mile} + 1.508 * \text{MUNI stop density within 3/8}$

¹³ “Table 7 Examples of Existing Pedestrian Volume Models” on page 36 of NCHRP 08-78 Task 1 Report. The original source of these models is the “Summary of Existing Pedestrian Volume Models” report written by R. Schneider, University of California, Berkeley, April 2010.

$$\text{mile} + 0.013 * \text{Population density within } \frac{1}{2} \text{ mile} - 0.262 * \text{Mean slope within } \frac{1}{16} \text{ mile} + 0.047 * \text{Patch richness density within } \frac{1}{16} \text{ mile} + 0.144$$

Liu & Griswold of San Francisco State University developed this model to estimate the total of pedestrian crossing at intersections during 2:30-6:30 p.m. on typical weekdays in the City of San Francisco, California. Data used for this study was collected during May, June, August, and September of 2002. In addition to factors associate with population, employment, and access to transit, this model also included variables such as presence of bike lane at the intersection, patch richness density, and mean slope.

- **Number of pedestrians approaching intersection on weekday between 7 a.m. and 7 p.m. = 54.0 * Total population within 0.25 miles + 18.1 * Total employment within 0.25 miles + 0.0530 * Urban residential area within 0.25 miles – 0.0260 * Mixed land use within 0.25 miles + 13.7 * Number of bus stops within 0.25 miles – 7.61**

Pulugurtha and Repaka of the University of North Carolina in Charlotte developed this pedestrian demand model for the City of Charlotte, North Carolina. Variables used in this model are very similar to the abovementioned California examples. Unlike the others, however, this Charlotte study focused on signalized intersections only.

Although all of these regression-based models accounted for population density, employment density, and transit accessibility, the measure of factors and their weights varies from model to model. Each model is unique in the sense that it was developed specifically for a given study location. Because of this, generalization of the functional form and/or transferability of these models to other regions may not be appropriate.

The latent demand models, on the other hand, produce an index that is a function of several geographic variables. Which variables are used in the model depends on what is available for the given jurisdiction. For the Santa Monica study mentioned above, Fehr & Peers also developed qualitative and validated latent demand models for their direct demand estimation. For each intersection, the variables (e.g., population density, intersection density) are calculated; each variable then becomes a component of the overall latent demand index. The index, which is calculated using weightings based on factors and engineering judgment, is a relative measure of an area's potential to produce bicycle or pedestrian volumes. Thus, for example, if a jurisdiction feels that more emphasis needs to be placed on facilities (e.g., bicycle path) rather than population density in the index, the adjustment can be made (e.g., place a higher weight on the facilities).

A bicycling model parallel to the pedestrian model was also developed under the Santa Monica study. Significant variables in this bicycling model include: land use mix, PM bus frequency, population density under 18, and bicycle network proximity.

2.1.3 Data Resources and Needs

The primary data identified by K&P to support the analytic tools described in the NCHRP08-78 report are results of household travel surveys, intercept and market surveys, and user volume counts. K&P further elaborate on two categories of data that are important in modeling and

planning, behavioral data and external/environmental data. These types of data are briefly summarized below.

Behavioral Data

This category includes data collected from:

- Household travel surveys: This includes regional household travel surveys and the national NHTS. These could be activity-based or trip-based surveys.
- Intercept and user surveys: Several types of surveys are included under this category: facility intercept surveys, outreach and awareness surveys, public opinion and preference surveys, and self-reported information and perceptions.
- Facility count and survey data: This type of data is widely accepted by transportation planners in measuring facility use as well as in validating travel demand models. Specific areas discussed by K&P under this category include: national perspectives, effects of exogenous events, natural or artificial NMT volume variability, and the lack of NMT facility survey instrument consistency.

External and Environmental Data

This type of information (e.g., location, setting and physical environment) is best represented through the use of GIS tools. Examples of measures derived using GIS include:

- Density (of population, households, employment, etc.);
- Diversity (e.g., land use mix);
- Design (e.g., street density, intersection density, connectivity); and
- Destination Accessibility (e.g., transit accessibility, number of jobs).

Pratt also discusses data and challenges in data collection in the special addendum report. Note that many of the studies referred to in the K&P report and the special addendum are local in scope. Though they may involve statistical sampling, they are case studies, anecdotal from a national perspective. Other national surveys include the “National Survey of Bicyclist and Pedestrian Attitudes and Behavior,” which is a national survey conducted in 2002 for NHTSA.¹⁴ The Alliance for Biking & Walking (mentioned above) compiles data from other sources, but also seeks to supplement it with additional data from the fifty states and fifty-one largest cities. Nevertheless, the NHTS and ACS data are considered the best sources of national-scope data about NMT. More detailed discussions on potential data sources for the current ORNL NMT study is included in Section 3 of this technical memorandum.

2.2 OTHER REFERENCES

No attempt is made under the current ORNL NMT project to duplicate K&P’s extensive literature review. Nevertheless a few additional references should be mentioned. A major one is a recent report prepared by the Transportation Planning Studio of the University of Washington’s

¹⁴ “National Survey of Bicyclist and Pedestrian Attitudes and Behavior,” NHTSA, August 2008.

Department of Urban Design and Planning¹⁵ (referred to as the Planning Studio from here on). This report contains a comprehensive review of best practices related to bicycle policy, facility design, and data collection. Primary indicators selected and used in the Planning Studio's bicycle count model are: time of day, season, population and employment densities, mix of use, bicycle facility type, traffic volume, rain and temperature, income, and age. The authors also give a good example of a fitted regression model for bicycle level of service (LOS):¹⁶

Bicycle LOS =

$$0.507 \ln(\text{Vol}_{15}/L) + 0.199 SP_t(1+10.38HV)^2 + 7.066(1/PR_5)^2 - 0.005 W_e^2 + 0.760$$

where

Vol_{15} = volume of directional traffic in 15-minute time period

L = total number of through lanes

SP_t = effective speed limit = $1.1199 \ln(SP_p - 20) + 0.8103$, SP_p is posted speed

HV = percentage of heavy vehicles

PR_5 = FHWA's 5-point surface condition rating (5 = best)

W_e = average effective width of outside through lane = $W_t + W_1 - \Sigma W_r$

W_t = total width of outside lane and shoulder/parking pavement

W_l = width of paving from outside lane stripe to pavement edge

ΣW_r = width reduction due to encroachments in outside lane

This approach is similar to the regression modeling method that ORNL is considering for bicycle travel analysis in the Phase I study.

Also discussed in the Planning Studio's report are the lack of and the need for year-round traffic counts for bicycles. Many case studies are also summarized in the report. The focus in many of the studies is on simultaneously considering both travel demand and travel safety:

“Cambridge, MA is an example where long-range policy and political will has created a setting in which conflicts involving cars, pedestrians, and bicyclists have remained static in spite of increased miles traveled¹⁷...”

“...careful consideration for existing conditions in the city has allowed Chicago to increase facilities while measuring safety and progress.¹⁸”

Various reports on walk scores, including the report of Manuagh and El-Geneidy¹⁹ are also particularly relevant to the current ORNL NMT project. Not only can these walk scores, or analogs for bicycle travel, be used as predictor variables in the models being considered, but also the scores themselves are computed with models related to travel. Although the very local nature of many of these models makes them less suitable for Phase I of the ORNL NMT project, they

¹⁵ “Bicycle Planning, Best Practices, and Count Methodology” Ginger et. al., The Transportation Studio, Department of Urban Design and Planning, University of Washington, 2011.

¹⁶ Ibid, page 16.

¹⁷ Ibid, page 22.

¹⁸ Ibid, page 30.

¹⁹ Op. cit.

might be incorporated under the Phase II study. More about walk scores, including its methodology report,²⁰ can be found at www.walkscore.com. A brief discussion on walk scores is also included under Section 3.4 or this report.

Furthermore, the NCHRP 08-78 Task 3 report, entitled “Estimating Bicycling and Walking for Planning and Project Development (Task 3 Report)” and written by J. Richard Kuzmyak et. al., elaborates on limitations of direct demand models. The authors argue that direct demand tools, as a principle shortcoming, do not deal with the non-linearity of travel flows in a network. Their point is that a given bicycle or pedestrian traveler has multiple options to consider when deciding whether he/she will use a particular facility if it is provided or if it is substantially improved. Projecting demand for a given facility (path, intersection) without considering its relationship to the network of possible paths would be equivalent to trying to optimize the timing of a single traffic signal in a street grid without accounting for how that change will impact flows elsewhere in the grid. Methods that try to predict bicycle or pedestrian activity levels for a given facility or intersection without accounting for redistribution in the network are, therefore, making a major—and probably dubious—assumption about the independence of that link or node²¹.

Another potential drawback of direct demand models, the issue of correlation versus causation, is also considered in the NCHRP Task 3 report. The authors of that report stated that “... *these (direct demand) models are not particularly effective planning tools. Their structure is not sound in terms of having causal links to the modeled behavior, so estimates of response to major changes in the environment that would affect walking or biking—proximity of attractive destinations, efficient, safe and connecting networks, changes in the cost of driving—cannot be reliably made.*” Bike lanes and bicycling, bike ownership and bike use, and walking to work and income are mentioned as specific examples in which causation and correlation might be confused. As an example of the correlation and causation discussion, they point out that a negative correlation between walking to work and income may have as much to do with the location of neighborhoods where lower-income households dominate than with the predilections of lower or higher income commuters.

Note that, however, the current ORNL NMT modeling effort is not to produce demand models for walk/bike travel, nor to predict utilizations of walk/bike facilities. Rather, the ORNL NMT project is aiming at identifying factors that best associate with walk/bike behaviors within a small geographic region (e.g., block group or Census tract). Specifically, factors that distinguishing those who do walk and those who do not within a given community (e.g., a block group). The goal is that, by understanding which factors are likely to influence the walk/not walk (or bike/not bike) activities, local planners might be able to develop strategies that allow improvements in certain factors that encourage more walking/biking. Lessons learned from the NCHRP 08-78 study is clearly beneficial to the ORNL NMT study. The drawback of direct demand models as identified by the authors of this NCHRP may not be applicable.

On the other hand, the scope of the Phase II study for this ORNL NMT project is to “develop a prototype of a Geographic Information System (GIS)-enabled and visualization-based sketch planning tool that will inform a community’s decision on the development and/or improvement

²⁰ See <http://blog.walkscore.com/wp-content/uploads/2010/12/WalkScoreMethodology.pdf>.

²¹ NCHRP 08-78 Task 3 Report, page 5-11

of non-motorized facilities.” This suggests that users of the sketch planning tool will be investigating changes in NMT due to specific changes contemplated at specific locations. A challenge in Phase II of ORNL's project will be to determine how to incorporate specific local detail input through a GIS interface into an analysis based on a regression-based model, in order to predict specific changes at specific locations. The correlation-versus-causation issue, as pointed out by Kuzmyak et. al., might be a challenge for the ORNL NMT Phase II project.

The NCHRP 08-78 Task 3 report lists numerous potential predictor variables for walking and bicycling. These are combined with variables which were identified from other references and included in Appendices B of this report. Because these variables are largely similar for both walking and cycling, and were also considered as potential predictor variables for NMT activities in the current ORNL NMT study, only one list is compiled. Of course, the extent to which the predictor variables affect cycling and walking (e.g., the elasticity) tends to differ. (Possible TRB references can be added here)

3. DATA SOURCES AND OVERVIEW

Major data sources used for the NMT study are the 2009 NHTS data from FHWA and the ACS data from the U.S. Census Bureau. Additional geospatial data variables were generated by the project team using GIS tools and information collected from other data sources. These data sources and overviews of selected variables are discussed in this section.

3.1 2009 NATIONAL HOUSEHOLD TRAVEL SURVEY (NHTS) DATA

The 2009 NHTS is an FHWA-sponsored national travel survey of over 150,000 households.²² According to the NHTS website²³ “The NHTS is the authoritative source of national data on the travel behavior of the American public.” The NHTS includes questions about trip frequency, distance, and travel time, and modes of transportation, including walking and bicycling.

Although the NHTS is a national survey, its sampling is more concentrated in areas with add-on components, which include individual participating states, cities, and regions.²⁴ Sampling weights provided in the NHTS data allow for proper accounting for the add-on sampling, as well as non-response and other adjustments. Of course, estimates (e.g., trips per person) computed for areas with add-on sampling do tend to be statistically more precise than estimates from areas without add-on sampling.

Census Tract and Block Group variables collected by Nielsen Claritas describe neighborhood characteristics of the NHTS samples. Considering that there are over 60,000 U.S. Census Tracts and over 200,000 Block Groups in the U. S., with a total of 150,000 households in the NHTS samples, this is equivalent to an average of fewer than 2.5 households per Census Tract that are included in the NHTS samples; while fewer than 0.75 households per Block Group are sampled in the NHTS. Thus, except perhaps for NHTS add-on areas which have much higher sampling rates, the NHTS data is not immediately suited to Census-Tract-specific or Block-Group-specific analyses. To the extent that the NHTS is a representative survey, however, it could be used to estimate characteristics that are consistent across classes of multiple Census Tracts or Block Groups.

The NHTS add-on areas not only have higher sampling rates, they also include more detailed information on destinations (i.e., geo-coded locations). Since geographic detail is a critical element in determining the ability, or access, to walk and bike facilities, special requests were made to every add-on agency for approval to use their add-on data. With each add-on agency that approved the request, the ORNL team signed a nondisclosure agreement which allows the team to use the private add-on data for this study. Add-on areas that permitted this project to benefit from their geographic detail data (aggregated at block group centroid level) are listed in Table 1 below.

²² U.S. Territories are not included in the NHTS sampling frame.

²³ <http://nhts.ornl.gov/>.

²⁴ The 2009 NHTS add-on areas were California, Cedar Rapids, Iowa, Florida, Georgia, Iowa, Indiana, North Carolina, New York, Omaha, Nebraska, Phoenix, Arizona, Piedmont Region, North Carolina, South Carolina, South Dakota, Tennessee, Tucson, Arizona, Texas, Virginia, Vermont, and Wisconsin.

Table 1. NHTS Add-on Areas That Provided Data Access

State	Agency
Arizona	Maricopa Pima
California	Cal Trans
Florida	FLDOT
Georgia	GADOT
Indiana	INDOT
Iowa	Cedar Rapids IADOT
North Carolina	NCDOT Piedmont-Authority
Nebraska	Omaha-MAPA
South Carolina	SCDOT
South Dakota	Rapid City Sioux Falls
Tennessee	TNDOT
Texas	TXDOT
Virginia	VADOT
Wisconsin	WDOT

Instead of the entire NHTS data set (i.e., national level), the geographic scope of the following data analysis and reviews were based on the NHTS add-on areas. Furthermore, with the focus of this project being walk and bike travel, data examined and summarized in this study report are mainly on those associated with walk/bike trips.

As stated in the objective of this report, the goals of this project are (1) to identify factors that influence communities to walk and bike and (2) to understand why, or why not, travelers walk and bike in their communities. Thus the main focus of this project is on examining characteristics of a neighborhood (i.e., Census block group or tract), rather than on the household or person as in other similar studies (e.g., TTI study). That said, profiles of household and person, however, are still beneficial to this study effort. To study the factors impacting travel behavior in a community, one needs to understand not only the characteristics of that community but also its residences (each household as a whole, as well as, the persons in that household).

3.1.1 Demographic Characteristics of Households Taking Walk/Bike Trips

Household demographic profiles are examined using the NHTS add-on data and findings are summarized in this subsection. For example, Figure 1 shows the comparison between the number of households that took walk/bike trips and those that did not by the geographic region of household locations. The “All HHs” group represents a general “overall” distribution of all household locations, i.e., all households in the NHTS add-on areas. Thus, the largest “share” is the South Atlantic region due to the large populations in states such as Texas, Florida, and Georgia. The same grouping method is used in all other charts presented in this subsection and also used in many other figures in other parts of this Section.

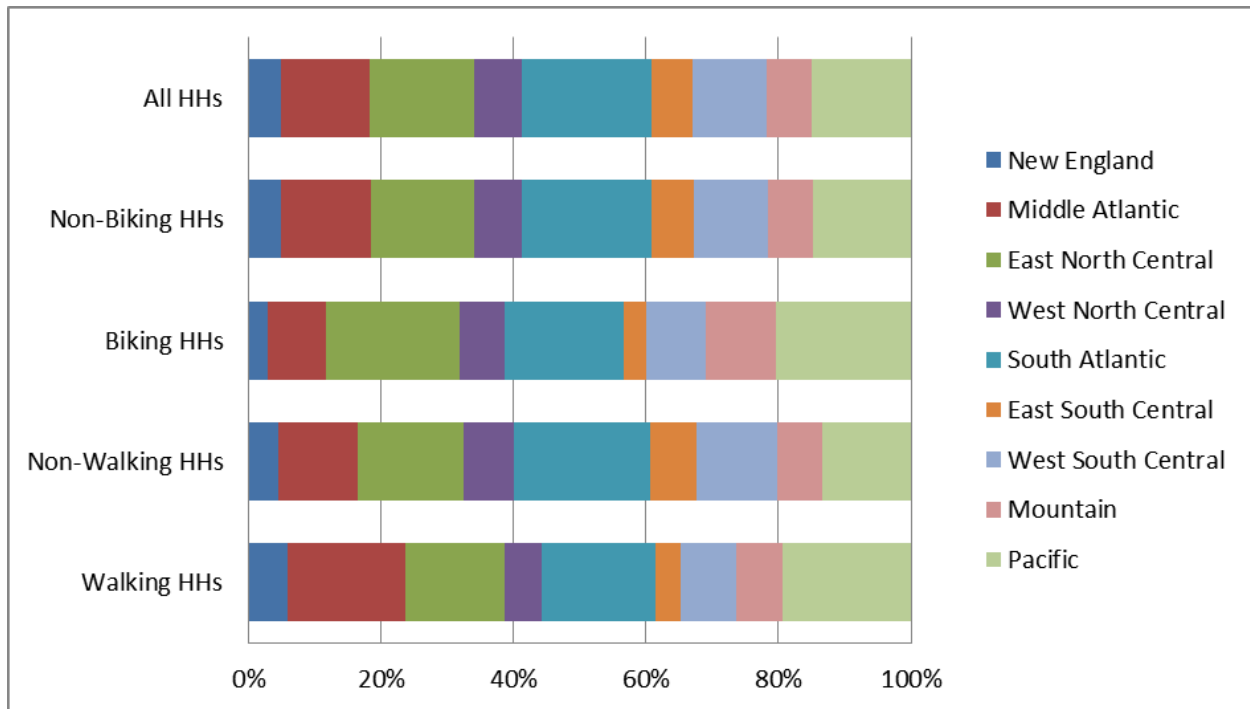


Figure 3. Distribution of number of households with/without walk/bike trips by geographic location of the household.

Because the number of households that took walk/bike trips is relatively small, in comparison to those that did not make any walk/bike trips during the survey period, distributions of the “Non-Walking HHs” and the “Non-Biking HHs” groups are generally similar to the distribution of the “All HHs” group. The “center of attention” when viewing the figure, thus, should be on comparing between the “walking HHs” and the “All HHs”; and between the “Biking HHs” and the “All HHs” groups.

This section also compares travel behaviors among households living in different geographic regions, specifically by the Census Regions. The map²⁵ in Figure 2 shows the definitions of nine Census Regions, which is used in 2009 NHTS.

²⁵ Map obtained from: http://www.eia.gov/emeu/reps/maps/us_census.html

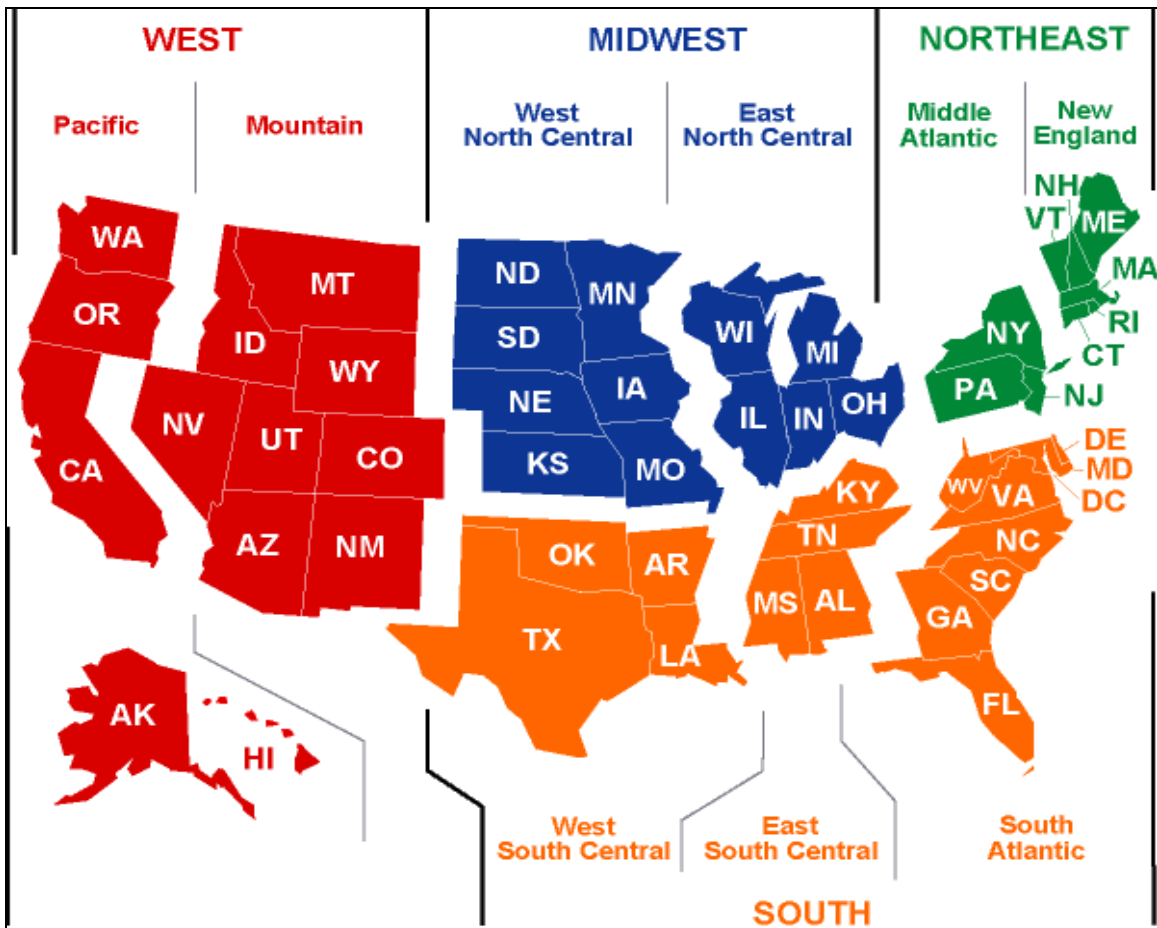


Figure 4. Nine Census geographic regions used in the 2009 NHTS.

Geography Region of the Household

Based on Figure 1, there is clearly a smaller number of walking households in the South Atlantic region, a share of 17% when considering walking households versus its share of 20% in the general population. On the other hand, a higher share of walking households can be seen in the Middle Atlantic region (18% share); and similarly in the Pacific region (over 19%). Figure 1 also shows that a much smaller share of households in the Middle Atlantic region have taken bike trips during the survey year; less than 9% of the “biking households” versus over 13% of households live in this region. East North Central and Pacific regions are the two that have the highest shares (over 20% each) in share of biking households.

Racial Difference

White households have a slightly lower share in making walk trips, but with a higher share in taking bike trips (Figure 3). Although they are as likely as every American household when considering walk trips, African American households are much less likely to take bike trips. All other non-white households, as a whole, also have a slightly higher share in taking walk trips; about 14% of walking households belong to this group while they are accounted for about 11% of the household population.

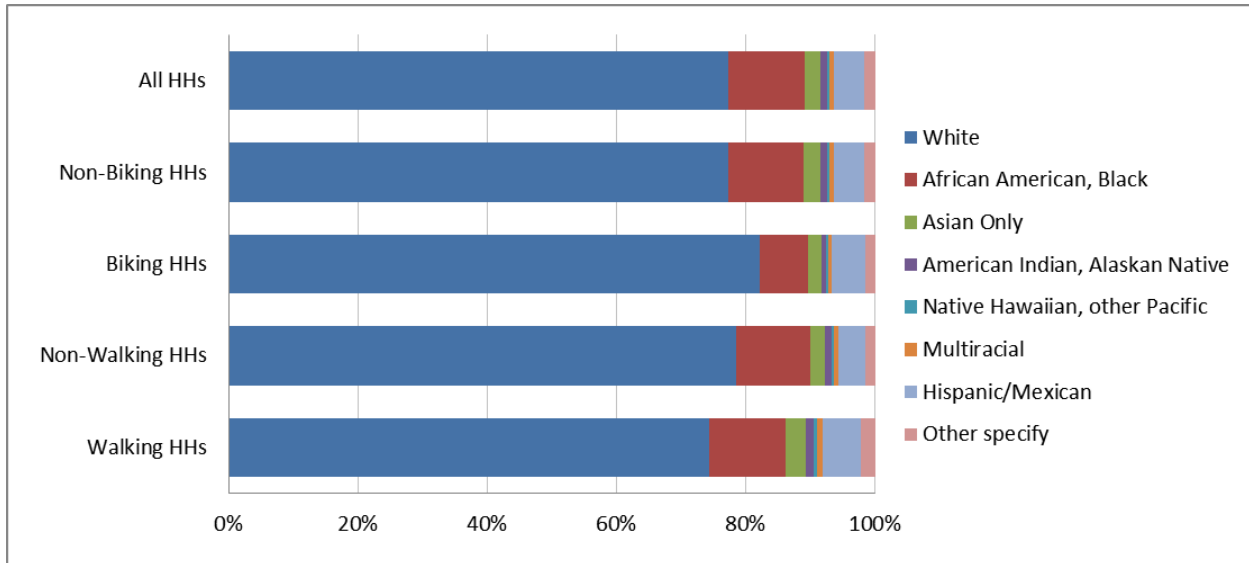


Figure 5. Distribution of number of households with/without walk/bike trips by race.

Household Size

Based on data shown in Figure 4, a single person household is less likely to take walking or biking trips than multiple person households. This is particularly apparent for bike trips, less than 12% of the single-person households took bike trips when considering that single-person households accounted for about 28% of total households in the country. Larger size households (with three or more persons) are much more likely to take walking and biking trips. This group accounted for over half of the total “walking households” and two-thirds of the “biking households,” although less than 40% of the households in the country are in this size (i.e., 3 or more persons).

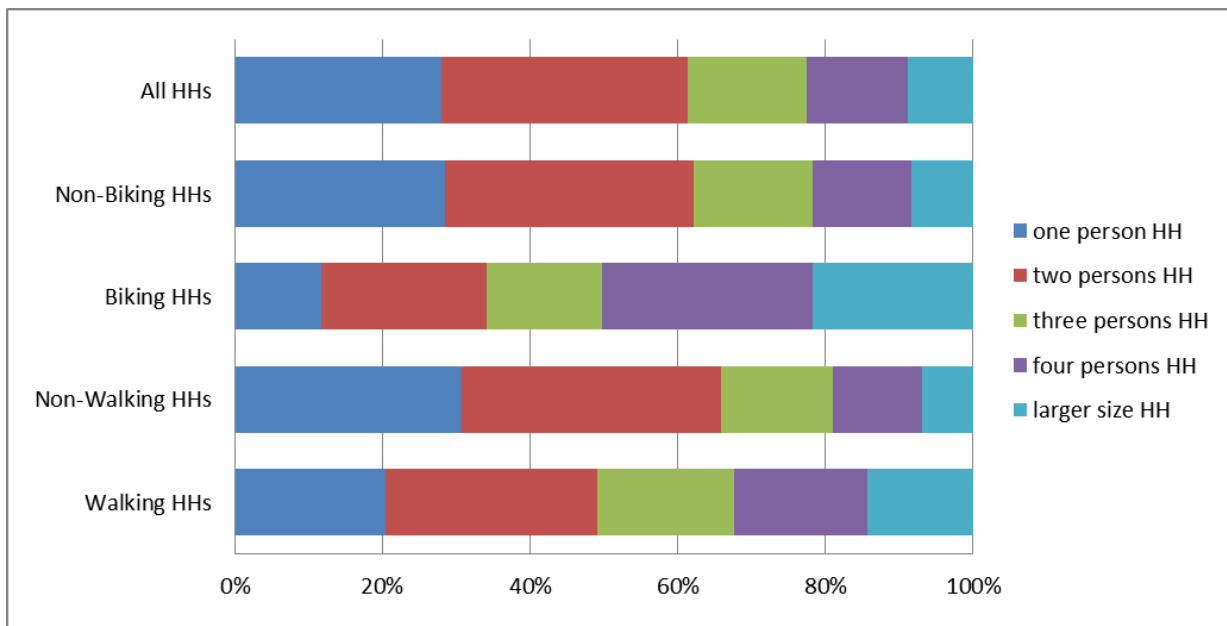


Figure 6. Distribution of number of households with/without walk/bike trips by household size.

Households With or Without Children

Not surprisingly, as shown in Figure 5, households without children are less likely to take walking or biking trips, and particularly biking trips. Households with children between ages 6 and 15 have the highest share of bike trips, and accounted for over 40% of the total households with bike trips; while this group accounted for approximately 15% of the population (i.e., all households). Over all, households with children of all ages are more likely to walk or bike than those without; about 46% of households who walked have children, while 64% of households who biked have children, significantly higher than their share in the population (~34% households have children between ages 0 to 21).

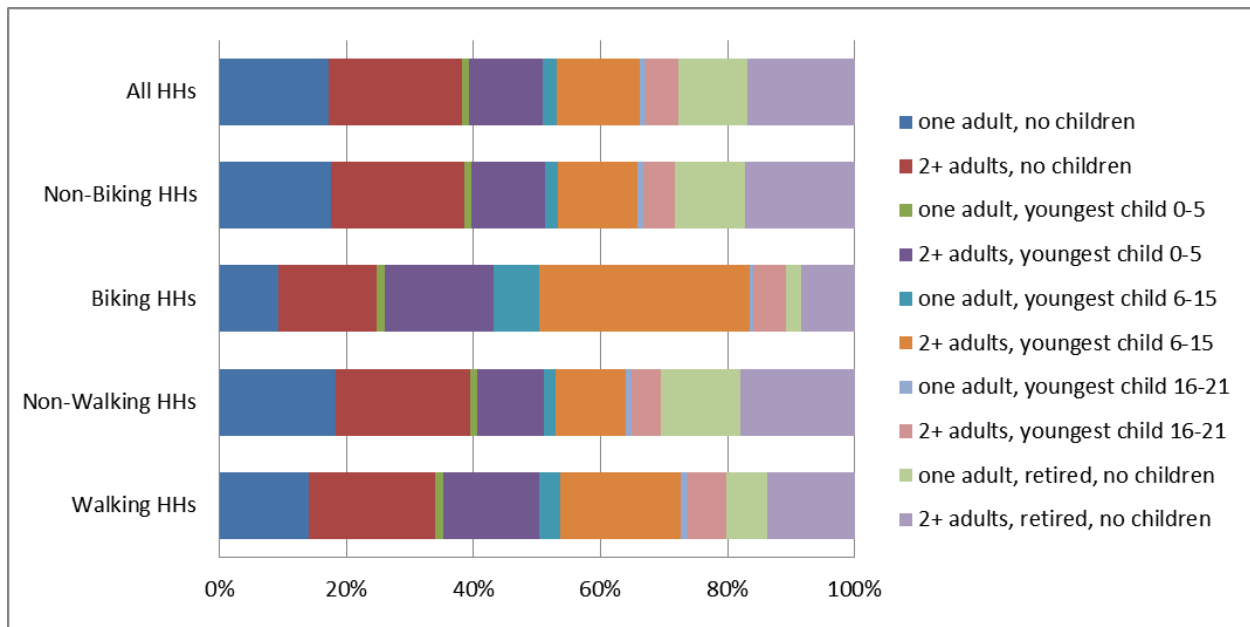


Figure 7. Distribution of number of households with/without walk/bike trips by type.

Urban Status

As one would expect, households in urban areas make up the vast majority of all households, in addition to a slightly higher majority of households taking walk and bike trips. Households not in urban areas, conversely, are less likely to be walking or biking households (Figure 6). Particularly for walking, almost 84% of households that walked live in urban areas (either in an urban cluster or in an urban area), while only 77% of total households are living in urban areas.

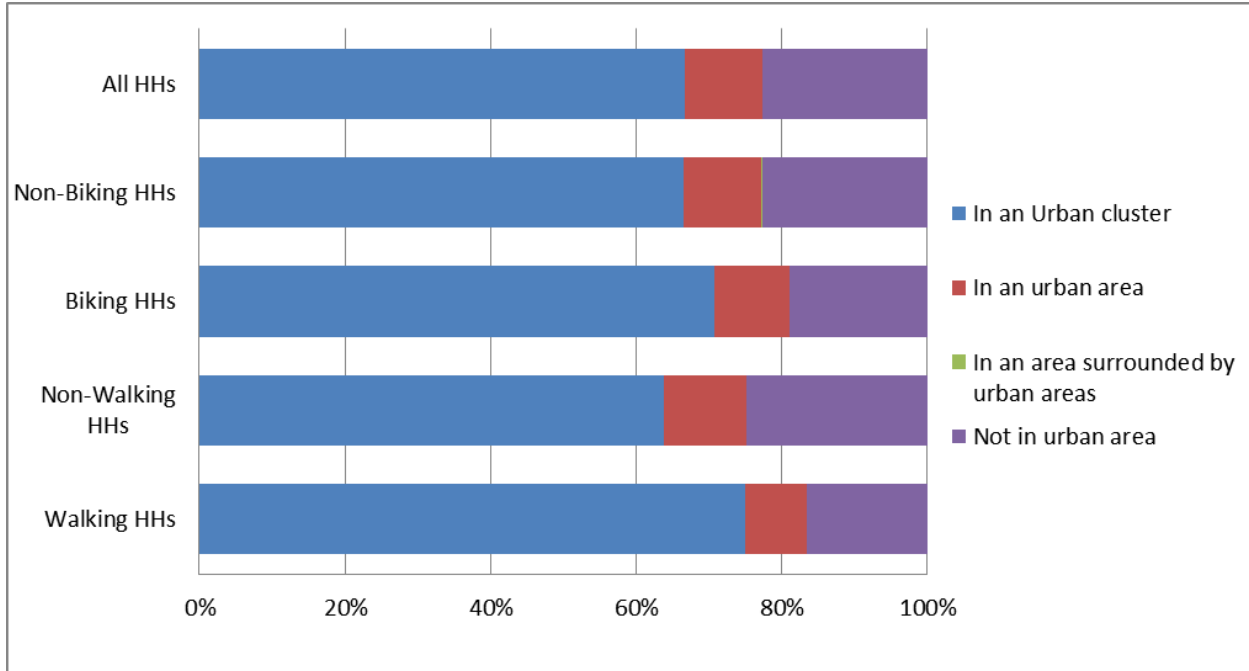


Figure 8. Distribution of number of HHs with/without walk/bike trips by urban status.

Workers in Households

The number of workers in a household has an impact on the likelihood of walking and biking. In particular, households without any workers are far less likely to walk or bike than households with at least one worker (see Figure 7). Also, close to half of all biking households have two or more workers, while the same group of households only accounts for 30% of the overall population (i.e., total households).

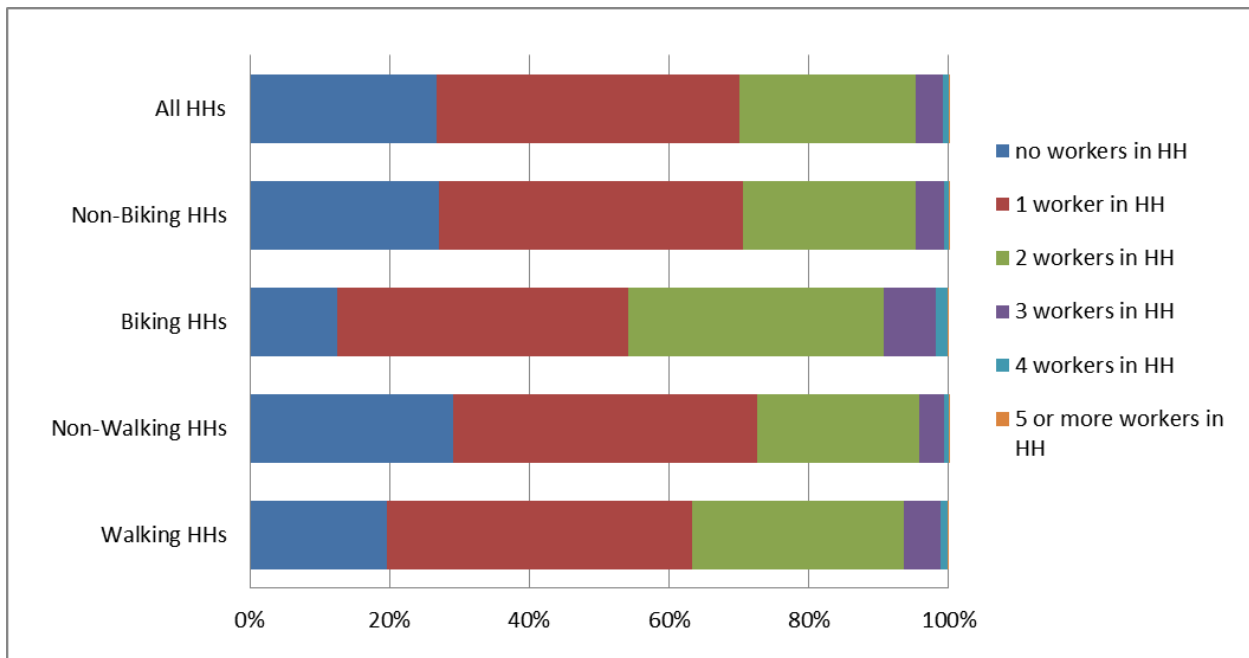


Figure 9. Number of HHs with/without walk/bike trips by workers in the household.

Rail Availability

While there is a limited impact on biking in terms of rail availability, MSAs with rail have a much larger percentage of households that take walk trips than MSAs without rail (Figure 8). The rail availability certainly provides additional mode choices to residents in surrounding areas. People living close would have the ability to walk to/from rail stations; while others might have to utilize park and ride lots near the rail stations.

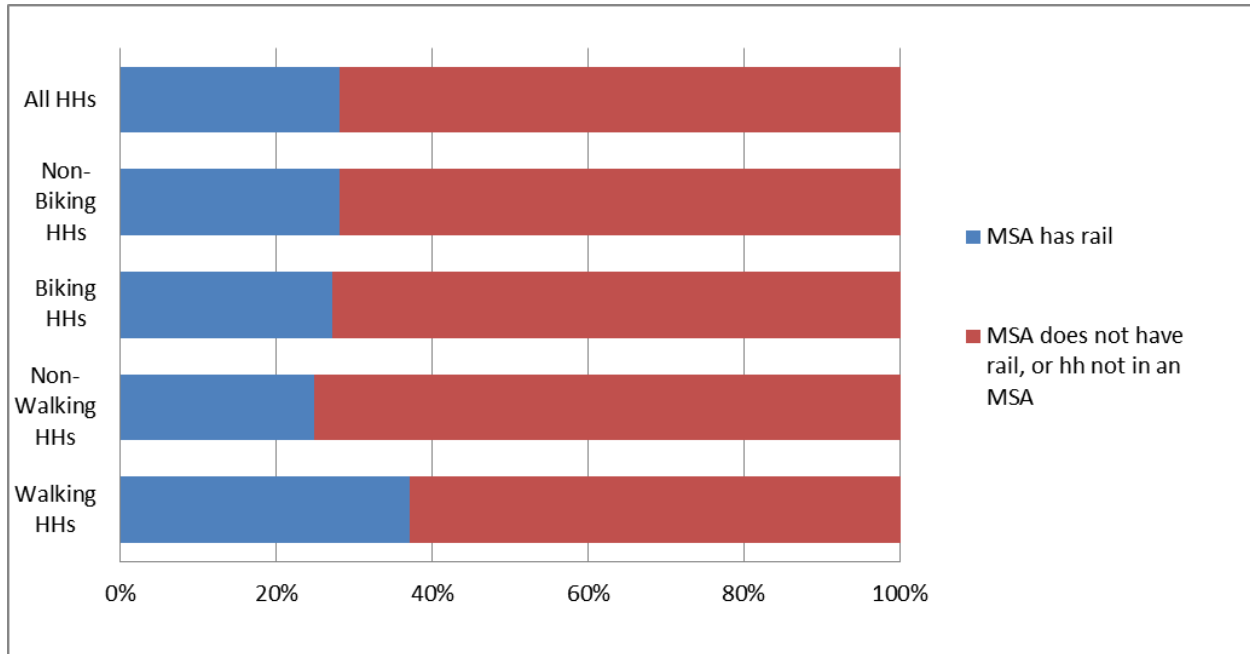


Figure 10. Distribution of number of HHs with/without walk/bike trips by rail availability.

3.1.2 Demographics of Persons Taking Walk/Bike Trips

In addition to examining the walk/bike profiles at the household level, a similar review on persons who walked/biked was also performed using the add-on data. Person demographic characteristics investigated under this effort include: birth status (i.e., born in the U.S. or not), driver status (driver vs. non-driver), education level, use of public transit, gender, age, and worker status. In most cases, the difference between persons who walk/bike and those who do not does not appear to be as significant as those comparisons made with households.

Birth Status and Driving Status of the Traveler

A foreign born traveler has a higher likelihood of walk/bike than their US-born counterparts. Non-drivers are also more likely to walk/bike. These characteristics (Figures 9 and 10) in travelers are likely to be correlated to one another and also can be expected to associate with the household racial difference seen in Figure 1.

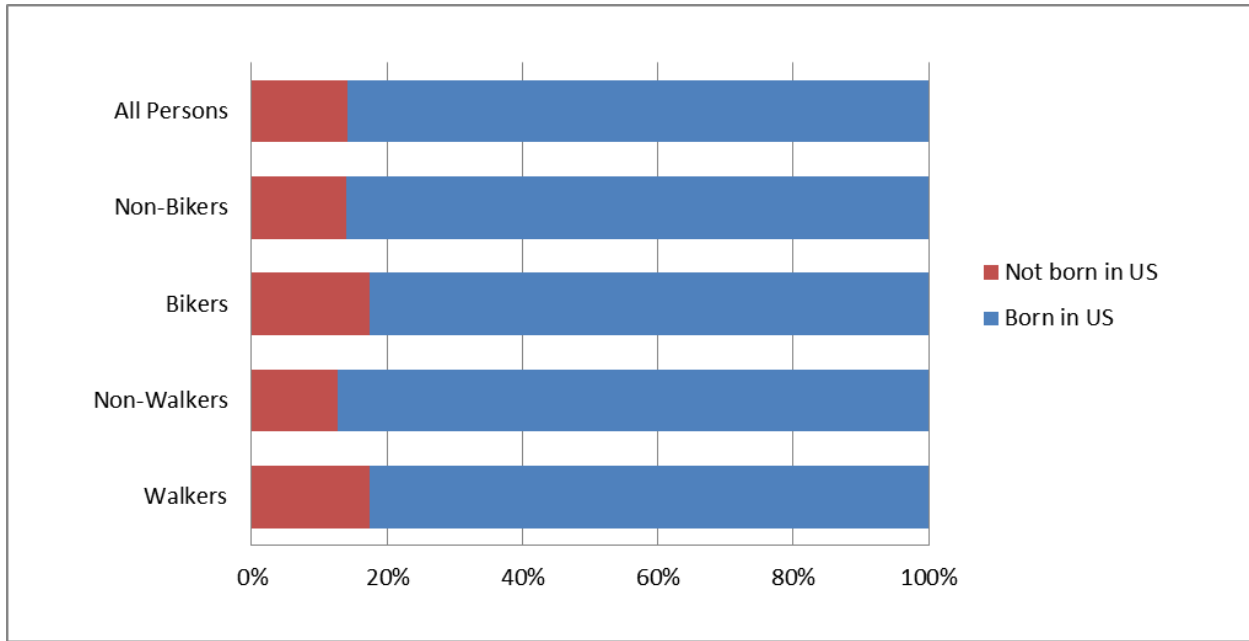


Figure 11. Percent of persons with/without walk/bike trips by birth status.

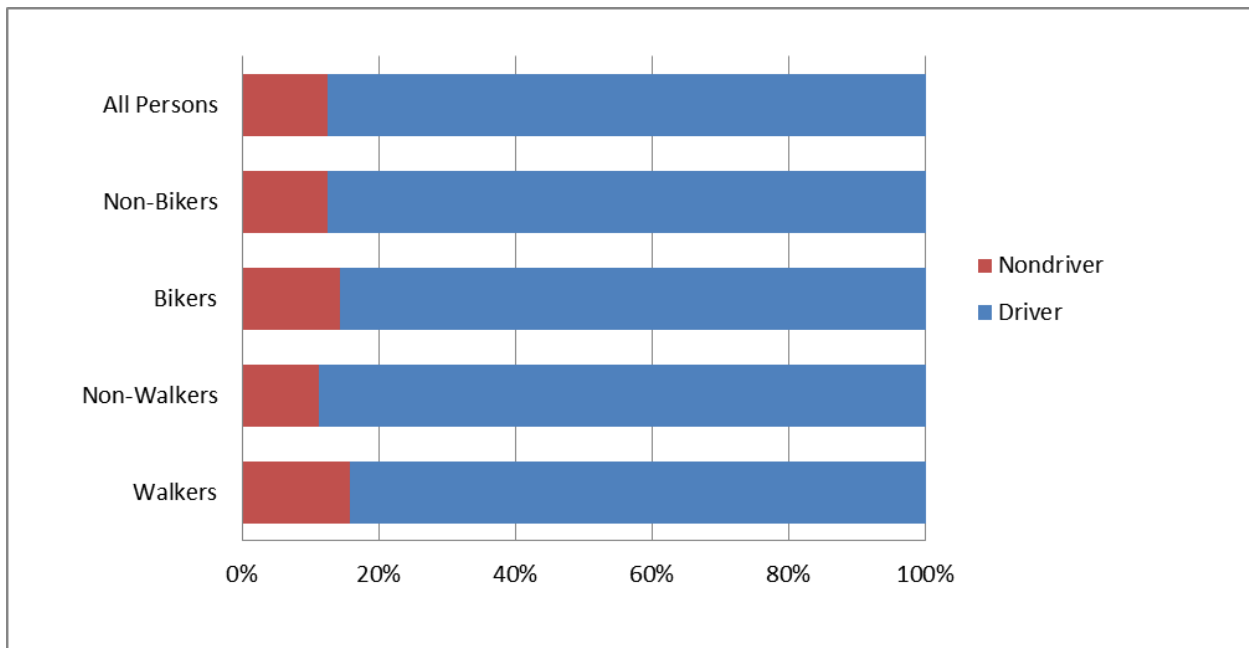


Figure 12. Percent of persons with/without walk/bike trips by driver status.

Education Level

Generally speaking, as shown in Figure 11, travelers with a higher level of education are more likely to walk/bike than those with high school or less educations. Those with the highest level of education (the rightmost group shown in cyan color) clearly have a significant level of bike and walk activities. This group accounted for about 13% in the population but has a share of over 16% for walking and 19% for biking.

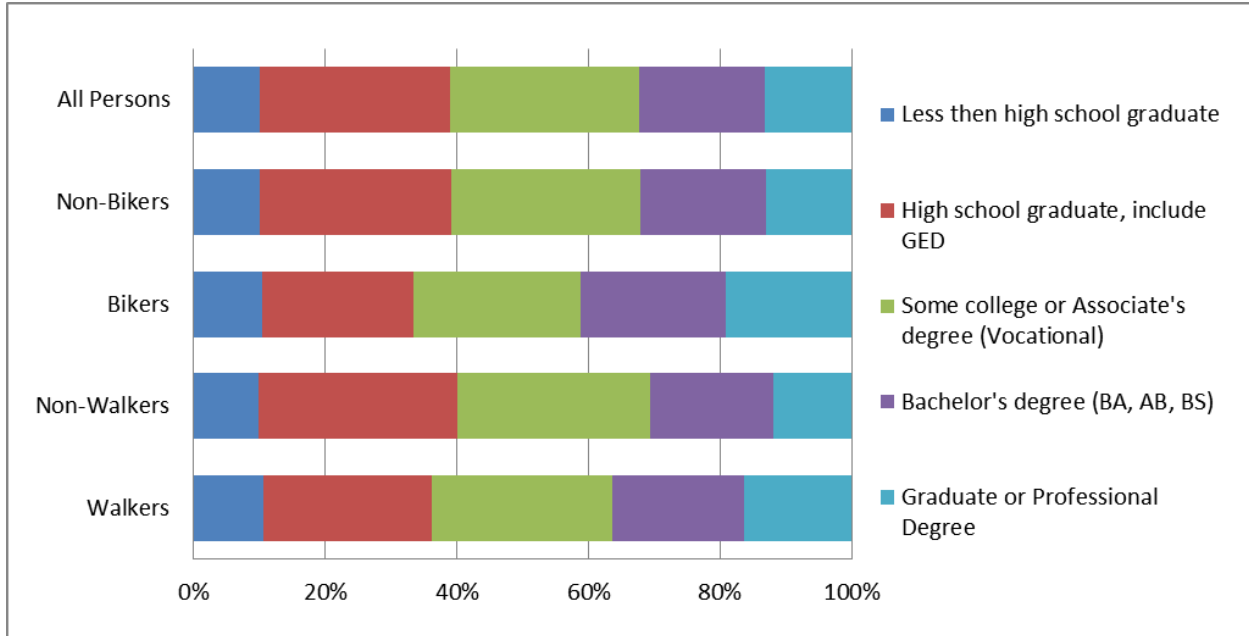


Figure 13. Percent of persons with/without walk/bike trips by education level.

Use of Public Transit

Similar to what was observed in Figure 8 (rail access by household), travelers who use public transit have a higher likelihood of walking/biking. About 35% of walkers used public transit at least one time during the survey year, while 24% of the general population and less than 20% of non-walkers did (Figure 12). Transit users also have a higher percent of biking, although not as high as walking. A correlation between transit and walk/bike clearly exists.

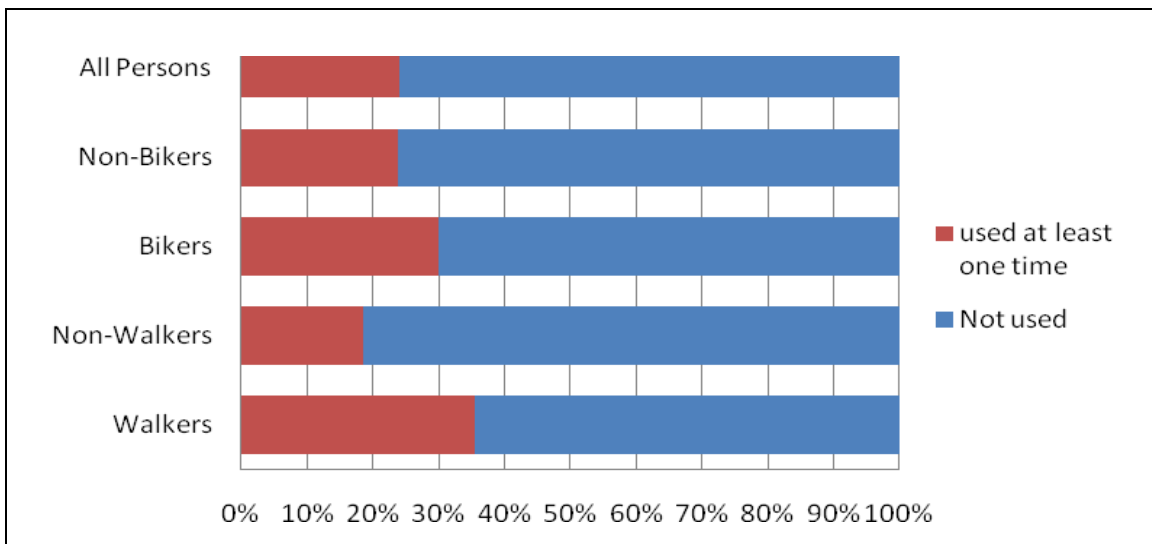


Figure 14. Percent of persons with/without walk/bike trips by use of public transit.

Gender and Worker Status

Both gender and worker status has a bigger impact on biking than walking (Figures 13 and 14). It is commonly known that safety and security are major concerns for females regarding biking. Thus, it is no surprise that males have a higher share of biking. No significant gender impact can be seen for walking, however. It is also interesting to see that workers are more likely to bike than non-workers, but show no significant difference in walking.

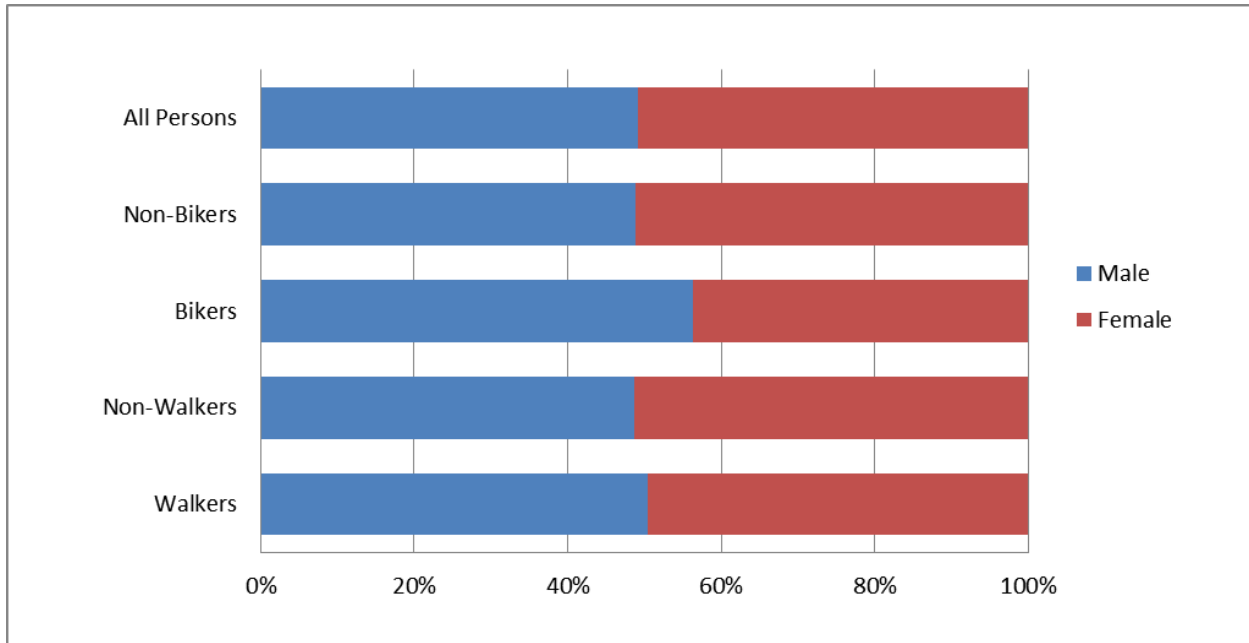


Figure 15. Percent of persons with/without walk/bike trips by gender.

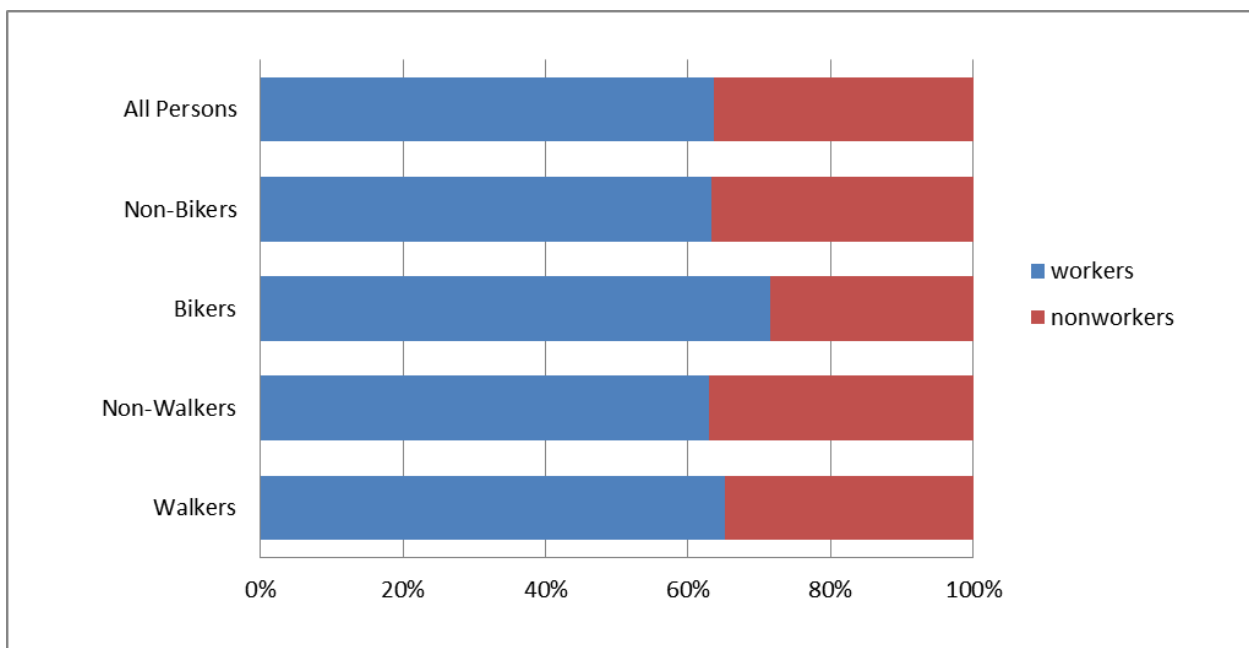


Figure 16. Percent of persons with/without walk/bike trips by worker status.

Age Group of the Traveler

Figure 15 shows the distribution of walk/bike trips by age group. The distribution for all trips is also included in the same chart as a reference. Clearly, bike trips are mostly (about 40%) taken by young children, ages 5 to 15 specifically (note that NHTS does not survey children younger than age 5). Walking for this young group also has a relatively higher share (i.e., red bar lines are taller than the green bar lines), although nowhere near the scale as in the biking. This reconfirms the earlier observation that households with children are more likely to bike (Figure 5).

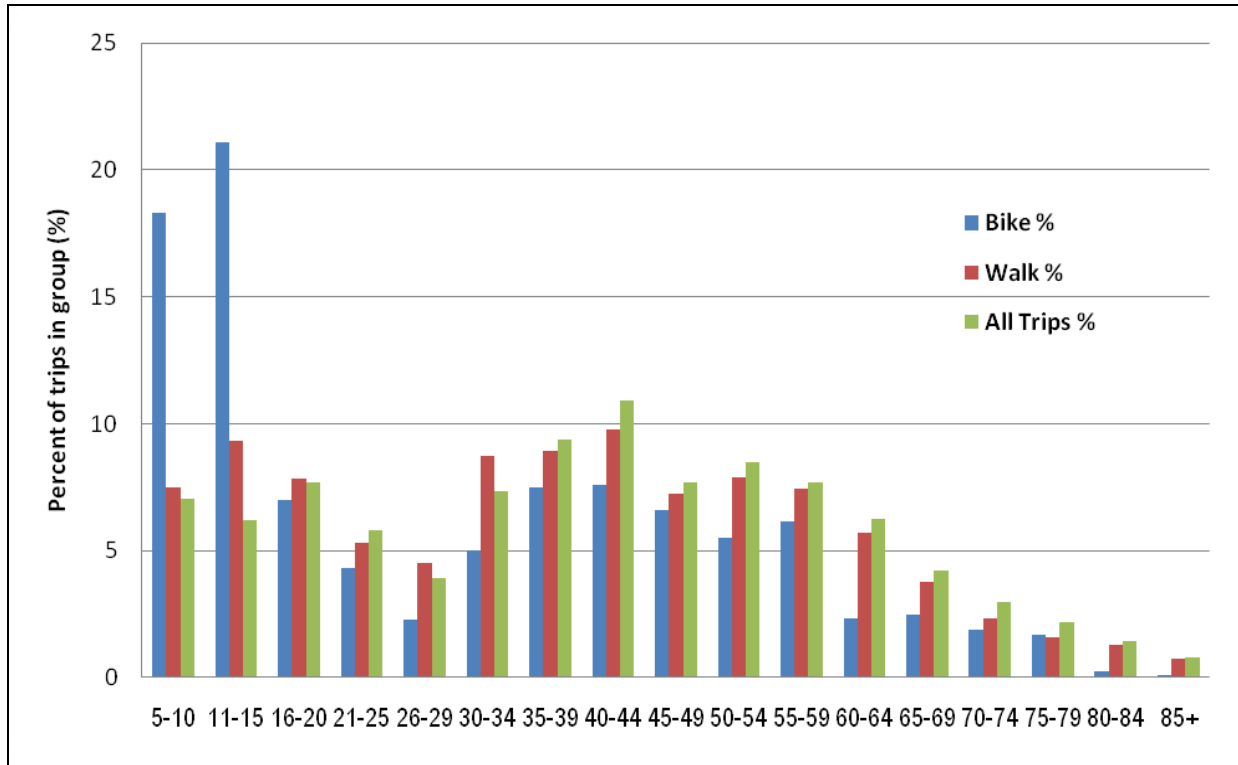


Figure 17. Distribution of walk/bike trips by age group.

As a whole, walk trips accounted for about 10% of the total trips, while bike trips only accounted for about 1%. When examining the percent of walking/biking trips made by each age group (Figure 16), 16% of the trips taken by the 11-15 age-group are walk trips, which has the highest percent of walking among all age groups. As observed in Figure 16, the age group that has the highest percent of biking trips is children age 11-15, followed by the younger age group of 5-10.

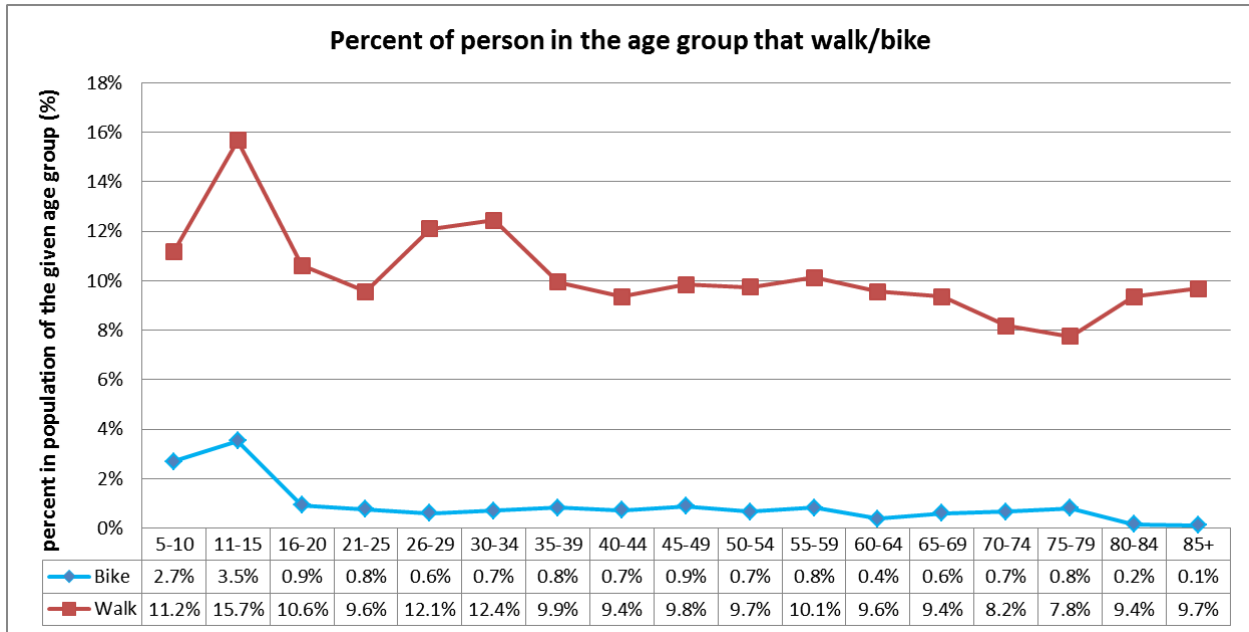


Figure 18. Percent of trips taken by the age group that is a walk/bike trip.

3.1.3 Land-use proxy with add-on data

Using the geographic destination information from add-on areas, a possible proxy variable for land use was computed using information contained under the WHYTO variable. This WHYTO variable provides the reason/purpose a trip was made to its given destination. The percentage of trips for each block group was computed for five aggregate WHYTO purposes: home, work, community (e.g., school/church), retail, and other. While other potential sources might provide much more detailed data for the purposes of a land use factor, this NHTS-derived measure can at least provide some information in the absence of other such data.

Figure 17 below shows an example of using this proxy to examine the land use distributions in a downtown neighborhood in Dallas, TX. A pie chart is shown in each block group displaying the distributions of the various land uses in that block group. One can see from this map, the primarily residential areas are those shown with pie charts predominantly dark green, while more retail areas are primarily orange. Multiple color pies are those with multiple land uses. One should keep in mind that NHTS was not designed for analysis in great detail of geographic levels, even with Add-on areas. Thus, sample size at Census block group and tract levels could be spotty. Consequently, it is possible that a certain characteristic of walk/bike trips in a given block group could in fact be driven from one or two samples.

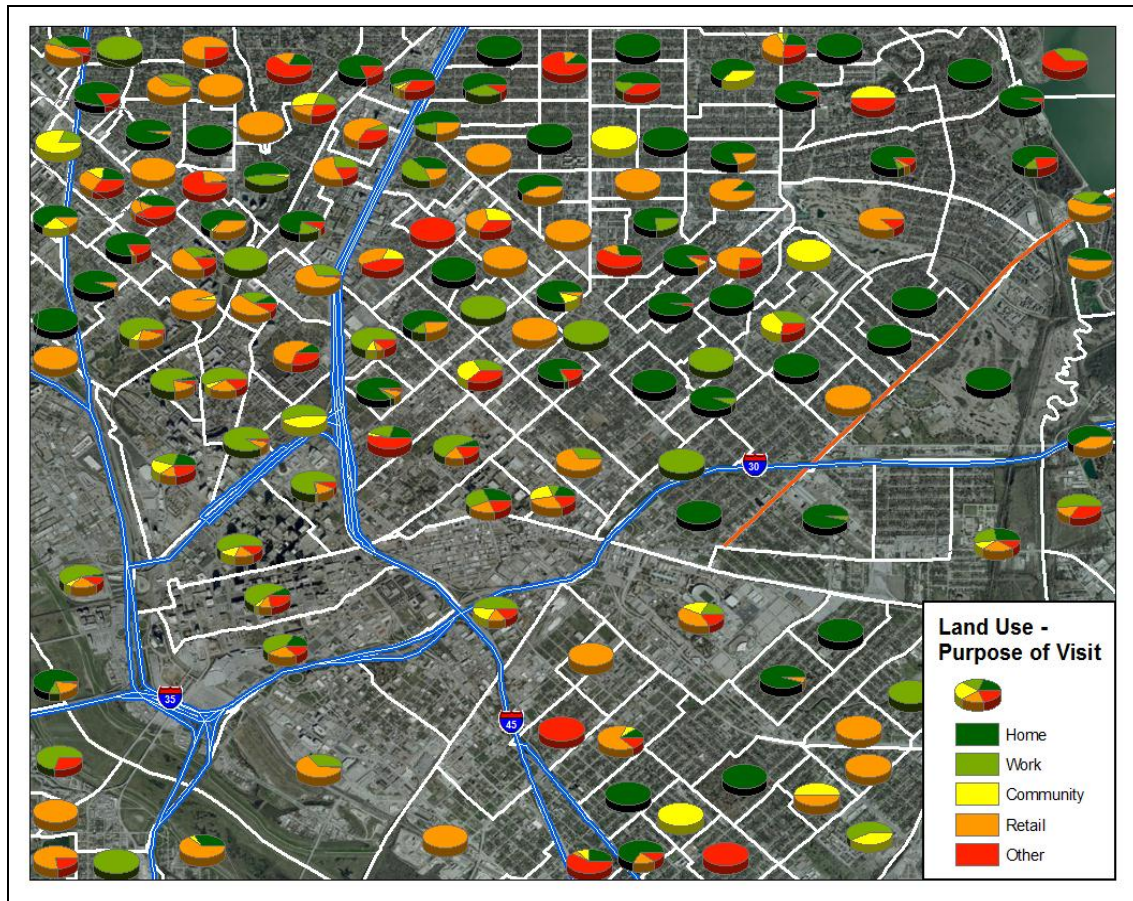


Figure 19. Use of NHTS "WHYTO" variable as a land-use mix proxy, example of a downtown neighborhood in Dallas, TX.

3.2 VARIABLES FROM NIELSEN CLARITAS

As part of the NHTS, FHWA purchased demographic data, by Census tract and block group, from Nielsen Claritas. Among the 30 variables at the block group level and 36 variables at the Census tract level were median income and income distributions, percent of households in poverty, housing density and median year housing units were built, own/rent information, race, population and employment density, urban classification, education, and percent of persons foreign born. These variables provide a base (e.g., denominators) for calculating additional measures from the NHTS data. However, the employment density data purchased from Nielsen Claritas for the 2009 NHTS was residential based data (i.e., where the workers live); therefore does not accurately reflect the density at places of employment. This is a major drawback with the NHTS data.

To accurately measure employment density, the ORNL NMT team entered into a license agreement with Nielsen Claritas and purchased an employment database. Due to resource limitations, data purchased for this project only covers Census block groups within the nine ORNL selected regions (see Section 4.3). Two variables are included under this agreement: the total employment count and its associated total retail employment count, for each of the block

groups within the selected regions. Using this new data, employment density in each block group can be correctly calculated.

To illustrate the difference between the two employment variables, employment densities using NHTS employment density variable and the density calculated based on ORNL purchased database for a community around Washington DC area are shown in Figures 18 and 19, respectively. The scales used in these two maps are the same which allows visual comparisons to be made. Differences in employment density are clearly visible in these maps for many block groups in the region.



Figure 20. Employment densities based on Claritas variable included in the 2009 NHTS.

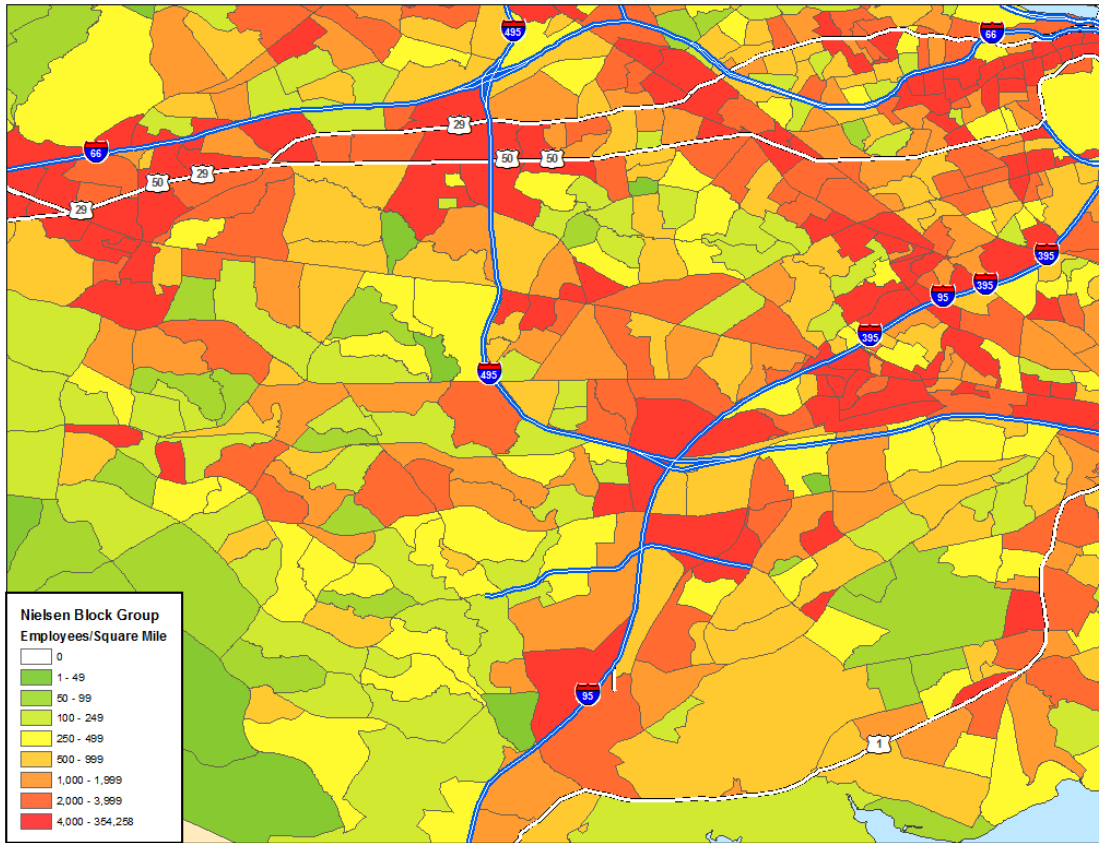


Figure 21. Employment densities based on ORNL-purchased Claritas data.

3.3 AMERICAN COMMUNITY SURVEY (ACS) DATA

The ACS²⁶ is a survey conducted by the U.S. Census Bureau of about 3 million households each year (250 thousand per month), subject to the constraint that households should not be surveyed more than one time in any five year period. Thus the ACS is very intensive - about 20 times as big as the NHTS (3 million versus 150 thousand) - and it is repeated every year. As the ACS does not have add-on components, it is also geographically more uniform than the NHTS. Each year's ACS sample includes, on average, almost 50 households per Census Tract and almost 15 households per Block Group.

Data on demographic, social, and economic characteristics is collected in the ACS. The ACS also collects data on commuting, i.e., the Journey to Work (JTW), including mode of transportation and travel time to work. The 5-year ACS for 2005-2009 is currently available and was used for this project. There are separate modes for bicycling and walking, allowing the data to be used for NMT studies.

²⁶ <http://www.census.gov/acs/www/>.

The ACS results are also used to produce the Census Transportation Planning Products (CTPP).²⁷ Prior to the ACS, the CTPP used the Census “long form” results. The next CTPP with tabulation at the tract or smaller geography (TAZ) will not be available until mid-2013, and was not available for this project.

As in many Census released data series, due to disclosure issues and the requirement to fulfill meet Census publication standards, ACS data is released only in various levels of aggregated tabulation, rather than a microdata format such as the NHTS data²⁸. For example, instead of listing detailed characteristics of each household in a separate record as in the NHTS household data file, the records in the ACS data file provide only aggregated summary for a group of households within a specific category. Specifically, current ACS data at the Census Tract and Block-Group levels is only available for five-year periods. The latest five-year release of 2005-2009 is the best ACS data available for the ORNL NMT project.

Table 2 summarizes ACS data involving JTW estimates. Over 99% of all Census Tracts (CT) and Block Groups (BG) have JTW data, with over 80% of tracts having at least one worker reporting their mode to work as “Walk,” and over 30% of tracts having at least one worker reporting “Bicycle.” The number of geographic regions with workers reporting that they used a bicycle to get to work is quite low, with only around 1% of Census Tracts having 5% or more of their workers reporting biking to work, and only 260 tracts having 10% of their workers report using bicycles. Of these 260 tracts, over 100 are from either California or Oregon.

Table 2. Geographic Coverage of the 5-Year 2005-2009 ACS Data

	Number of Census Tracts	Number of Block Groups
Total number of areas (CT/BG)	65,461	208,797
CT/BG With ACS JTW Data	64,942	207,252
JTW Walk Trips		
> 0 workers	53,323	101,915
> 5% of workers in CT/BG	6,196	19,139
> 10% of workers in CT/BG	4,086	16,661
JTW Bicycle Trips		
> 0 workers	20,913	29,032
> 5% of workers in CT/BG	718	3,869
> 10% of workers in CT/BG	260	1,933
> 10% of workers in CT/BG (California)	65	399
> 10% of workers in CT/BG (Oregon)	41	165

²⁷ CTPP is an American Association of State Highway and Transportation Officials (AASHTO) sponsored data program and is funded by member state transportation agencies. The program operates with support from the U.S. Department of Transportation (including FHWA, RITA, and FTA), Census Bureau, MPOs, and the TRB.

²⁸ Note that ACS also releases Public Use Microsample (PUMS) Data that is similar to NHTS; however, it is limited in geography, and does not allow for estimation at the Census tract or block group levels.

Using the 2005-2009 ACS JTW data, Figure 20 shows the numbers of workers that walk to work at the Census block group level. An insert of a zoomed view of the greater Washington DC area is also shown as an example in Figure 20 so that more details are visible. Similarly, Figure 21 displays the numbers of workers that bike to work; while the zoom view shows a selected smaller region around Milwaukee, Wisconsin. Instead of block group, the geography of bike to work data is shown at the Census Tract level.

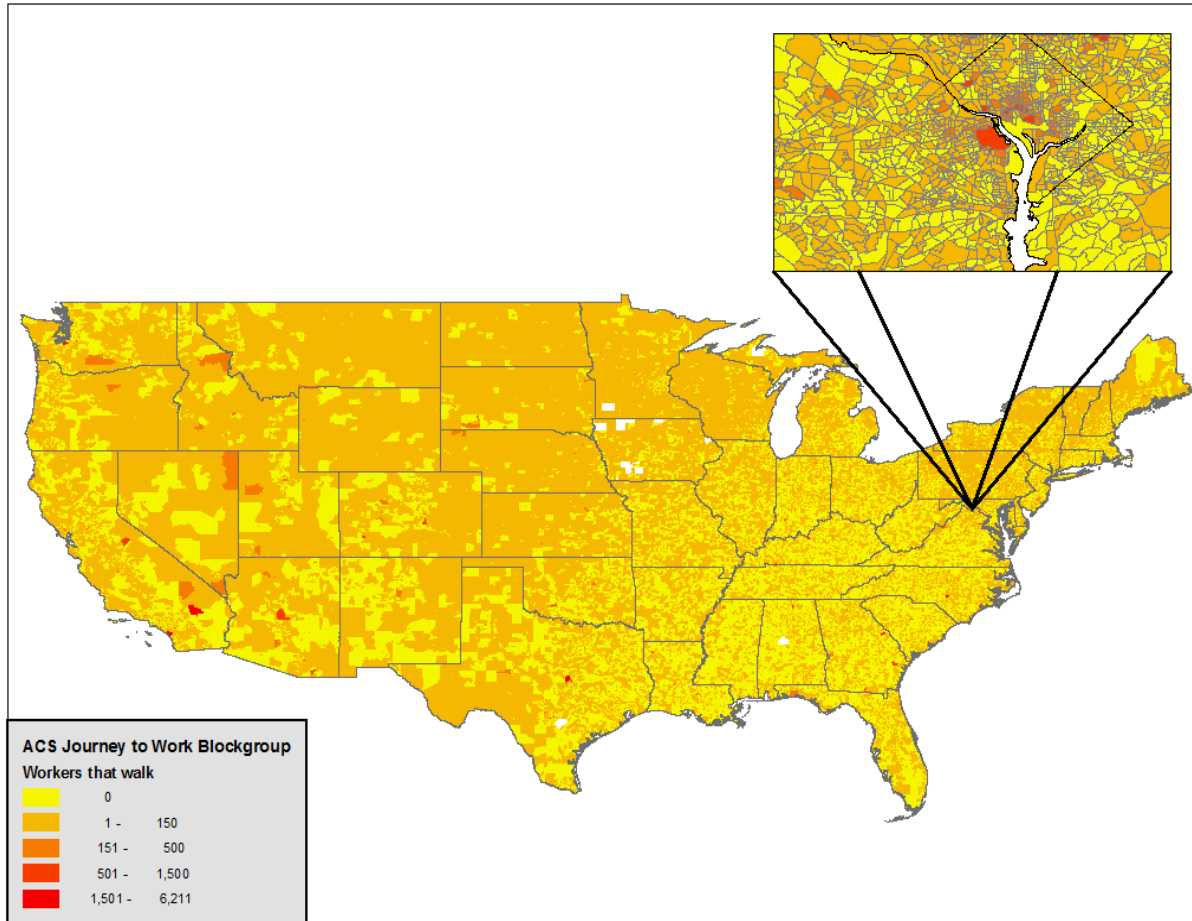


Figure 22. Number of commuters that walk to work by block group, with insert showing more details in the greater Washington, DC area. (Source: ACS Journey to Work data, 2005-2009.)

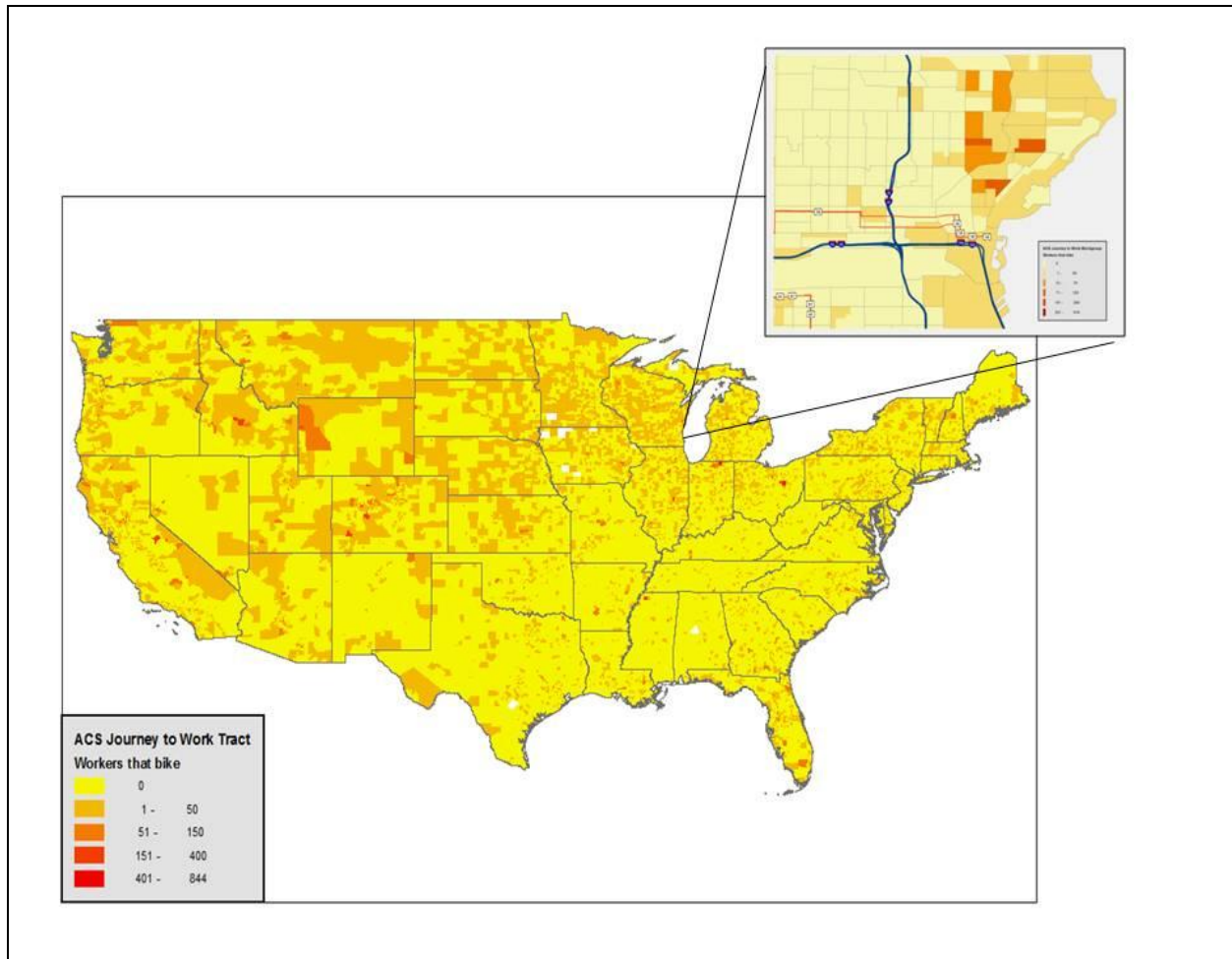


Figure 23. Number of commuters that bike to work in Census Tracts with an insert view of the greater Milwaukee, Wisconsin area. (Source: ACS Journey to Work data, 2005-2009.)

3.4 WALK SCORE DATA

Walk Score (www.walkscore.com) is a web site that lets users calculate a walkability score for any specific address or region (e.g., cities or neighborhoods). This score is defined as a numeric measure between 0 and 100 and can be used to indicate the walk-friendliness of any given location. The scoring algorithm²⁹ awards points based on distance to a diverse set of nearby amenities, e.g., grocery stores, restaurants, parks, schools, etc., and certain categories are weighted more heavily than others to reflect destinations associated with more walking trips. For instance, amenities within .25 miles receive maximum points and no points are awarded for amenities further than one mile. Road connectivity such as intersection density and average block length are also factored into the scoring algorithm. The Walk Score definition of degree of walkability is measured and classified in five general categories shown in Table 3.

²⁹ The Walk Score algorithm can be found at <http://blog.walkscore.com/research/methodology/>.

Table 3. Walk Score Categories

Walk Score	Description
90 - 100	Walker’s Paradise – Daily errands do not require car
70 – 89	Very Walkable – Most errands can be accomplished on foot
50 – 69	Somewhat Walkable – Some amenities within walking distance
25 - 49	Car-Dependent – A few amenities within walking distance
0 - 24	Very Car-Dependent – almost all errands require the use of a car

Walk Score published a city walkability ranking for the largest 2,500 cities in the U.S. (<http://www.walkscore.com/rankings/cities>). Walk scores for these cities, as classified in the 5 categories defined in Table 3, are shown in Figure 22. Based on Walk Score’s walkability measures, Cambridge, Massachusetts has the highest ranking with a score of 88.8; followed by New York City, New York with an 85.3 for its walk score.

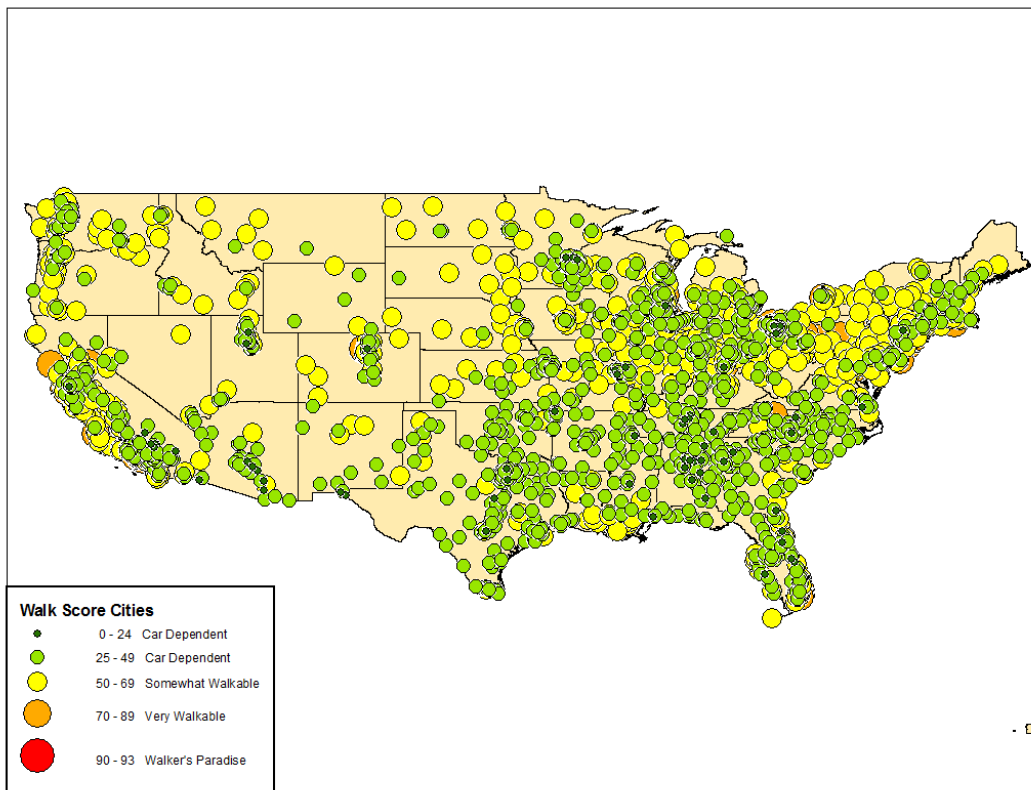


Figure 24. Walk scores for the top 2,500 cities in the U.S.

In recent years, Walk Score has added the Transit Score application to its web site. The transit score measures how well a location or community is served by public transportation. Based on information provided from Walk Score³⁰, the Transit Score algorithm calculates the score for a specific point by summing the relative “usefulness” of nearby transit routes. Walk Score defines

³⁰ See <http://www.walkscore.com/transit-score-methodology.shtml> for description of transit score methodology.

the “usefulness” of a location based on its distance to the nearest transit stop on the route, the service level, and type of route (bus, heavy/light rail, and ferry/cable car/other).

To calculate the score, values from all nearby transit routes are added. The service level of a route is measured as frequency per week. For type of route, bus is used as the base (i.e., a bus route has the score of 1), heavy/light rail is weighted 2 times and ferry/cable car/other are weighted 1.5 times. The same distance decay function as used in calculating the Walk Score is also used for Transit Score. The raw transit scores are then normalized to generate a value from 0 to 100. This score works in any city where transit agencies publish data in the Google Transit Feed Specification (GTFS) format. As in Walk Score, Transit Score is generalized in 5 categories (see Table 4) by the Walk Score.

Table 4. Transit Score Categories

Transit Score	Description
90 - 100	Rider’s Paradise – World-class public transportation
70 – 89	Excellent Transit – Transit is convenient for most trips
50 – 69	Good Transit – Many nearby public transportation options
25 - 49	Some Transit – A few nearby public transportation options
0 - 24	Minimal Transit – It is possible to get on a bus

3.4.1 Walk Scores for the Study Area

According to the Walk Score methodology report, the average walk score of a city could be calculated based on over 8 million walk score points. These individual points are weighted by population density and then aggregated into one average walk score for that city. Upon reviews of the methodology, and the geographic-specific walkability measures it produced, the ORNL NMT project team decided to include this Walk Score measure as one of the factors being considered under the study. Walk Score developers offer their fee-based services to others that are interested in specific regions or point locations. Due to resource limitations, walk scores were obtained only for block groups in selected NHTS Add-on regions (see Section 4.3 on regions identified for this study). The project team compiled a database with almost 20,000 geographic locations of block group centroids (i.e., lat-long of the center point of the block group polygon), and their respective city names and states. This database was then used by the Walk Score to generate walkability scores. Based on the resulting walk scores, Figure 23 shows an example of walk score patterns for a selected region in downtown Dallas, TX. In this display, the darker the green, the higher the Walk Score. Because these walk scores were calculated based on block group centroid locations, each block group is assigned with only one walk score regardless of the block group size. These walk scores might be more representative in very small block groups (e.g., typically in the downtown region) than those in the larger ones, since the walk score algorithm considers only amenities within up to a one-mile radius of the given location/point. Although a finer geography would be preferred, as mentioned before, resource limitations prohibited the project from obtaining walk scores at a more detailed geography.

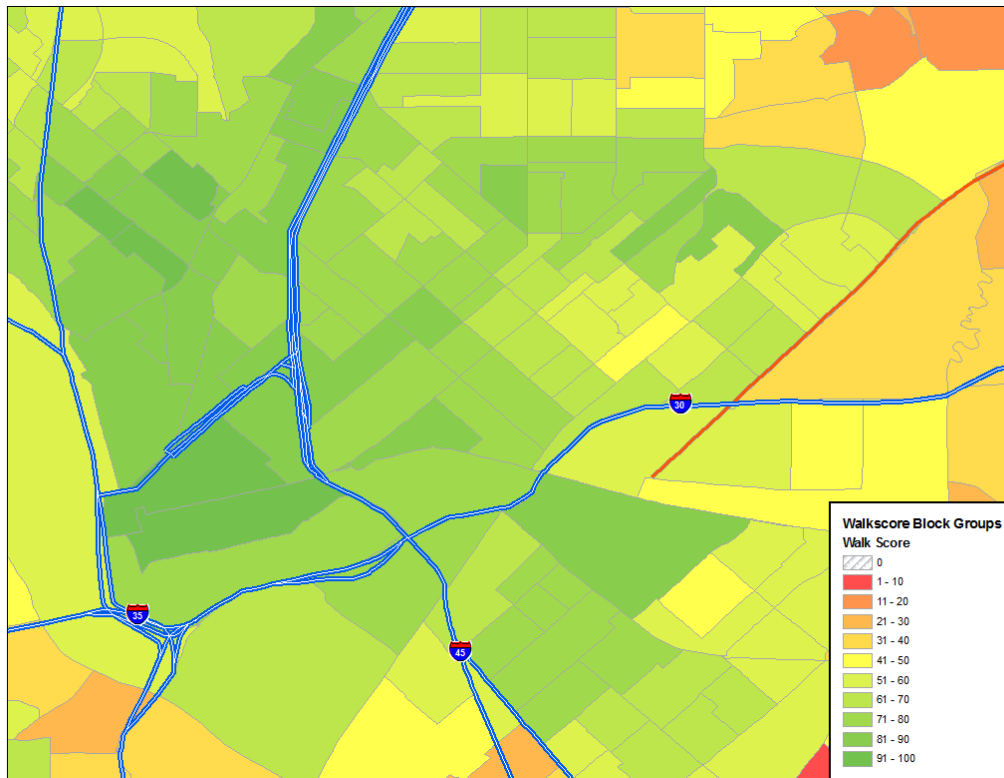


Figure 25. Work Scores in the Dallas, TX neighborhood.

3.4.2 Transit Scores in the Study Area

The same database prepared by the project team was also used by Walk Score to generate transit scores. Of the nearly 20,000 block groups, only about one-third of the block groups returned with a value. The rest of the block group records were returned with an “unavailable” as its result. The ORNL team performed additional checks against the full list of cities from the General Transit Feed Specification (GTFS)³¹ and identified all but about 3,500 block group records to be within cities that have no transit services. These were reassigned to have a transit score of zero. The remaining block groups, about 3,500, were treated as missing value for the subsequent modeling efforts. Figure 24 shows transit score patterns in the same downtown neighborhood of Dallas, TX, so that visual comparisons with the patterns from walk scores can be made. It is clear that Census block groups with higher transit scores (darker green areas shown in Figure 24) generally have a higher walk score (see Figure 23).

³¹ See <https://developers.google.com/transit/gtfs/>.

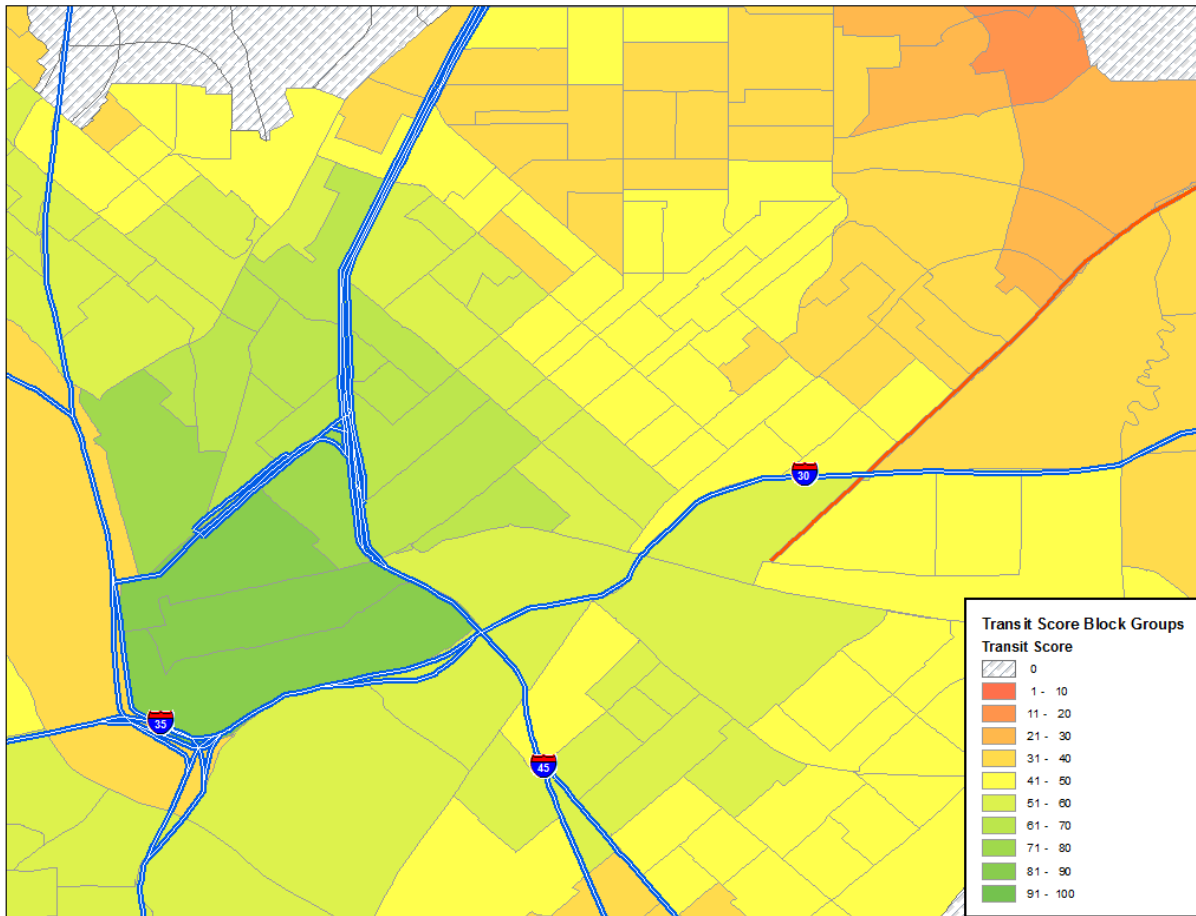


Figure 26. Transit score patterns at the downtown neighborhood of Dallas, TX.

3.5 GEOSPATIAL DATA

Walkable and bikable places are streets and districts with physical attributes that encourage walking/biking for functional and recreational purposes. Because of this, knowledge on land use characteristics of a neighborhood's physical environment is also an important component needed for studying factors that impact walking/biking activities. The geospatial nature of a neighborhood/community is best addressed by the use of Geographic Information System (GIS) software. Factors such as roadway density, intersection density, and population density are commonly used as attributes in studying neighborhood land-use mix.

At the beginning of this project, a set of roadway density and intersection density data developed by the University of Illinois at Chicago (UIC), for a FHWA-funded 2001 NHTS transferability study, was obtained. Dr. Kouros Mohammadian of the UIC, the principal investigator of the UIC study, provided the data at the Census tract level to ORNL. Unfortunately, block group level data from UIC was not available. Since block group level geography is needed for the ORNL NMT study, the ORNL research team decided not to use the UIC data. Instead, the team decided to generate new sets of densities at the block group level for this study. Due to the intensive effort required in generating these densities, and the short time frame needed to

complete this task, the ORNL team decided to limit this estimation effort to the NHTS Add-on areas only. Although much smaller than a national scale, the total size of NHTS Add-on areas is more than adequate to cover the study regions selected for this project (Section 4.3).

3.5.1 Roadway Density

Road density is a measure calculated by dividing the total length of roadways within a given area (in this case, block group) by its size. To complete this task, roadway network links had to be “cut” where they intersect with a block group boundary using the GIS software. Then, the lengths of all roadway segments within a block group were aggregated to provide a total roadway length (in miles) for that given block group. The roadway density is then estimated by dividing this total roadway length by the size of the block group (in square miles). The Census Tract level roadway density can be generated in a similar matter; by aggregating all roadway lengths and area (square miles) from block groups within a given tract to obtain its tract roadway length and tract area first. These tract level roadway lengths were then divided by their corresponding tract sizes to create the tract level roadway densities. Figure 25 is an example showing the roadway density in the same downtown Dallas TX neighborhood as used in illustrating other data.

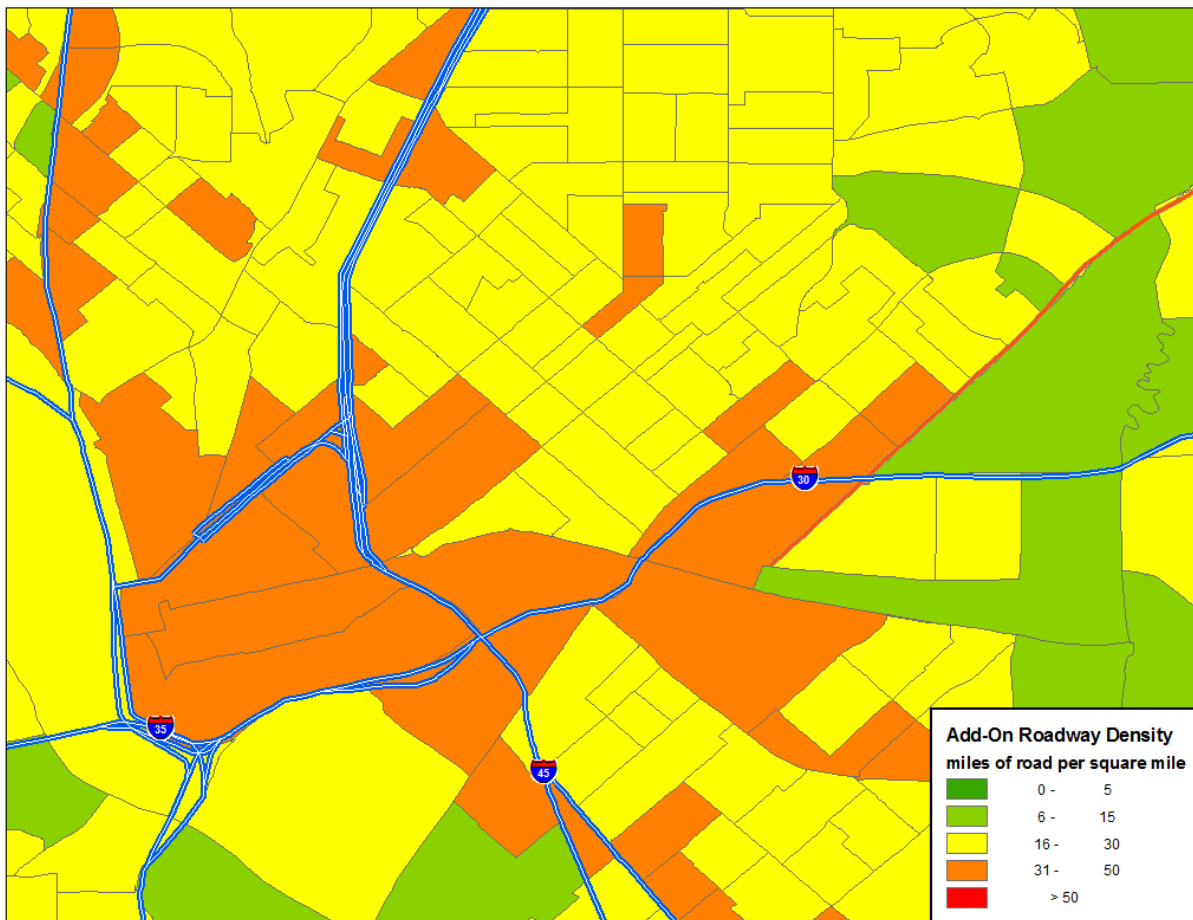


Figure 27. Example of roadway density in the Dallas, TX downtown neighborhood.

3.5.2 Intersection Density

Intersection density is a measure calculated by dividing the number of intersections within a given area by the size of the area. To calculate the number of intersections within an area (e.g., block group), road network nodes have to be filtered using the GIS software to remove nodes where two or more streets do not intersect (e.g., dead ends, connecting nodes, etc.). Intersections are then aggregated to give a total number of intersections within each block group. The total number of intersections was then divided by the size of the given block group (in square miles) to create its intersection density measure at the block group level. Similarly, the Census Tract level intersection density is created by aggregating the number of intersections and the area of all block groups within a tract to obtain the tract level number of intersections and area. The tract level number of intersections was then divided by its area to create the intersection density measure at the tract level. Figure 26 shows an example of intersection density in the Dallas TX neighborhood. Generally, especially in urbanized areas, higher roadway density areas are more likely to have higher intersection density. This can be seen when comparing Figure 26 to Figure 25.

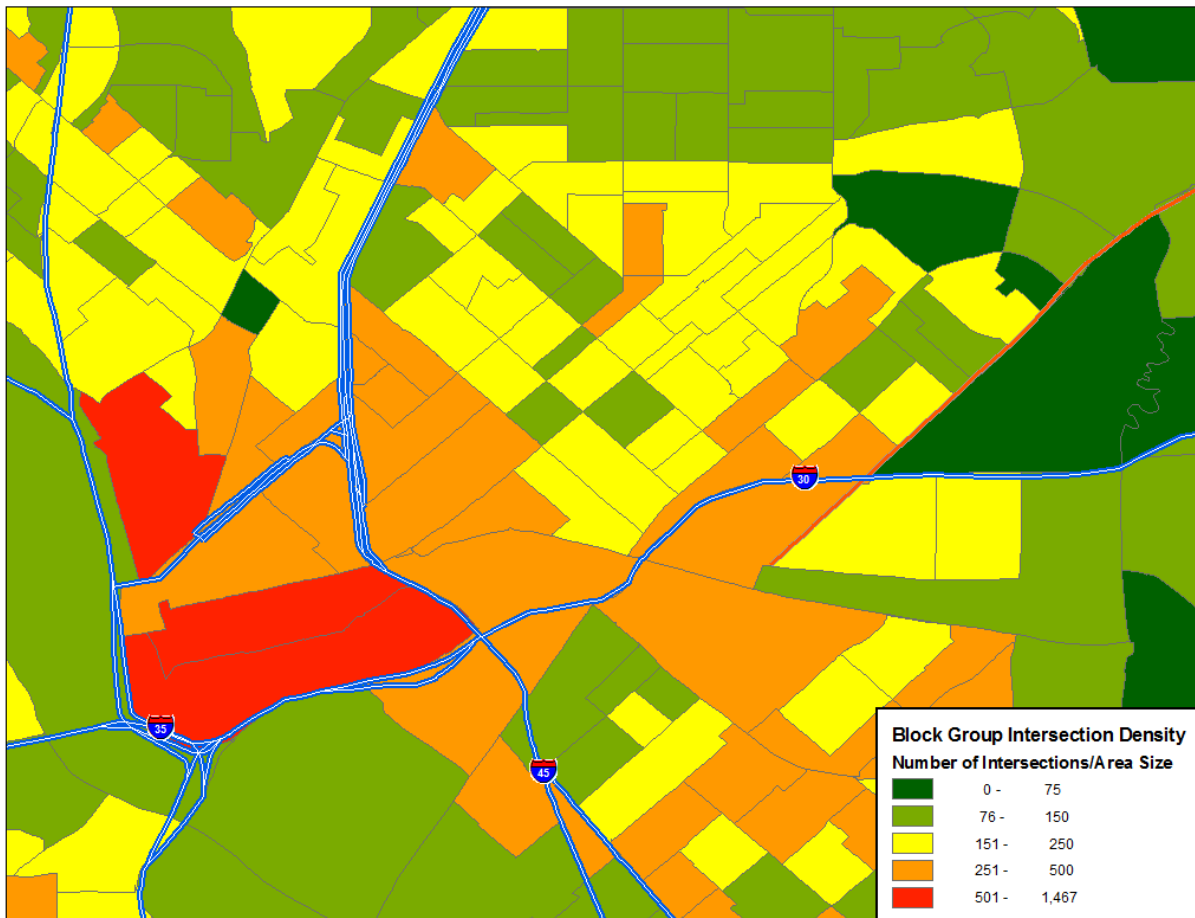


Figure 28. Intersection density at the downtown neighborhood of Dallas, TX.

3.5.3 Population Density

Population density for each block group was also generated using the GIS tool. The population density represents the density of residents living in the area (i.e., Census block group or tract); which is calculated by dividing the total population within the region by its size (in square mile). Figure 27 shows the general population density in the same example location in the Dallas, TX area. Population density for children ages 5 to 17 in the same region is shown in Figure 28. There is no surprise that most of the higher children concentration areas also are the higher general population density areas.

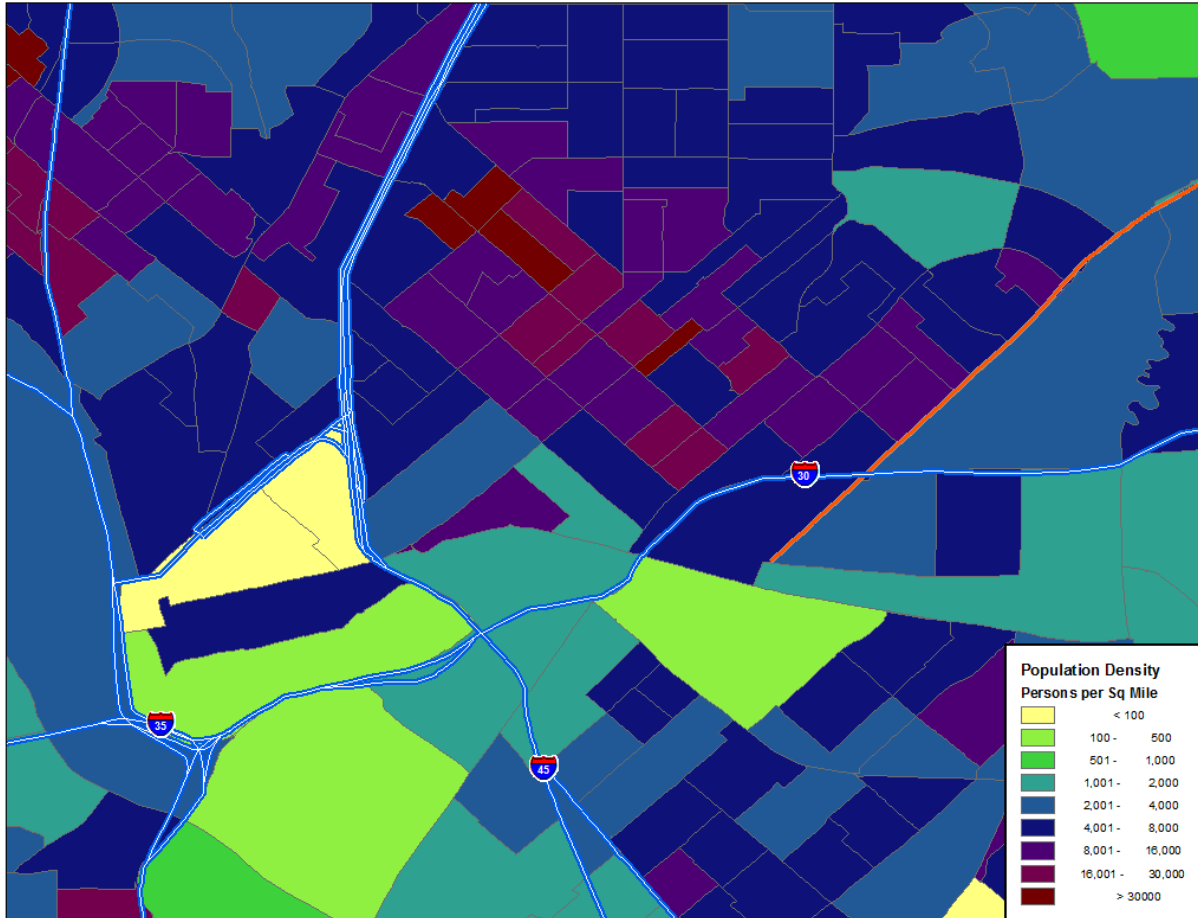


Figure 29. Population density in a downtown Dallas TX neighborhood.

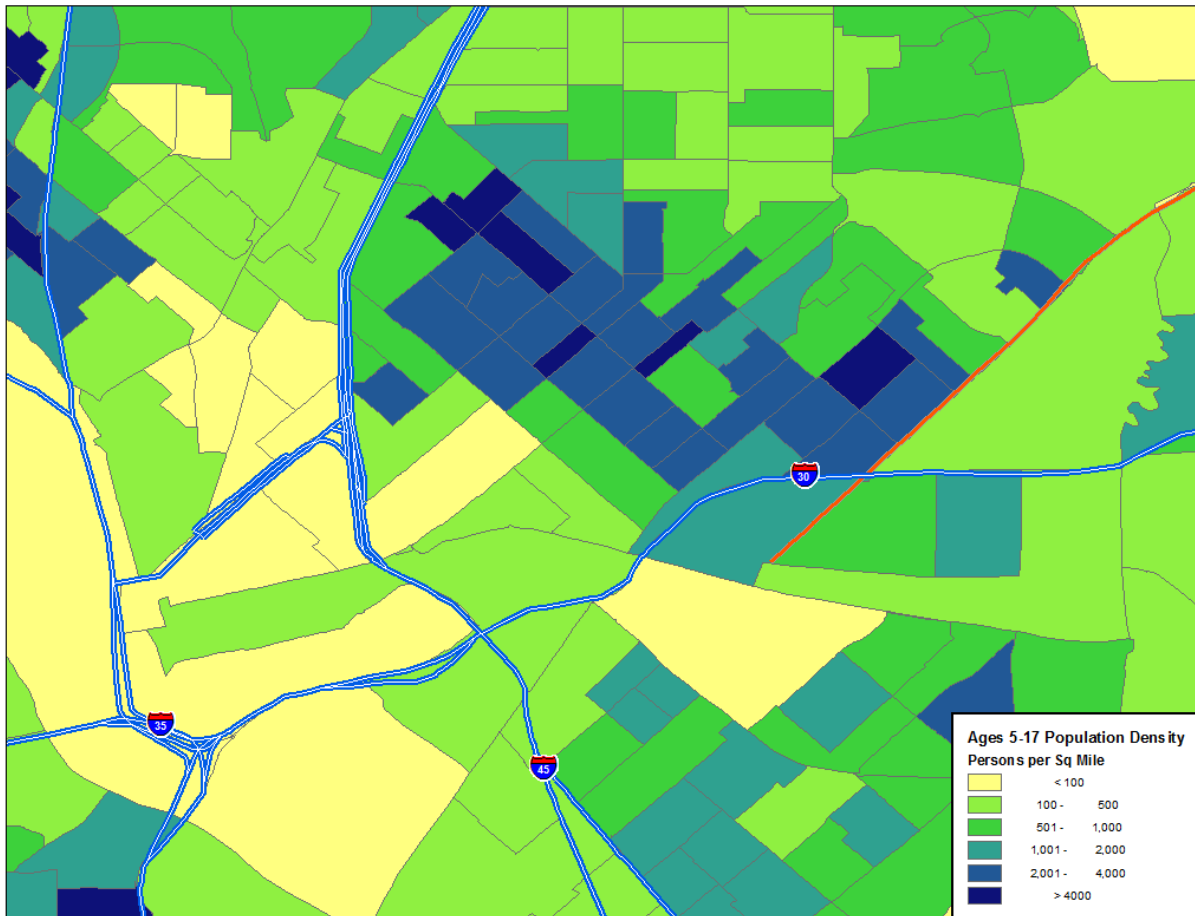


Figure 30. Population density for children ages 5-17 that reside in the Dallas, TX downtown neighborhood.

3.6 OTHER DATA SOURCES

Under this study, several “non-typical” data sources were also investigated during the research. These include data sources such as Center for Disease Control (CDC), National Center for Education Statistics (NCES), and Federal Emergency Management Agency (FEMA); where these data were collected for non-transportation study purposes. Some relevant data from these data sources are discussed in this section.

3.6.1 Center for Disease Control Obesity and Physical Inactivity Data

The association between obesity and physical inactivity has been studied, mainly in health care related research. The CDC provides data on physical inactivity and obesity for each county in the country.³² In addition to data on the percent of persons who are inactive or obese, categorical data on these percentages, coupled with the count of adults in a given county, was obtained from the CDC website. The data that best matched the NHTS survey timeframe was for the year

³² http://apps.nccd.cdc.gov/DDT_STRS2/NationalDiabetesPrevalenceEstimates.aspx?mode=PHY, accessed May 2, 2011.

2008. During that year, the percent of adults physically inactive in a given county ranged from 9.7% to 44.8%, while the obesity percentages for counties ranged from 11.7% to 43.7%. Figures 29 and 30, which were taken from the CDC website and are based on CDC data for 2008, provide visual displays of the estimated percentage for adults (20 years and older) that are physically inactive and those that are obese, respectively.

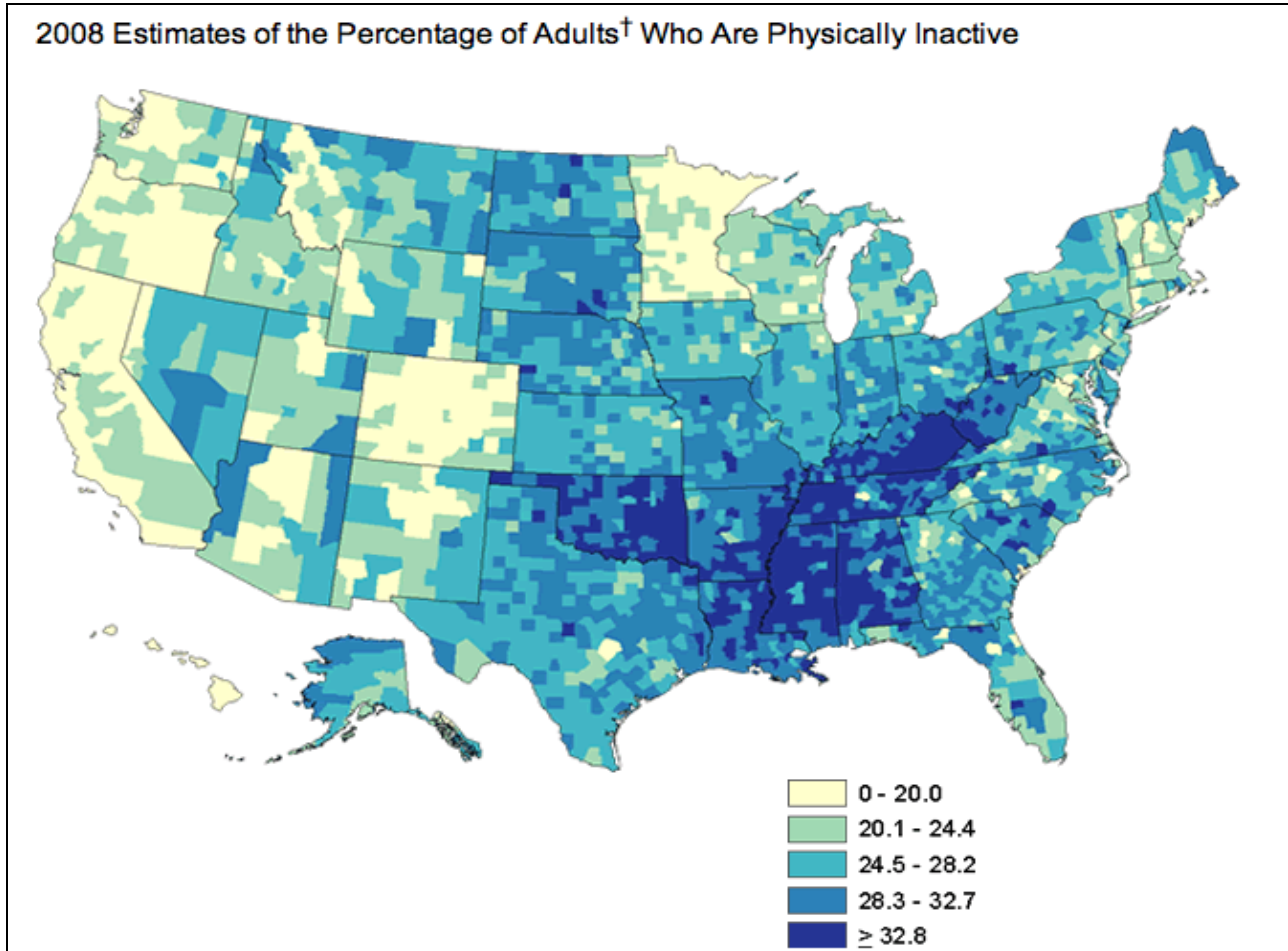


Figure 31. Percentage of adults who are physically inactive in 2008.

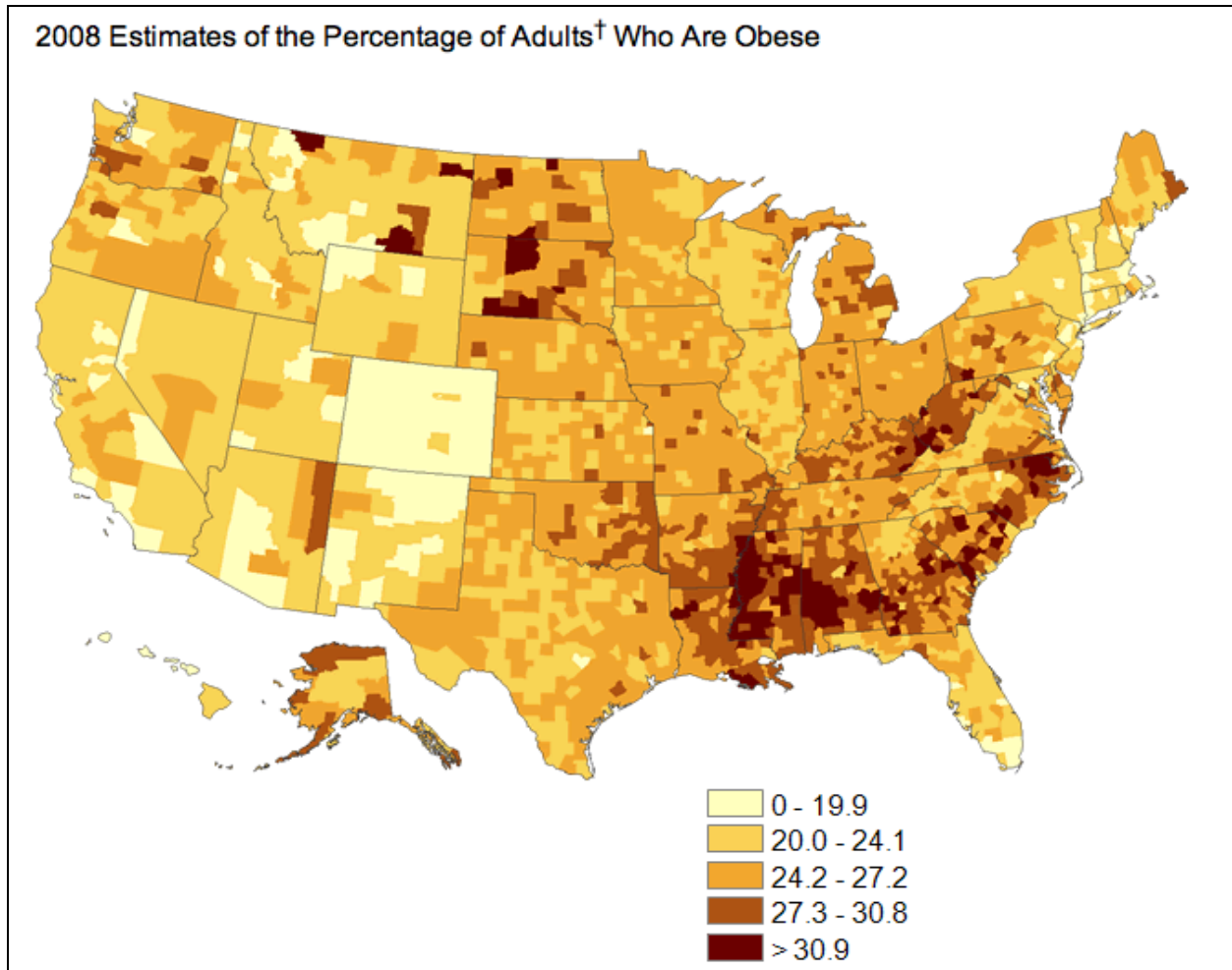


Figure 32. Percent of adults who are obese in 2008.

The obesity data from the CDC was integrated into the study database and used in the NMT modeling effort. Since the data is only available at county level, all block groups and tracts within the given county were given the same percentage (i.e., assuming uniform distribution within each county). This is certainly a data limitation.

3.6.2 School Location and Enrollment Data

Data on all elementary and secondary school (K-12) locations and their student enrollments are based on the 2008-2009 surveys conducted by the NCES³³. Both public and private school survey data were combined for the purpose of this study. Figure 31 shows an example of K-12 school student enrollment at the block group level. Note that the block group used here is based on the lat-long coordinates of the school (typically based on the address). It is the block group location where students attend the school, not necessarily the same block group where these students live.

³³ Information can be found at <http://nces.ed.gov/>.

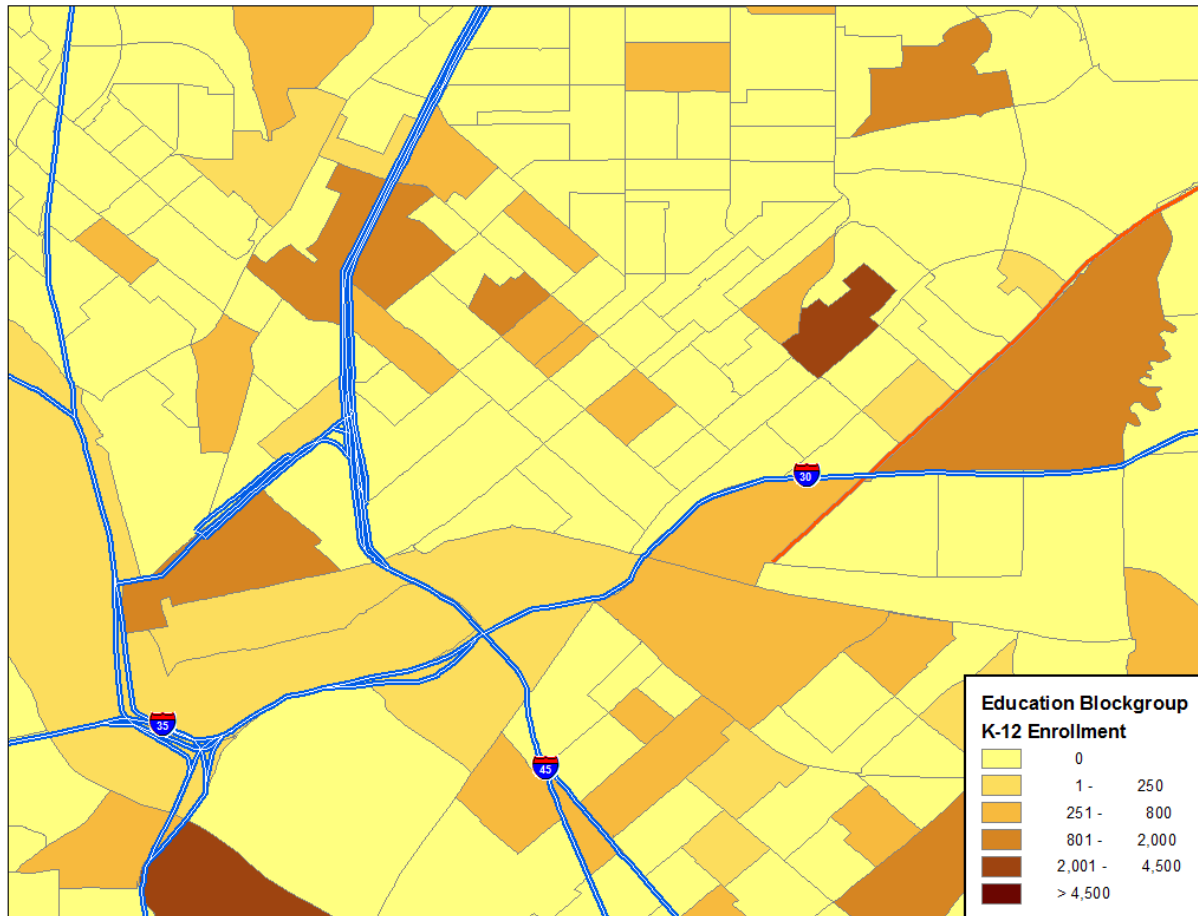


Figure 33. K-12 school student enrollments at the downtown Dallas, TX neighborhood.

3.6.3 Square Footage Occupancy Data for Land-Use Proxy and Land-Use Diversity

It has been suggested that activity-friendly neighborhoods promote healthy living of their residents. Mixed land use in a neighborhood is important in this regard, as it reflects the availability of destinations to which residents can walk or bike. There is no direct measure of land-use mix at the geographic scale covered by this study, however. A potential proxy measure, using NHTS Add-on variable “WHYTO,” was discussed in section 3.1.3. As pointed out previously, this measure might be misleading in some cases due to sampling limitations in the NHTS.

Square footage occupancy is a data set that the project team identified as another potential measure of land use mix. This data set is available from FEMA (data released as of May 2011) at the Census block level and is collected for use in emergency management purposes³⁴. FEMA categorizes the data into seven classes: residential, commercial, industrial, agricultural, religious, government, and education. With the exception of religious and agricultural, further sub-

³⁴ HAZUS is FEMA’s methodology for estimating potential losses from disasters. For more information see: <http://www.fema.gov/plan/prevent/hazus/>

classifications of land use are also provided. For example, the residential class can be further categorized into number of units per residence, temporary lodging, nursing home, etc.

To be consistent with the rest of the data used in this study, the block level square footage data was processed and aggregated to block group and Census tract levels for areas in the selected study regions. Figure 32 provides an example on the square footage occupancy data, showing the distribution of each building category in a pie chart for the downtown Dallas, TX neighborhood.



Figure 34. Square footage data for the downtown Dallas, TX area.

To examine how this data can be used to better reflect activities that might impact walk and bike in an area (e.g., shopping), several ratios based on various combinations of the major building occupancy categories were also analyzed. The “commercial” category was further separated into “retail” and “non-retail commercial” groups; where “retail” consists of FEMA’s subcategories of retail store, restaurant, and theater; and the “non-retail commercial” consists of the remaining subcategories within the original “commercial” category.

Moreover, the Shannon diversity index³⁵ is applied to measure the land-use mix diversity in each individual block group and tract. This method utilizes the concept of entropy to combined information, in this case, using the square footage data on major building occupancy types to generate a quantitative measure of the land-use mix. The higher the Shannon index is, the more mix of land-use in the region. The hypothesis here was that more diverse land uses would encourage more activities being conducted by walking or biking in a community. Figure 33 displays the block group level land-use diversity index in the Dallas, TX neighborhood. This diversity index and other ratios and densities created based on the square footage data are added to the long list of variables and examined under the modeling effort (discussed in Section 4).

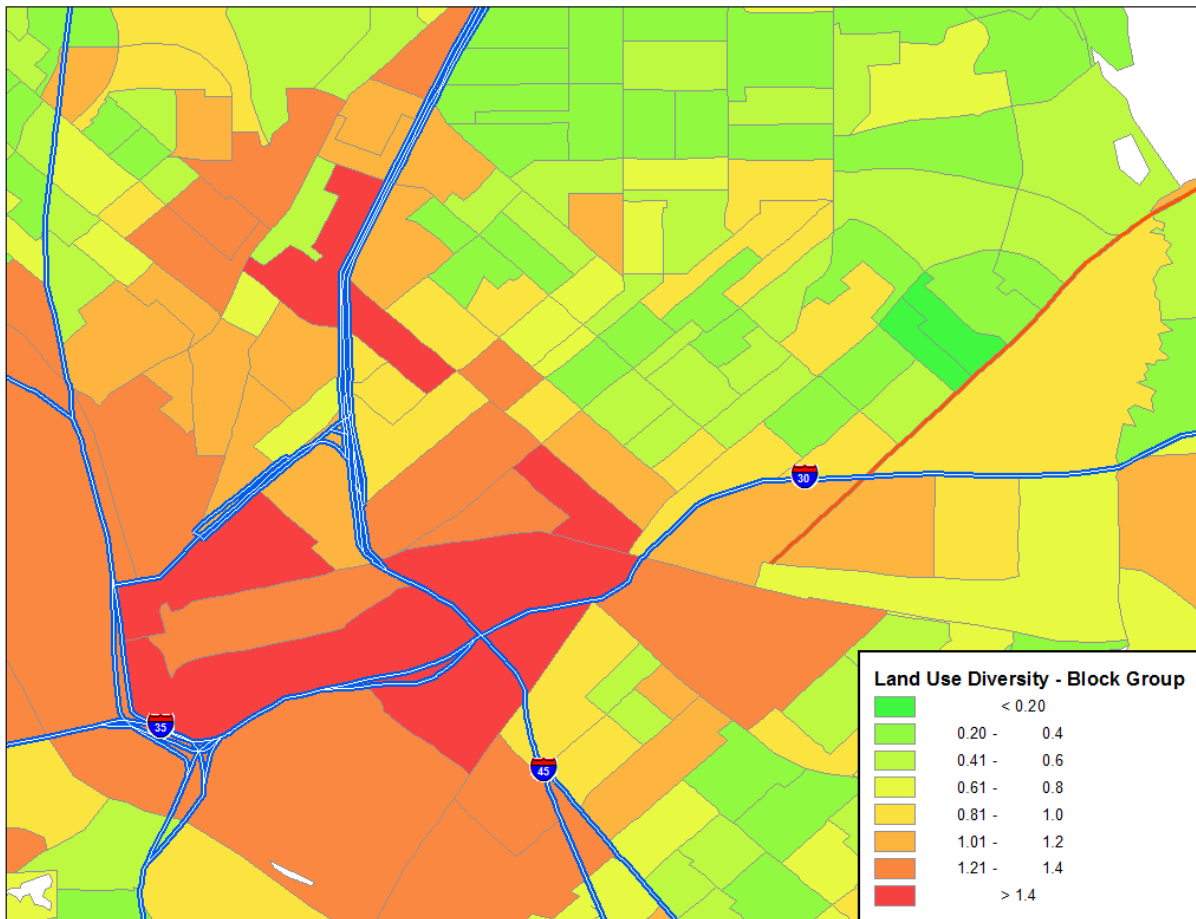


Figure 35. Block group level land-use mix diversity for the Dallas, TX neighborhood.

³⁵ Information on Shannon diversity index can be found at: <http://www.tiem.utk.edu/~mbeals/shannonDI.html> .

4. ANALYSIS AND MODELING

4.1 BACKGROUND

The database compiled from NHTS, ACS, and other data sources as discussed in Section 3 of this report, contains several hundred potential predictor variables. Not all of these potential predictor variables are expected to be correlated with walking or biking (i.e., the dependent variables) in a way that might be useful in practice. Some of the correlations may be too weak to be useful and some may be redundant, in the sense that, if anyone in a group of variables is included in a model, others in that group would no longer improve upon predictions. For this reason, as well as because the number of potential predictors is so large, it is critical to reduce the number of variables to a more tractable size. This can be carried out by utilizing statistical tools to eliminate some of the redundancies and weakly related variables.

Obviously the prediction model itself is critical in determining whether potential predictors might be useful in practice. An accurate physical model might expose very fine details about relationships between walking or biking and the various predictors. Unfortunately such physical models are unknown. Rather, data and judgment have to be combined in a way that leads to an approximate model that, while not necessarily a true physical description, at least emulates one in the sense that known changes in the predictors correspond to predictable changes in the walking or biking outputs.

As stated above, for the derivation of an approximate model to be feasible, the number of potential predictors should be tractable. To achieve this, screening procedures such as stepwise regression are often used to identify potential predictors that appear most likely to be useful in an approximate model. Because linear models are generally most straightforward, the screening procedures are generally based on underlying linear models. Thus procedures such as stepwise linear discriminant analysis or stepwise linear/logistic regression are used for the screening. Once the set of potential predictors is reduced to a tractable subset, more flexible nonlinear models can be developed.

The initial modeling effort conducted under this study is to perform screening procedures with the primary objective of reducing the extremely large number of potential predictors to a tractable subset of “discriminating factors for identifying walk-and-bike friendly communities.” The screening models ultimately derived and reported here reflect trial and error, and judgments made, in multiple passes through the data at varying levels of geographic detail. The objective was to derive a reasonable screening approach, not to demonstrate that the approach is optimal in some sense. Thus, the final model and results of the screening are described in more detail, but the steps to deriving the model and screening it are sketched in much less detail.

The discussion in Section 4.2 below focuses on screening at the national and NHTS add-on levels. A two-stage model was derived that, on the basis of one or more predictors, (1) identifies areas (block groups or census tracts) that support significant pedestrian or cycling activity, and (2) for areas so identified, describes the degree of the activity. The predictors in the two stages may be, but need not be, the same.

Because some of the potentially most useful predictors, specifically, walk scores and Nielsen employment data, are not feasible to obtain at the national or at the entire NHTS add-on levels within this study's budget, the study region for this project was further limited to a few selected areas of the NHTS add-on regions. Criteria used to select the refined study region and the resulting geographic areas are discussed in Section 4.3. The variables ultimately selected for these selected regions are addressed in Section 4.4.

4.2 SCREENING AT THE NATIONAL AND NHTS ADD-ON LEVEL

This section describes the method used to identify discriminating factors of walk/bike friendly communities, and sketches the model development. This model was the basis for the final choice of discriminating factors discussed in the later part of this Section. The development of this model involved multiple, some iterative, steps and judgments. For example, one of the earlier questions during the initial screening was to determine whether the dependent variable should represent walking or biking activity totals or rates per person.

Due to sampling limitation of the NHTS, many of the potential predictor variables are incomplete in the sense that they are missing for some data records (i.e., no data at block groups or census tracts level). Potential predictors to be considered in the stepwise models were determined based on both the degree of data completeness and the strength and importance of the predictors themselves. The strength and importance, specifically, are best assessed by considering the stepwise procedure output. That is, which variables are selected and what is the strength of their correlations in the model based on them. Thus, the model derivation is an iterative process, involving many steps, only some of which are summarized and discussed here.

Preliminary variable screening models were run after the majority of the input data was prepared, and initial screening results were used to check and refine the choice of data. For example, an initial trip purpose classification based on the NHTS WHYFROM variable (why traveler took a trip at trip origin) was replaced with one based on the NHTS WHYTO variable (why the traveler went to the trip destination, which was described in Section 3.1.3). The SAS logistic (logistic regression), REG (linear regression), and STEPDISC (discriminant analysis) procedures were used to perform the stepwise model fits.³⁶

4.2.1 Review of Walk/Bike Distributions

Based on examination of initial results and the consideration of available NHTS samples, the project team decided to base models of pedestrian activity on block group data and models of biking on Census tract data. Using the NHTS data for block groups where households were sampled, distributions of annual walk trip and mileage trip rates per person were examined. Similar analysis was also conducted to study the distributions of annual bike trip and mileage rates per person at the Census tracts. Since the NHTS, including the Add-on, was not designed to represent detailed geographic regions such as block groups or Census tracts, it is not a surprise to find a large number of the block groups (or Census tracts) with no NHTS sampled households; especially when considering walk/bike travel activities in particular.

³⁶ SAS Institute Inc. 2010. SAS/STAT® 9.22 User's Guide. Cary, NC: SAS Institute Inc.

Based on 2009 NHTS data, about 62% of the total 63,953 block groups that contain NHTS sampled households have no walking trips in 2009, and about 91% of the total 34,536 Census tracts with NHTS sampled households have no biking trips. Although pedestrian and cycling activity was observed for relatively few block groups or census tracts, where it was observed, it was sometimes substantial. Figure 34 shows the distribution of block groups with positive walk activity rates; while Figure 35 displays the distribution of Census tracts with positive bike activity rates. Based on the NHTS, a majority of block groups or Census tract with positive walk/bike activities are in the “1-5 times a week” category, and accounted for over 40% of these regions for either walk or bike.

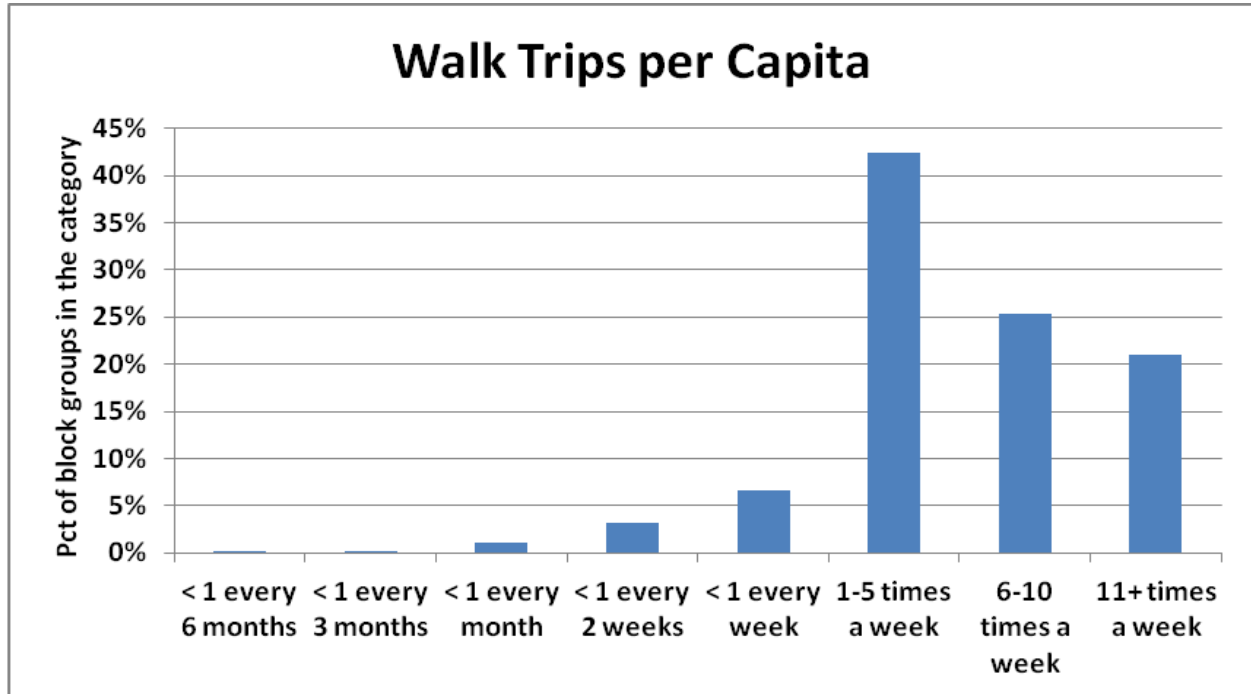


Figure 36. Distribution of block groups by walk trips frequency, based on block groups with positive walk trip rates from 2009 NHTS data.

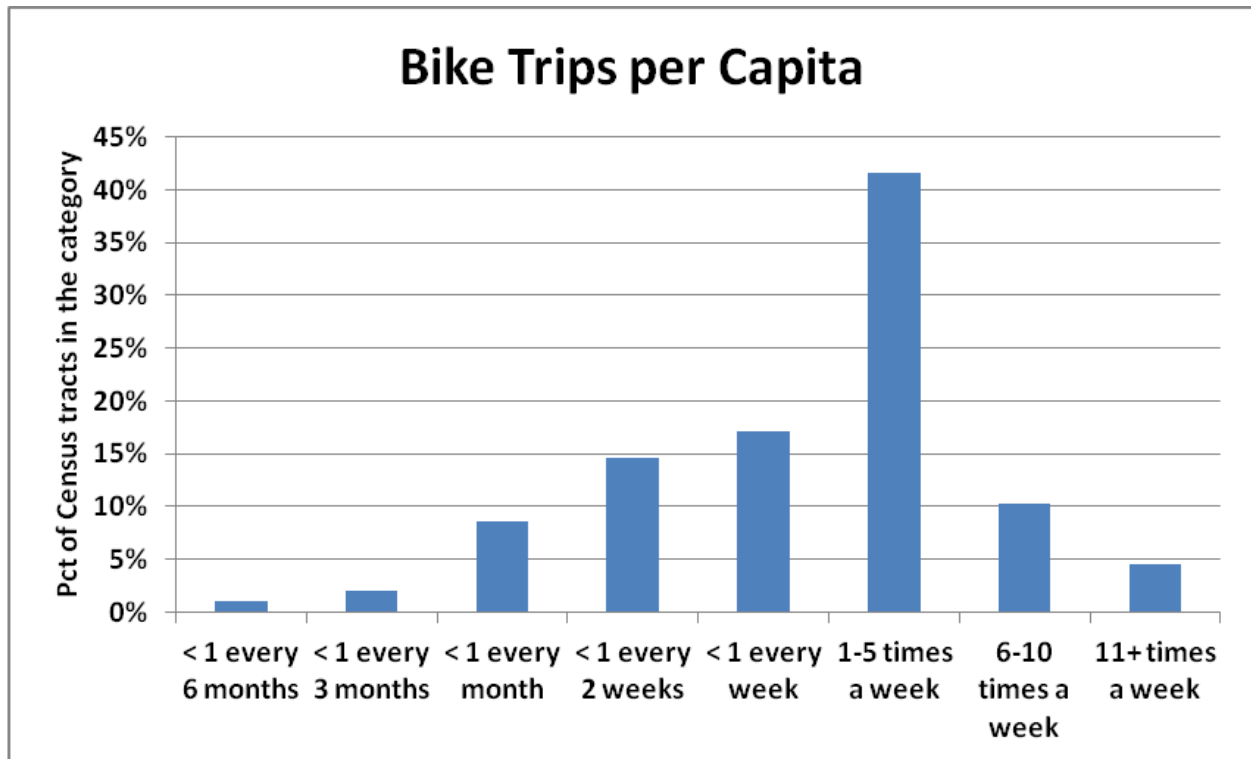


Figure 37. Distribution of Census tracts by bicycle trips frequency, based on Census tracts with positive biking activities from 2009 NHTS data.

Furthermore, as shown in these figures, for areas where walking/biking activity is observed (i.e., at least one walk/bike trip reported in the block group or Census tract), the distributions are somewhat skewed. This suggests a two-stage model, with a first stage to distinguish between areas with and without appreciable activity, and a second stage to model the rate of the activity. In the second stage, with areas of zero activity excluded, a log transformation of the positive activity rates could be used to accommodate the skewed distributions if necessary.

4.2.2 First Stage Models

For the first stage models, areas (block groups or Census tracts) are classified according to whether there was activity in them or not. Two statistical procedures were then considered for modeling the classification: stepwise linear discriminant analysis and stepwise linear logistic regression. Results of both approaches are discussed in this report. However, because discriminant analysis assumes multivariate normality of the predictors and many variables under consideration in the models are clearly not fulfilling this assumption, logistic regression should be the preferred approach of the two. In both approaches, variables are iteratively included or excluded from a predictive model. Specifically, variables were added to the tentative model if their statistical significance levels were 0.0001 or less. At the same time, variables were dropped from the tentative model if their significance levels exceeded 0.0001. The use of this more restricted than usual significant level (typically at 0.5) was intended to reduce the number of variables being included into the model; as mentioned at the beginning of this Section, the source database contains several hundreds of candidate variables.

Detailed discussions of the stepwise logistic and linear discriminant procedures can be found in the SAS manual³⁷ and in the many references cited there. In particular, the Cox & Snell R-Square statistic³⁸, an analog of the R-Square statistic from linear regression, was calculated as a measure of goodness-of-fit for the logistic regression models. Note that, the purpose of this study is not to develop demand models for predicting the frequencies of walk/bike travels nor trip lengths from such activities. Instead, the modeling effort of this project is aimed at identifying major factors that influence the communities (block groups or census tracts) to, or not to, walk/bike. In other words, the screening effort conducted here is to determine factors that are not only correlated with walk/bike activities but also can be used to distinguish members of the two groups, i.e., walk or not walk block groups. This applies similarly to the two groups of bike and not bike Census tracts. A low value of the R-Square, which indicates the fitted regression model is not a good predictor of the response under consideration, is therefore not a concern for the purpose of this study.

As an example, Table 5 illustrates the first stage results of the stepwise logistic regression computed from a preliminary run with pedestrian trips per person. This table lists the variables selected in the stepwise logistic regression, their corresponding parameter estimates, standard errors, and nominal significance levels.³⁹ Predictors in these tables are listed in ascending order of their nominal significance levels (i.e., the most significant one listed first) in the stepwise analysis. Additional modeling runs were conducted using pedestrian mileage per person, bike trips per person, as well as bike mileage per person as the dependent variables. All results produced from these screening processes are included in Appendix C of this report.

Table 5. Estimated Walk Trips Per Person Stepwise-Selected Logistic Regression Model
Number of Observations Used: 46,354, Cox & Snell R-Square: 22.2

Predictor Variable	Parameter Estimate	Standard Error	Nom. Sig. Level
Percent Weighted Trips in Destination Block Group for the Purpose Home-Based Social-Recreational, All Modes (NHTS+: HBSOCREC_DestPct)	1.6486	0.0542	4.E-203
Percent Persons in BG with Educ of HS or Less (NHTS: PctHSLess)	-0.6153	0.0459	5.1E-41
Sample HHs in BG (NHTS: NHHs)	0.2452	0.0199	9.7E-35
Average Worker Count of HHs in BG (NHTS: AvgWorkers)	0.1961	0.0190	5.1E-25
County-level Percentage of Adults Who Are Obese (Other: PctObese)	-0.0369	0.0037	1.4E-23
Average Vehicle Count of HHs in BG (NHTS: AvgVehCount)	-0.1546	0.0161	6.7E-22
Sample Persons in BG (NHTS: NPers)	0.0922	0.0097	1.6E-21
Percent Range CY Pop 25+, College Graduate (Claritas: HBPCOLGRD)	0.0073	0.0008	1.6E-21

³⁷ Ibid.

³⁸ See the section “Generalized Coefficient of Determination” on page 3,945 of the SAS documentation.

³⁹ Significance levels are “nominal” in stepwise regressions because the variables are selected on the basis of significance or correlation or related criteria.

Table 5. Estimated Walk Trips Per Person Stepwise-Selected Logistic Regression Model (Cont.)

Predictor Variable	Parameter Estimate	Standard Error	Nom. Sig. Level
Percent Drivers in BG (NHTS: PctDrivers)	-0.5156	0.0595	4.2E-18
Percent Weighted Trips in Destination Block Group for the Purpose Home-Based Other, All Modes (NHTS+: HBO_DestPct)	0.3637	0.0495	1.9E-13
Average Driver Count of HHs in BG (NHTS: AvgDrivers)	0.1901	0.0272	2.6E-12
CY Median Year HU Structure Built (Claritas: HBHMEDYR)	-0.0055	0.0008	7.8E-12
Percent Range CY Housing Units, Renter Occ (Claritas: HBHTNRNT)	0.0039	0.0006	3.2E-11
Intercept	10.3434	1.5786	5.7E-11
Percent Persons in BG with Some College/Bachelor's (NHTS: PctCollege)	-0.2772	0.0440	2.9E-10
Wtd Pct HHs in BG with English Interviews (NHTS: PctEngHH)	-0.3568	0.0599	2.6E-09
CY Own Occ HUs Median Value (Claritas: HBHMEDHS)	0.0000	0.0000	3.5E-09
Wtd Pct HHs in BG Owning HH (NHTS: PctOwnHH)	-0.1918	0.0349	3.9E-08
Households by Occupancy Status, Vacant (ACS: B25002e3)	0.0003	0.0001	3.8E-07
Employee Range CY Employees per square mile (Claritas: HBEEMPDN)	0.0000	0.0000	4.4E-07
Percent Range CY Pop, Asian (Claritas: HBPRCASN)	-0.0067	0.0013	4.6E-07
Percent Weighted Trips in Destination Block Group for the Purpose Home-Based Work, All Modes (NHTS+: HBW_DestPct)	-0.3245	0.0661	9.2E-07
Percent Males in BG (NHTS: PctMales)	0.1986	0.0418	2.0E-06
Workers by Means Of Transportation To Work, Public transportation (excluding taxicab): Subway or elevated (ACS: B08301e13)	0.0017	0.0004	4.0E-06
Town & Country Indicator (Claritas: HBHUR_TC)	-0.1197	0.0277	1.5E-05

4.2.3 Second Stage Models

In the second stage model, stepwise linear (least squares) regression was used to model the logs of the activity rates as functions of the potential predictors. Similar to the first stage models, statistical significance levels were set at 0.0001 or less. Table 6 illustrates the results of the stepwise linear regression which was based on a preliminary run with pedestrian trips per person. Similar to first-stage tables, it lists predictors selected in the stepwise regression in ascending order of nominal significance level. Corresponding parameter estimates for the predictor variable and their standard errors are also shown in the table. Since only positive activity rates are considered in the model (i.e., excluding block groups that has no walk trips from NHTS sampled households), the value of the R-square is higher than the R-square from its first stage

counterpart model. This means there is a stronger correlation between the predictor variables and the activity rate of walk trips, where such activities exist (e.g., a positive walk trip rate).

Table 6. Estimated Walk Trips Per Person (When Positive) Stepwise-Selected Least Squares Regression Model

(Number of Observations Used: 18,388, R-Square: 43.8)

Predictor Variable	Parameter Estimate	Standard Error	Nom. Sig. Level
Intercept	16.2957	0.6283	1.E-145
CY Median Year HU Structure Built (Claritas: HBHMEDYR)	-0.0073	0.0003	2.E-115
Households in BG (NHTS: TotHHs)	0.0001	0.0000	4.E-105
Percent Weighted Trips in Destination Block Group for the Purpose Home-Based Social-Recreational, All Modes (NHTS+: HBSOCREC_DestPct)	0.3644	0.0214	8.6E-65
Wtd Pct HHs in BG Owning HH (NHTS: PctOwnHH)	-0.1895	0.0125	8.8E-52
Average HH Size of HHs in BG (NHTS: AvgHHSIZE)	0.0756	0.0051	7.9E-49
Sample HHs in BG (NHTS: NHHs)	-0.0579	0.0042	1.7E-43
Persons by Sex By School Enrollment By Level Of School For The Population 3 Years And Over, Total (ACS: B14002e1)	-0.0001	0.0000	4.7E-38
CY Own Occ HUs Median Value (Claritas: HBHMEDHS)	0.0000	0.0000	1.2E-36
Median Age, Total (ACS: B01002e1)	0.0061	0.0005	1.2E-33
County-level Gradation of percentage of physically inactive adults by total number of physically inactive adults in 2008 (Other: LIPERTOT08)	-0.0204	0.0019	4.0E-26
Average Vehicle Count of HHs in BG (NHTS: AvgVehCount)	-0.0517	0.0052	2.0E-23
Sample Persons in BG (NHTS: NPers)	0.0160	0.0019	9.4E-17
Wtd Pct Afr Am. HHs in BG (NHTS: PctAfAmHH)	0.1537	0.0187	2.0E-16
Employee Range CY Employees per square mile (Claritas: HBEEMPDN)	0.0000	0.0000	2.0E-16
Percent Weighted Trips in Destination Block Group for the Purpose Home-Based Other, All Modes (NHTS+: HBO_DestPct)	0.1491	0.0200	1.0E-13
Wtd Pct Hisp HHs in BG (NHTS: PctHispHH)	0.1067	0.0151	1.4E-12
Persons in BG (NHTS: TotPers)	0.0000	0.0000	3.1E-12
Workers by Means Of Transportation To Work By Travel Time To Work, Car, truck, or van - drove alone: 20 to 24 minutes (ACS: C08134e15)	-0.0002	0.0000	1.3E-11
Percent Weighted Trips in Destination Block Group for the Purpose Not Home-Based, All Modes (NHTS+: NHB_DestPct)	-0.1303	0.0201	1.0E-10
Percent Range CY Pop, White (Claritas: HBPRCCAUI)	-0.0016	0.0003	3.0E-10
Percent Range CY Pop, Foreign born (Claritas: HBPFORBN)	0.0022	0.0004	9.7E-10

Table 6. Estimated Walk Trips Per Person (When Positive) Stepwise-Selected Least Squares Regression Model (Cont.)

Predictor Variable	Parameter Estimate	Standard Error	Nom. Sig. Level
Percent Workers in BG (NHTS: PctWorkers)	0.1400	0.0229	1.0E-09
Household Income In The Past 12 Months, \$100,000 to \$124,999 (ACS: B19001e14)	0.0004	0.0001	2.4E-09
Percent Drivers in BG (NHTS: PctDrivers)	-0.1150	0.0196	4.1E-09
Percent Range CY Families Below Poverty (Claritas: HBPLTPOV)	0.0026	0.0005	6.1E-09
County-level Percentage of Adults Who Are Obese (Other: PctObese)	0.0070	0.0012	1.7E-08
Persons by Sex By Educational Attainment For The Population 25 Years And Over,Female:7th and 8th grade (ACS: B15002e23)	-0.0009	0.0002	5.8E-08
Persons by Race, American Indian and Alaska Native alone (ACS: B02001e4)	-0.0004	0.0001	1.0E-07
Average Worker Count of HHs in BG (NHTS: AvgWorkers)	0.0549	0.0112	9.6E-07
Percent Persons in BG with Grad/Prof Degree (NHTS: PctGrad)	0.0780	0.0163	1.7E-06
Persons by Sex & Age, Male: 50 to 54 years (ACS: B01001e16)	-0.0004	0.0001	1.8E-06
Persons by Sex & Age, Female: 30 to 34 years (ACS: B01001e36)	0.0003	0.0001	4.4E-06
Persons by Race, Black or African American alone (ACS: B02001e3)	-0.0001	0.0000	4.5E-06
Percent Range CY Housing Units, Owner Occ (Claritas: HBHTNOWN)	-0.0010	0.0002	1.8E-05
Percent Males in BG (NHTS: PctMales)	0.0659	0.0155	2.1E-05
Percent Range CY Pop 20-64 (Claritas: HBP20TO64)	0.0021	0.0005	2.2E-05
Percent Range CY HUs, Built 2000 or Later (Claritas: HBHRCENT)	0.0013	0.0003	3.7E-05
Workers by Travel Time To Work, 90 or more minutes (ACS: B08303e13)	0.0004	0.0001	3.9E-05
Household Income In The Past 12 Months, Less than \$10,000 (ACS: B19001e2)	-0.0003	0.0001	4.7E-05
Town & Country Indicator (Claritas: HBHUR_TC)	0.0388	0.0095	4.8E-05

Models developed using only positive walk/bike trip activity, however, might not necessarily be appropriate for uses to describe relationships among the predictor variables in the models and the walk/bike activity variables for regions that do not have any walk/bike trips. Nevertheless, these second stage models do provide indications to predictor variables that could have strong associations with walk/bike activity, which indeed supports the goal of the screening process.

Note that, because so many variables are being considered in so many analyses, a few of the selected predictors are likely to be false positives. Similarly, variables may be excluded because their effects are masked by other variables or simply because of the statistical limitations. For the bicycle models, which are subject to the more substantial statistical limitations of the NHTS data, many of the predictors were selected for only one particular model.

As part of the evaluation of the models, regression residuals from the national models on various potential predictors selected in the stepwise regression were examined. The distribution of these regression residuals appears to be symmetric and no apparent patterns exists, which suggests that the log transformation was appropriate. Note that the national models were conducted as a part of the preliminary analysis effort to screen the data as well as to generate the conceptual modeling framework for the NMT study, and as such results from this part of the project were preliminary in nature and are included in the Appendix C for references. The main focus of this ORNL NMT project was on the rescaled geographic regions, which is discussed next under Section 4.3 of this report.

4.3 DATA LIMITATIONS AND REDUCED GEOGRAPHIC COVERAGE

Data availability precluded the use of national level data in the analysis in several ways. First, walk and transit scores at the block group and Census Tract level for the entire country were prohibitively expensive. Second, block group level employment data from Nielsen Claritas with a national coverage is also beyond the affordability of the project budget. Next, the computational effort involved in computing roadway and intersection density for the entire country at the desired levels of geography was similarly restricted. Lastly, the availability of large samples in the NHTS data is limited to 14 states (including instances of additional sampling in certain areas of two of these states), and areas in two additional states. Given the presumed importance of these factors, the decision was made to reduce geographic coverage to a handful of areas representative of the country as a whole.

The following criteria were taken into account when selecting areas for use in the remainder of the analysis:

- Within an NHTS add-on area,
- Transit score availability,
- GTFS availability, and
- Population size.

Specifically, given the NHTS' role as the primary data source, no area outside of an add-on was considered for inclusion in the reduced geographic coverage. Furthermore, transit score or GTFS data was preferred in the determining areas of selections, but it is not necessarily the sole determining factor for an area. Finally, a reasonable mix of large and small-to-medium sized metropolitan areas were sought to complete the reduced area used for analysis.

Table 7 below details the 9 geographic areas selected for the continuing analysis. The areas range from small cities (Cedar Rapids, IA) to large metropolitan areas (Los Angeles, CA), with a

primary focus on larger areas that would presumably be more likely to have transit and walking and bicycle use.

4.4 MODELING USING DATA FROM SELECTED REGIONS (WITHOUT NIELSEN EMPLOYMENT DATA)

The two-stage approach to stepwise variable selection discussed in Section 4.2 is applied here. The reduced geographic region as described above allows the project to expand the modeling effort to include additional candidate predictor variables, such as walk scores, intersection and road densities, real estate square footages, and land-use diversity measure. The employment data purchased from Nielsen was not included in the modeling effort until the late stage of the project due to delays in reaching a disclosure agreement between ORNL and Nielsen.

Table 7. Selected Study Regions

Region	Definition	Population	Transit Score for Main City?	GTFS?	Block Groups with NHTS HHs	Total Block Groups
Washington, DC (Virginia portion)	Cities of Alexandria, Falls Church & Fairfax, and counties of Arlington and Fairfax	1,464,216	yes (in part)	yes	315	798
San Diego, CA	San Diego county	3,095,313	yes	yes	1,542	1,762
Los Angeles, CA	Los Angeles county	9,818,605	yes	yes	2,382	6,351
Miami, FL	Miami-Dade county	2,496,435	yes	yes	701	1,222
Dallas, TX	Collin, Dallas, Denton, Kaufman & Rockwall counties	3,994,781	yes	yes	1,391	2,215
Austin, TX	Hays, Travis & Williamson counties	1,604,052	yes	yes	460	709
Cedar Rapids, IA	Linn county	211,226	no	no	122	133
Research Triangle area, NC	Chatham, Durham, Orange & Wake counties	1,365,886	yes	yes	232	480
Madison, WI	Dane county	488,073	no	yes	110	260

Note that, the NHTS block group or census tract population estimates (NHTS TotPers) were used for rate-per-person denominators, rather than the ACS population counts (variable B01001e1). The ACS are probably better counts in their own right; however, the NHTS counts are more suited to the NHTS numerators in the rates because of the way NHTS samples were weighted.

Initially, the modeling approach was applied to each of the four dependent variables: walk trips and walk mileage, and bike trips and bike mileage. A fifth dependent variable was also considered: total walk trips that include access/egress to/from public transit. Ordinarily in the NHTS, walking to/from public transit is counted as a transit trip (along with the transit trip itself) because transit in that case is the primary transportation mode. Pedestrian trips to/from public transit are often significant in their own right, however. The NMT team decided that it makes sense to also consider a version of walk trips augmented to include transit access/egress. Because pedestrian mileage was not recorded in the NHTS for the pedestrian transit access/egress, a transit-augmented version of walk mileage is unavailable.

Figure 36 below is the analog for the augmented pedestrian trip frequencies of Figure 34. Not surprisingly, walk trip frequencies are higher when transit access/egress trips are included. Based on Figure 36, more than a quarter of the block groups (with positive walk activity and transit availability) made an average of more than 10 walking trips per week; while only about 5% of the block groups showed the same level of activity in Figure 34.

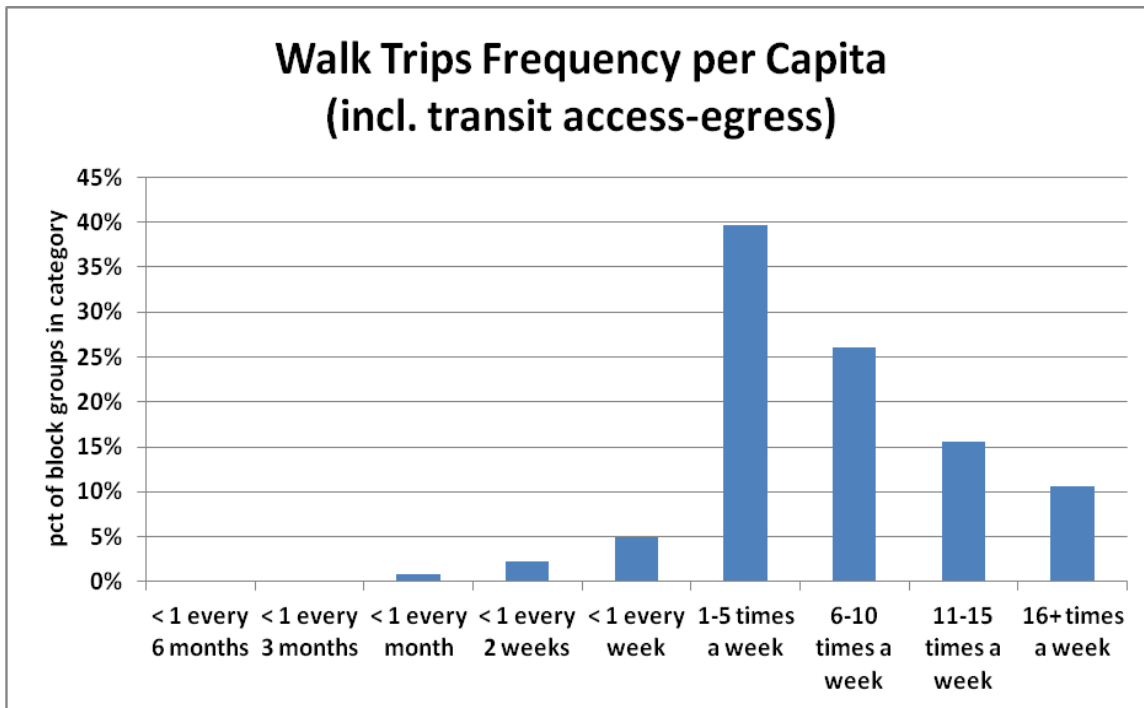


Figure 38. Distribution of block groups by frequency of walk trips (including transit access/egress) per capita.

Sets of tables for each of the five dependent variables mentioned above, i.e., modeled with all but the Nielsen employment data, are included in Appendix D. For each dependent variable, the two-stage modeling approach produces:

- (1) a first-stage table of predictor variables selected by stepwise logistic regression for classifying block groups (pedestrian analyses) or census tracts (bicycle analyses) according to whether they have significant pedestrian or bicycling activity at all;

- (2) a table of variables selected by stepwise discriminant analysis, as an alternative to the table of variables selected by stepwise logistic regression in the first stage classification; and
- (3) for block groups or census tracts where walking or biking activity is indicated in the first stage, a second-stage table of predictor variables selected by stepwise least squares regression for quantifying (i.e., estimating the amount of) activity.

Thus, the following four sets of three tables, each set analogous to Tables 5-6 for walk trips (see Section 4.2), are listed in Appendix D. The first 6 tables are based on stepwise regression models, including Tables D.1-D.2 for walk trips, Tables D3-D4 for walk mileage, and Tables D.5-D.6 for walk trips augmented with pedestrian transit access/egress. Additional Tables D7-D9 are results generated from the Discriminant Analysis models for walk trips, walk mileage, and walk trips including transit access/egress, respectively.

Similar to the national level data analysis, the symmetric distributions of the residuals from these models support the use of the log transformation and the linear modeling to quantify walking and bicycling activity. The number of block groups or Census tracts used in each model varies, depending on the data availability of selected predictor variables for the given model. For example, bike mileage per person model contains many more points (i.e., Census tract records) than the bike trips per person model. This is because certain variables were selected for the bike trips per person model but not the bike mileage per person model, and values of these variables were missing in some tracts thus these records were left out from the model. The variable selection procedures were nevertheless restricted to block groups or census tracts; it uses the largest subsets of the data records available for all candidate predictors.

For these screening models (using data in the reduced geographic regions), variables selected in the stepwise procedures are required to have nominal significance level of 0.01 or less to enter and remain in the corresponding linear or logistic or discriminant model. The 0.01, rather than the 0.0001 as used in modeling the national data set, allows a greater variability of selecting a good set of predictor variables for further analyses. Of course the selected variables also depend on the underlying models, which for these screening analyses are all linear in their parameters. Dependency of variables will be dealt with once the major factors (i.e., predictor variables) are identified.

4.4.1 Review of Initial Results – from the Linear Screening Model

The estimated results from the linear logistic or discriminant models described above are a set of probabilities. These probabilities can be considered as “scores” which, when comparing with a “cutoff” value, allows one to classify the data into two groups. In other words, a “good” model will produce scores (estimates) that allow separation of the two be as “far apart” as possible.

Figure 37 shows a hypothetical case of what might be considered as an “ideal” outcome from the model estimates (probabilities). In this hypothetical case, the red and blue groups are mostly separated. Using the cutoff point of 0.5, the overlapping portion on the right (i.e. >0.5) can be viewed as false positive cases (i.e., “no walk trip” case being classified in “with walk trip” group) and those on the left are false negatives (i.e., “with walk trip” case being classified in “no walk trip” group).

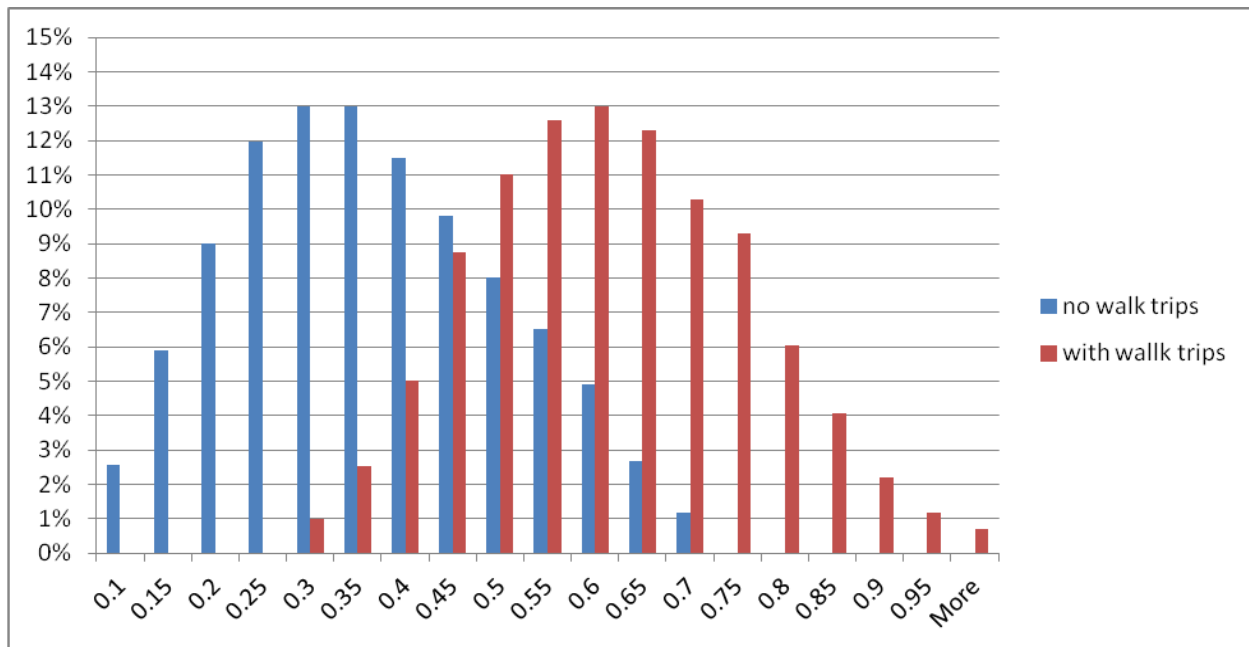


Figure 39. Distributions showing a hypothetical "ideal" scenario with discriminant analysis model.

Because of NHTS sampling limitations, a block group recorded as “no walk trips” might be based on only one sampled household that happened to make no walking trips on the assigned travel day. This block group might have similar characteristics, measured by the selected predictor variables, as those block groups that did walk; and as a result, the model assigns a higher probability for that block group. The reverse could also be possible, where a low value of probability is assigned to a block group containing walk trips, based on NHTS samples. Therefore, a pattern as shown in the hypothetical Figure 36 is not expected, nor realistic, for the NHTS data. The results generated from the linear screening model discussed in this Section (i.e., selected geographic region and without Nielsen employment data) are shown in Figure 38.

For the walk trips model, using 2009 NHTS data for the selected geographic region, about 7,100 block groups contain usable data (i.e., non-missing values in the predictor variables) for NHTS sampled households. The number of block groups is about evenly split between those with and without walk trips. Thus, a reasonable “cutoff” point would be at 0.5. Clearly seen in Figure 38, the overlapping portion between the two groups is quite large, indicating further improvements on the model is desirable.

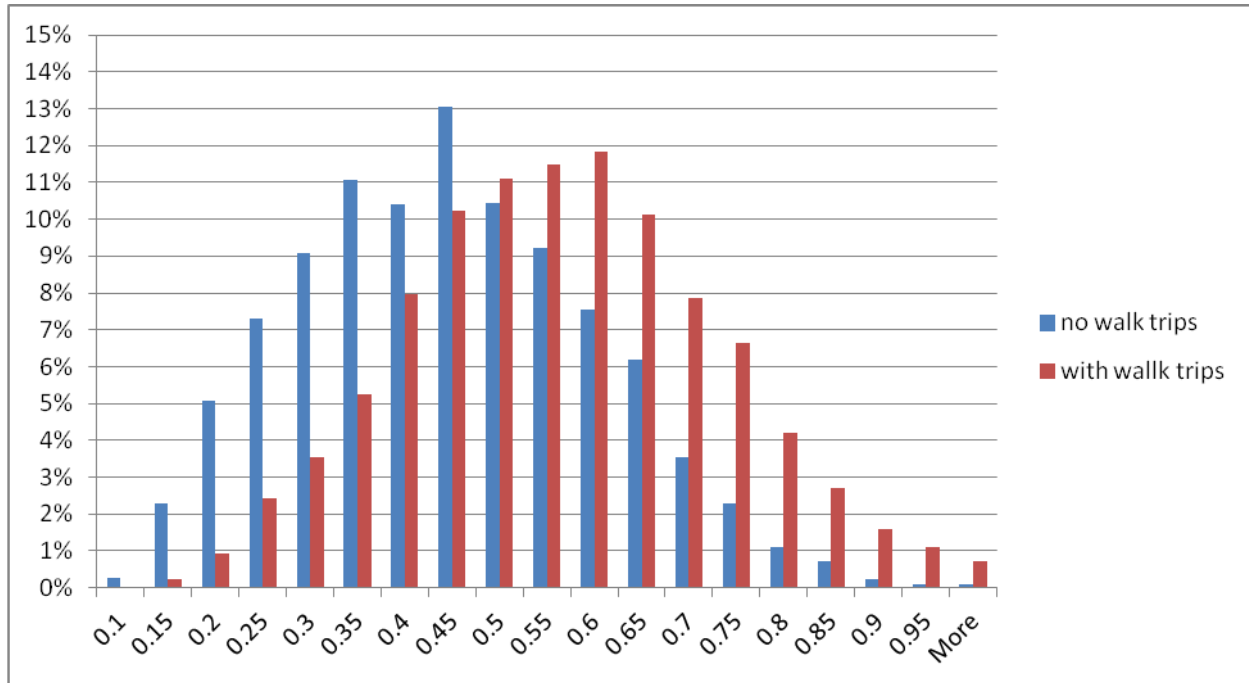


Figure 40. Distribution of the estimated probability on block groups with or without walk trips.

Similar reviews on other models were also conducted; i.e., for walk mileages, walk trips including transit access/egress, and bike trips. The distributions of the estimated probability for other walk models are very similar (Figure 39) to what is seen in Figure 38. On the other hand, the distribution patterns for the bike trip and bike mileage models are very different (Figure 40). Due to the “less common transportation mode” nature of biking, only small sets of samples were collected in NHTS. At the Census tract level (rather than block groups as for the walking side), only about 12% of these Census tracts recorded bike trips from NHTS sampled households. Although the “biking” distribution does show a slight shift toward the right (higher probabilities), the small set of biking tracts made it extremely difficult to draw conclusions from either of the preliminary bike trip and bike mileage models.

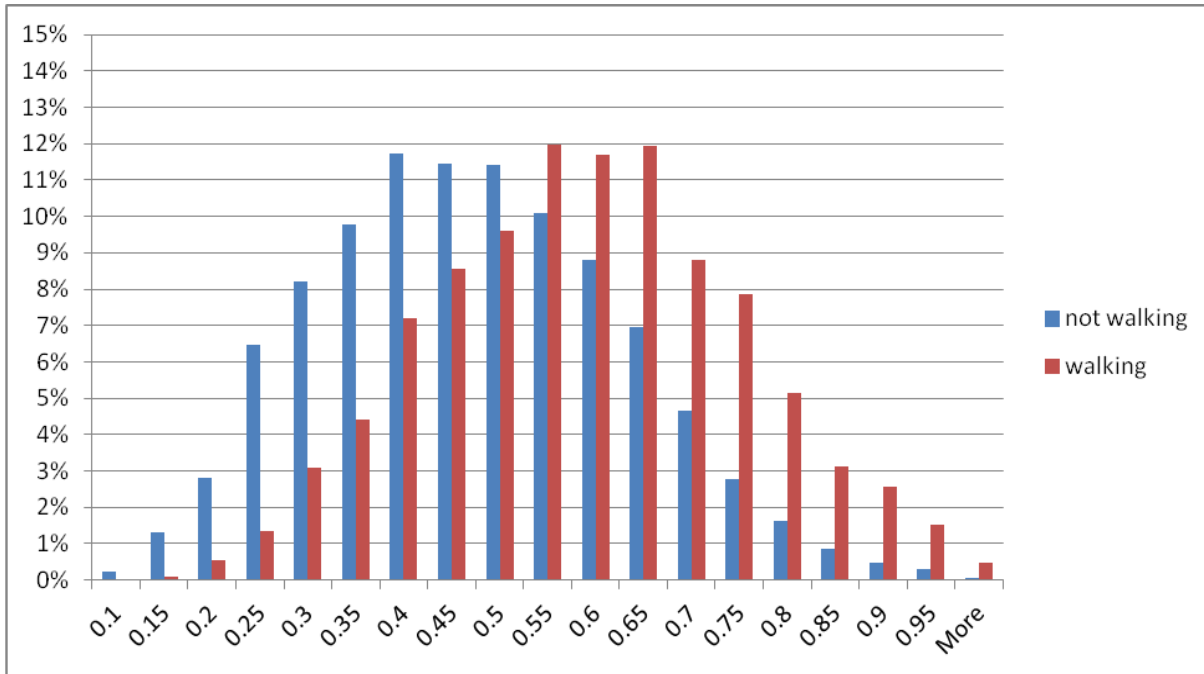


Figure 41. Distribution of the estimated probability on block groups with or without expanded walk trips (including transit access/egress trips).

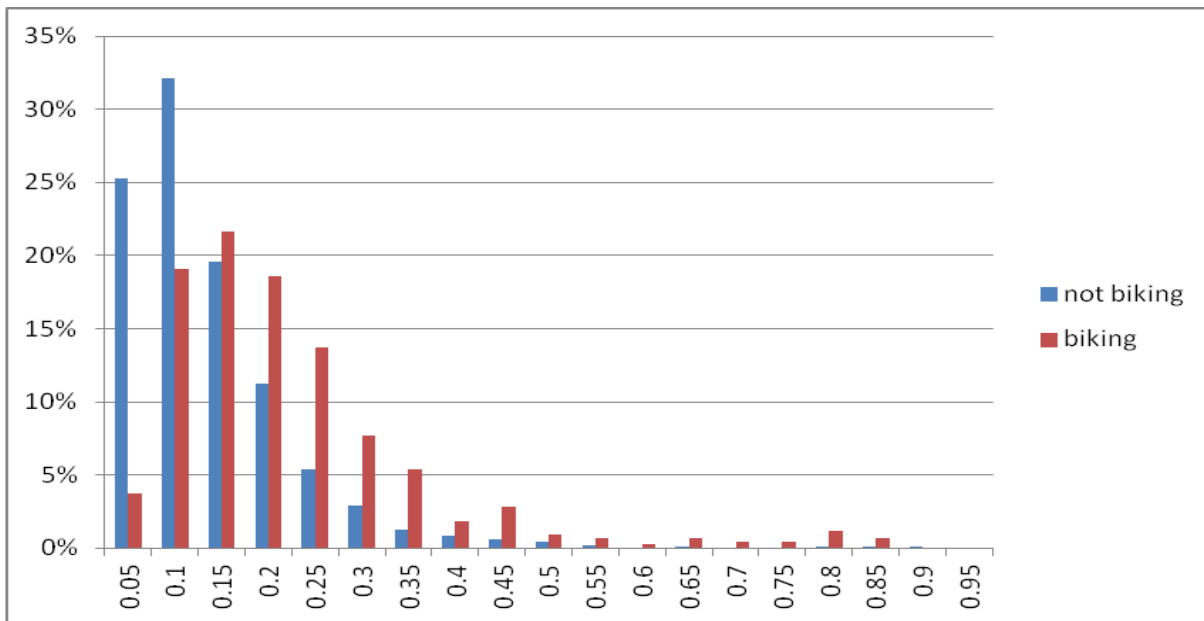


Figure 42. Distribution of the estimated probability for Census tracts with or without bike trips.

Since the Nielsen employment data was not received by the ORNL NMT team until a little more than a month prior to the end of this study, both the FHWA program manager and the team decided to focus the remaining effort on improving the walk models. Furthermore, based on preliminary results examined so far, there was no apparent advantage on using other measures (e.g., walk mileage) instead of walk trips in determining major factors for the purpose of

discriminating communities that walk/not walk. Thus, all efforts from that point on were limited only to the modeling of walk trips. As a result, all discussions starting from Section 4.5 in this report are based on walk trip models.

4.5 MODELING EFFORT WITH NIELSEN EMPLOYMENT DATA ON WALK TRIPS

The Nielsen employment data obtained by the research team include two variables, the employment number for all sectors and the employment number for the retail sector. As stated before, the data was provided for all block groups within the ORNL NMT project selected regions. These two Nielsen variables were added to the list of potential predictor variables used in the preliminary walk trip model, and new models were developed. Results from these new runs of stepwise linear logistic and discriminant analysis models are summarized here.

4.5.1 Results Based on the Linear Model with Final Selection of Predictors

As shown in Figure 41, the distributions of the estimated probability from this new model (i.e., including Nielsen employment variables) is quite similar to the one not including these variables (Figure 38). The two distributions, however, are better separated using this new model. This indicates an improvement of model performance when the employment data were added. To ensure the model selects most potential factors (i.e., predictor variables) for the discriminant analysis, and also because many of the redundant or irrelevant variables in the original data set were eliminated from previous modeling efforts, the significance level for the new runs was relaxed to the typical significance level of 0.5.

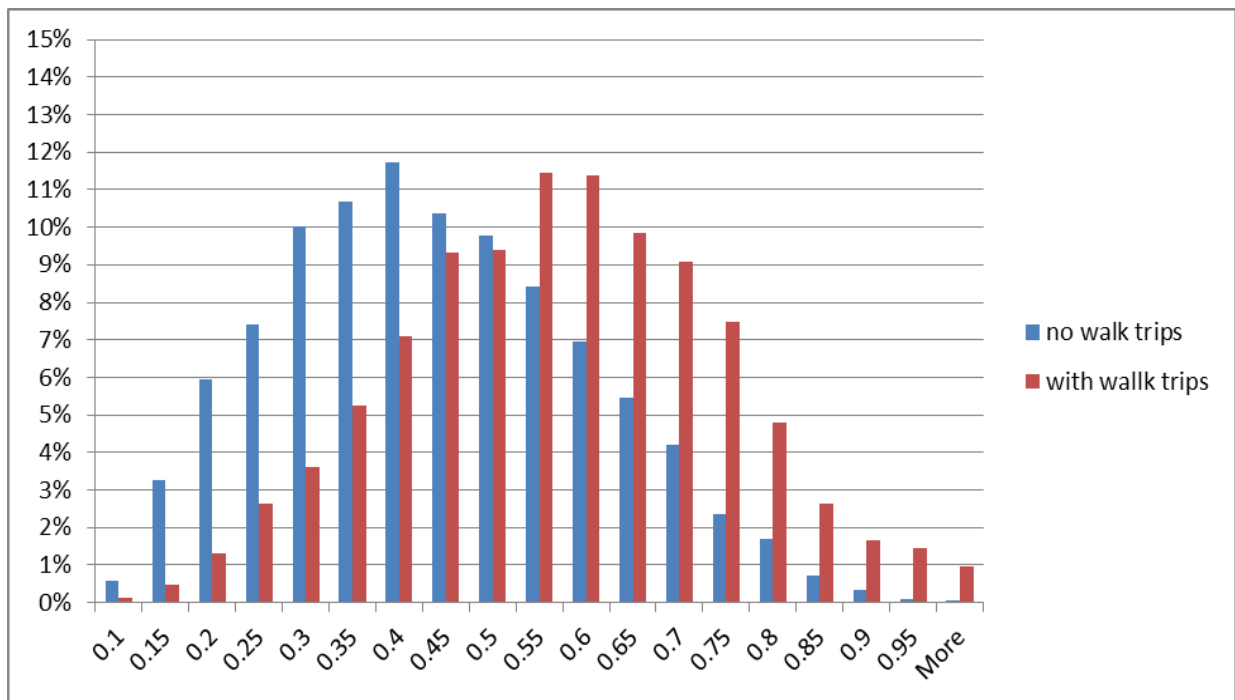


Figure 43. Distribution of the estimated probability based on linear model with Nielsen employment data.

4.5.2 Consideration of Dependency and Use of Non-Linear Predictor Variables

Previous models assumed linearity for all predictor variables, since those were mainly used for variable screening purposes. Now that factors were basically determined, additional steps were taken to investigate ways to improve the model performance, i.e., to separate the two groups (“with walk trips” and “without walk trips”) as far apart as possible. Transformation of variables, such as log or square root, is most commonly used when the relationship between a predictor variable and the dependent variable is not linear. For example, although retail density is known to associate with walking, the relationship is not linear (this can be easily done with a scatter plot). Instead of associating the *Retail Density* variable with the *walk/no walk* dependent variable, a log transfer or square root transfer of *Retail Density* might be considered.

Figure 42 shows the resulting probability estimates based on a model that includes Nielsen employment data and transformation terms of several variables. Although the distributions on the walk and no walk groups are still overlapping with each other, the two peaks appear to be “pushed” slightly further apart from the previous curves (Figure 41).

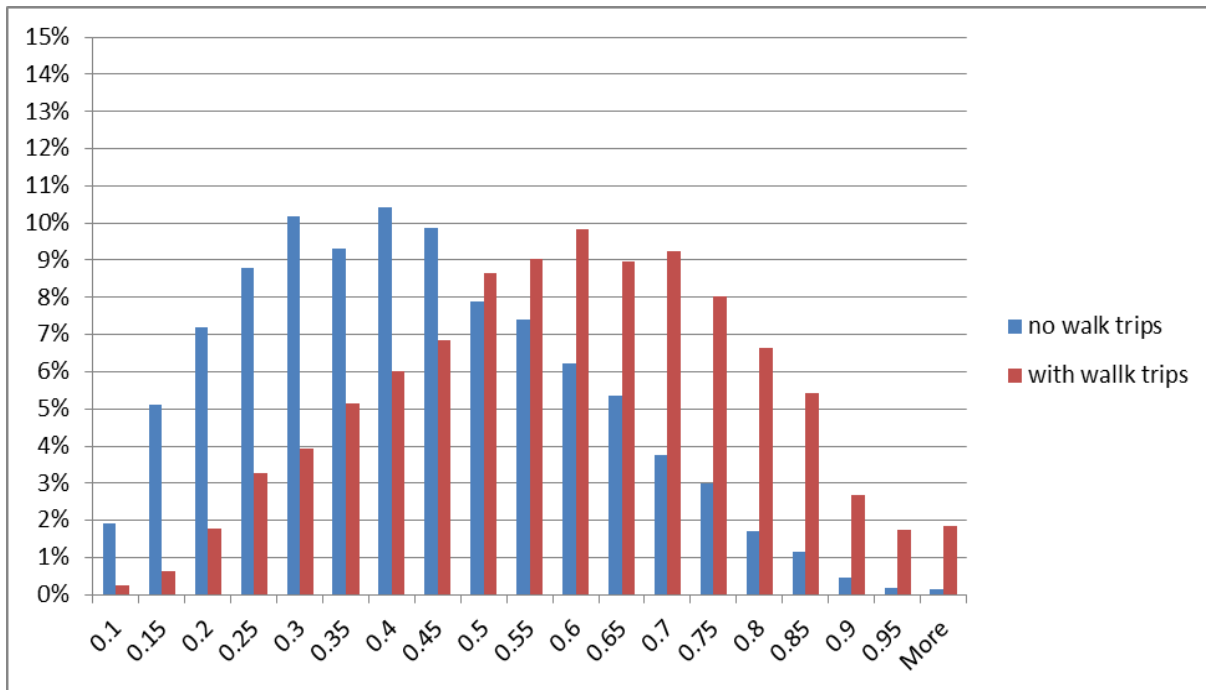


Figure 44. Distribution of the estimated probability based on model with Nielsen employment data and transformations of selected variables.

In addition to transformation terms in the model, interaction terms are also considered. Clearly many of the selected predictor variables are not independent from each other; for instance, percent of higher household income within a block group and the medium house value in that block group; or percent of non-white in the block group and percent of education level within the block group, just to name a few. Thus, the final modeling effort was to include interaction terms of predictor variables, along with some transformation terms. The probability estimates from the final model is displayed in Figure 43, and clearly, the peaks of the two distributions are further

apart than those from any of the previous models. Due to limitations in time and resources, the set of predictor variables selected by this model was determined to be the final one; and the list of factors are used to continue the analysis for discriminating the walk/non-walk block groups. These major factors are further discussed in Section 5 of this report. Because of the number of variables used, summary result tables on the predictor variables selected by this model are included in Appendices F & E, for models with and without interactions, respectively.

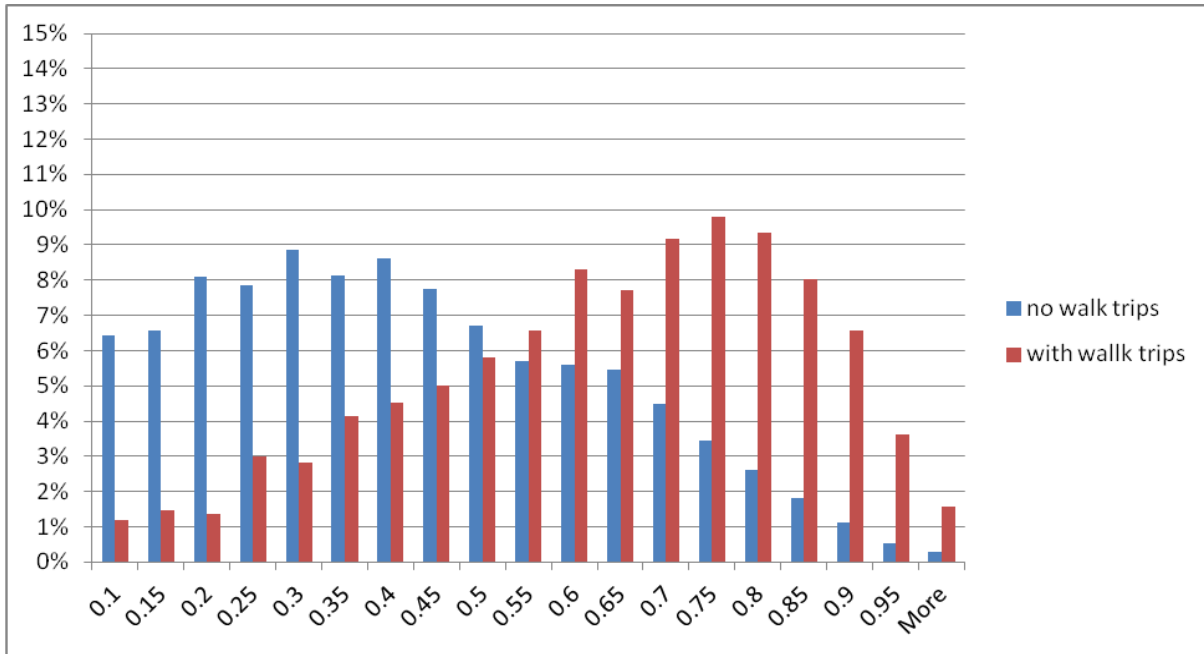


Figure 45. Distribution of estimated probability from model with transformation and interaction terms.

4.5.3 Review of NHTS Coverage Issue

Besides identifying factors that make a community walk or not walk, the ORNL NMT study also calls for a database that includes information on these factors for all areas under the study region (i.e., block groups). An index for measuring walk or not walk is also to be included in this database. Although there are nearly 14,000 block groups in the study region, only about half of the block groups have information on all of the predictor variables (thus were used in developing the models). Since variables based on data source such as ACS, Walk Score, or Nielsen Claritas are available for all block groups within the study region, only very few, if any, have missing values.

The NHTS sampling limitation is clearly visible in Figure 44, which shows the block groups within the Research Triangle region of the North Carolina. Block groups with white color, or “blank,” are areas without NHTS information, thus have missing values for NHTS variables. Since San Diego, California is a region that was oversampled under the 2009 NHTS, the region was used to illustrate the area with a better coverage under the 2009 NHTS. Clearly, one can observe only a few “blank” spots in Figure 45, which indicates San Diego coverage by the NHTS is much better and, therefore, will have less missing value in its NHTS variables. In order

to the build the desired database for all block groups within the study region, some imputations on missing NHTS values are necessary.

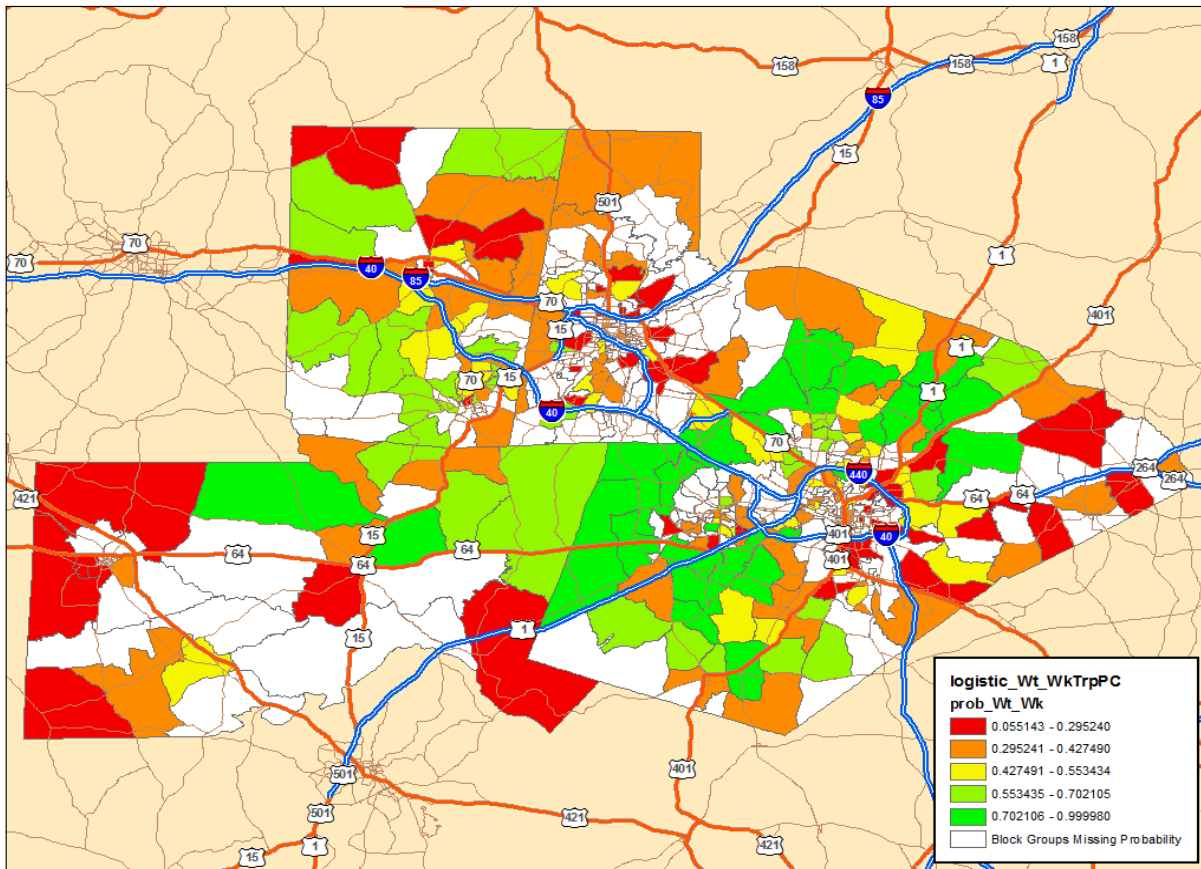


Figure 46. NHTS sample coverage for block groups within Research Triangle Region, NC.

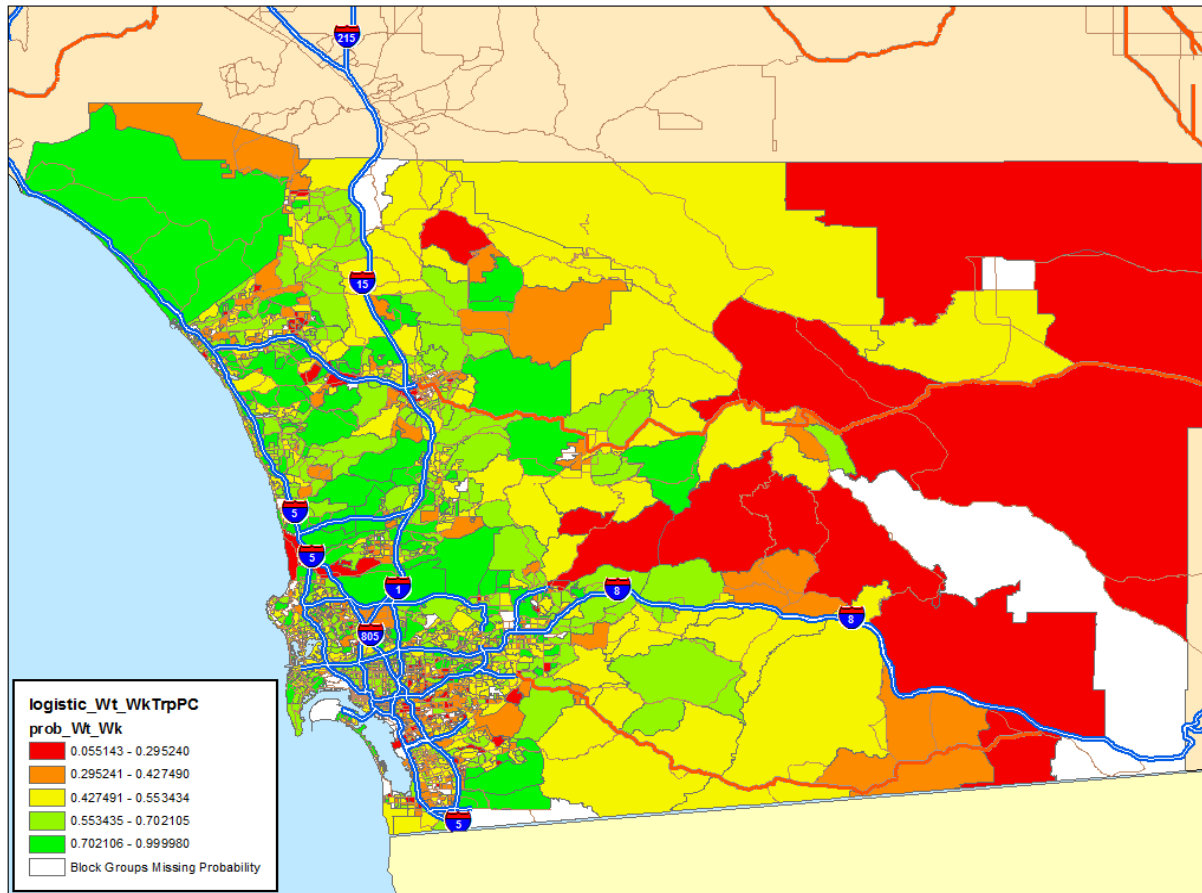


Figure 47. NHTS sample coverage for block groups within San Diego, CA.

4.5.4 Imputation of Missing NHTS Variables for Walk Trips

As pointed out above, the lack of NHTS sampled data in non-sampled block groups prevents the model from including those block groups to produce estimates. Consequently, the desired index for walk/not walk would not be available for those non-sampled block groups. Since only about half of the block groups in the study regions were included in the final model, this would mean the walk/not walk indices would not be available for the remaining half of the block groups. To reduce this impact, a series of imputation procedures was used to fill in missing NHTS values for each associate predictor variables.

Depending on which NHTS variables are involved, different approaches are used to obtain estimates from the imputation process. Specifically, two types of procedures are applied in handling the missing NHTS values.

Substituting Missing NHTS Variables with Equivalent Variables from the ACS

For NHTS variables that have equivalent variables from the ACS (e.g., total number of household in a block group), simple regressions were generated based on data from known block groups. The resulting regression equations are then used to produce estimates for each of the

missing NHTS variables on non-sampled block groups. The reason that missing NHTS variables can't be substituted using ACS equivalent data directly is due to the way sampling weights were calculated in the NHTS. Since NHTS data was not sampled for block group level analysis, NHTS block group level estimates are not consistent with ACS estimates on the same block group.

Imputing missing NHTS variables via geospatial method

For NHTS variables without existing equivalent ACS variables (e.g., average number of drivers in the household), a geospatial method was applied to impute necessary values for the non-sampled block groups. The method applies the NHTS values from a higher aggregated level of geography (i.e., Census tract) and uses them to fill the missing NHTS values for the block groups within that tract.

Under situations that a Census tract level value is also missing, the average value based on all known intersecting block groups around the targeted missing-block group can be calculated via geospatial software and used as an estimate for the missing block group. In some case, this process has to be conducted multiple times to ensure all block groups are covered. Figures 46 and 47 show different levels of imputation processes applied on two groups of NHTS variables.

For NHTS variables *Average Vehicle Counts* and *Percent of Drivers* (Figure 46), over half of the block groups in the NMT study region were covered by the NHTS sampling; leaving about 48% of block groups that needed to be imputed. Among those that require imputations, two-thirds of these block groups (i.e., 32% of the 48%) can be filled with their corresponding tract level values; and almost all the remaining block groups can be geospatially filled by the average values from their surrounding known block groups within the first pass.

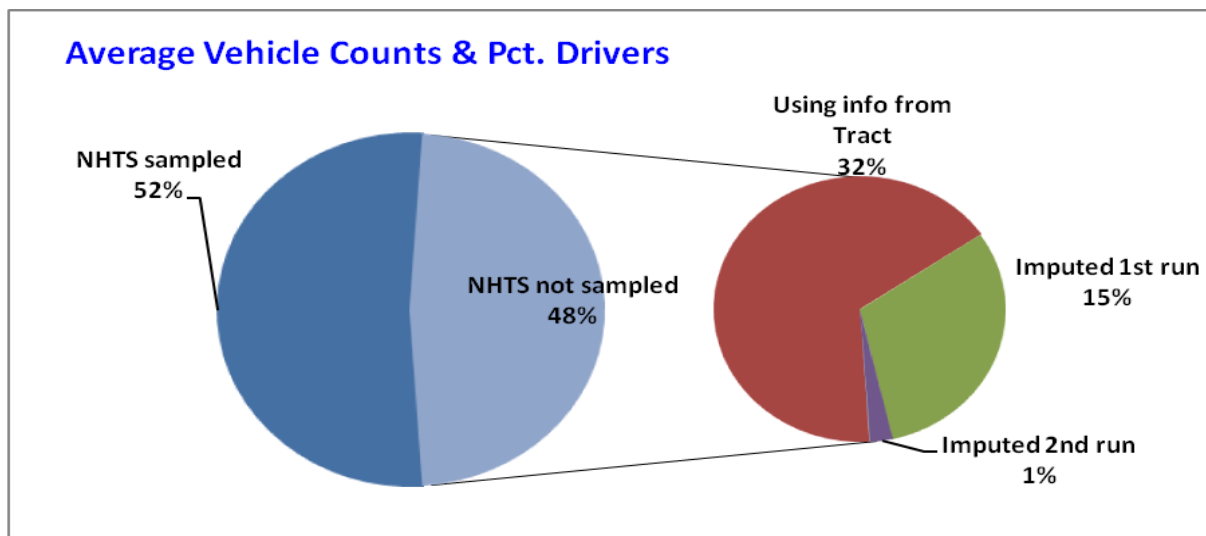


Figure 48. Imputation methods used for NHTS variables on average vehicle counts and percent drivers.

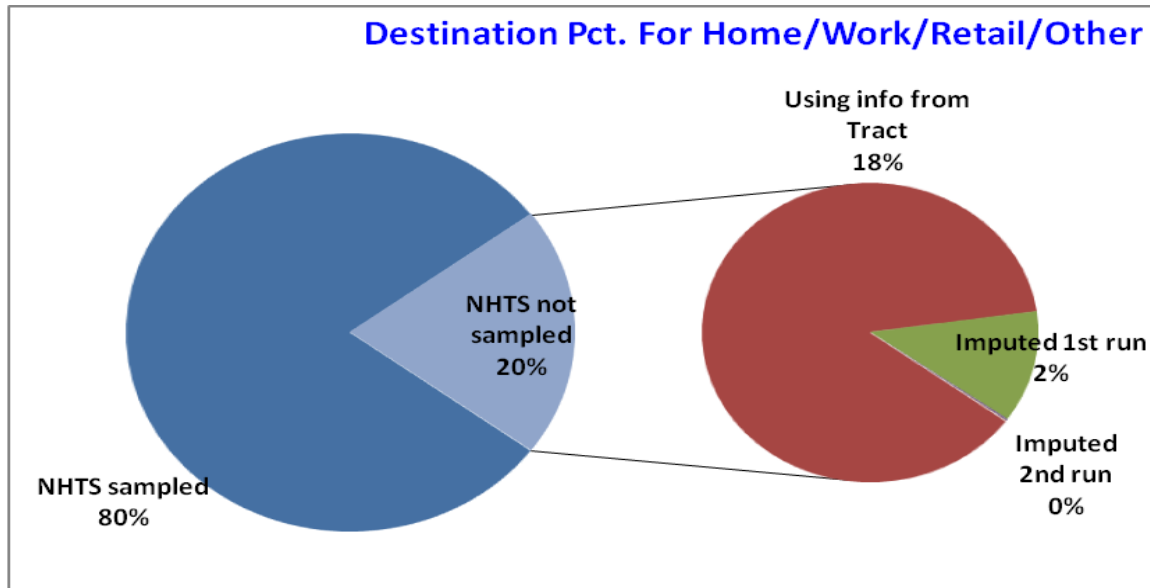


Figure 49. Imputation method used for NHTS variables related to share of trip purpose at destinations.

For NHTS variables related to the share of trip purposes to a destination, namely *Destination Percent for Home*, *Destination Percent for Work Purpose*, *Destination Percent for Retail Purpose*, and *Destination Percent for Other Purposes*, only about 20% of block groups within the study region have missing values. Among those NHTS non-sampled block groups, about 90% (i.e., 18% or the 20%) of the block groups can be imputed with their corresponding tract data. Similar to the previous case shown in Figure 46, the remaining block groups were geospatially filled by the average values from their surrounding known block groups within the first pass.

Note that block groups with imputed values were not used in developing the model. The values of predictor variables from those block groups are only used with the selected model formulation to generate the resulting walk/not-walk index for those non-sampled block groups in the study region. Thus, ranking of all block groups within the study region, in terms of walk/not-walk index can be conducted.

5. DISCRIMINATING FACTORS

A large number of tables, models, and selected predictors were generated from the data analysis and modeling effort discussed in Section 4. The final set of predictor variables, based on the model described in Section 4.5.2, were selected as major factors that can be used to discriminant walk/not-walk in block groups. These major factors were grouped in a few general categories for ease of discussion and are briefly summarized in this section. Note that interaction effects may cause the effect of the standalone variable to be nonsensical.

5.1 HOUSEHOLD DEMOGRAPHICS

5.1.1 Drivers

Percent of Drivers

(NHTS variable) was selected as an overall negatively correlated predictor in the logistic regression model, once interactions are taken into account.

Average Drivers per Household

(NHTS variable) was selected as a negatively correlated predictor in the logistic regression model.

5.1.2 Household/Person Counts

Number of Households in BG/CT

(NHTS variable) was selected as a positively correlated predictor in the logistic regression model, with interactions having both positive and negative correlations.

Population in BG

(ACS variable) The log of ACS population was selected in the stepwise discriminant model, and was thus added to the final discriminant model.

5.1.3 Income

CY Own Occ HUs Median Value

(Claritas variable)⁴⁰, which is presumably a measure of affluence, was selected as a correlated predictor in all of the pedestrian models.

Household Income in the Past 12 Months

The candidate predictor for *High Income* (ACS variable) was selected as a positively correlated predictor in the logistic models for walk trips per person. Interaction terms also were almost exclusively positive.

⁴⁰ “CY” is used in many of the Claritas variable labels to indicate data for the corresponding calendar year.

5.1.4 Education

Percent Persons in BG/CT with Education of Some College/Bachelor's Degree

(ACS & NHTS variables) was selected as a positively correlated predictor in all of the walking logistic models. Higher education equates to a higher likelihood of walking.

5.1.5 Language and Race

Persons by Race, Black or African-American

(ACS variable) was selected as a net negatively correlated predictor in the logistic walking models once interaction terms are taken into account.

Claritas Percent Race Variables

(Claritas variables) Percent of the population that was Hispanic/Latino was found to be negatively correlated with walking, while percent white and percent "Some Other Race" were found to be positively correlated.

Race of NHTS Household Respondent

(NHTS variables) Mirroring the Claritas variables, the percent of household respondents who were Hispanic was negatively correlated with walking in the logistic models, while the percent white was positively correlated.

Language of NHTS Interview

(NHTS variables) The percent of households in a block group with interviews in English was negatively correlated with walking. This almost certainly is a counteraction of the positive correlation in prior race variables.

5.1.6 Vehicle Ownership

Average Vehicle Count

(NHTS variable) was selected in all of the pedestrian models as a negatively correlated predictor (the more vehicles, the less walking), although interaction terms involving vehicle count are mixed, with positive and negative correlations.

5.1.7 Gender

Females Aged 22 to 34

(ACS variable) was selected as a predictor in the stepwise discriminant model for walk trips per person.

5.2 EMPLOYMENT

5.2.1 Worker and Job Type

Female State Government Workers

(ACS variable) was selected as a negatively correlated predictor in the walking model.

Employees per Square Mile

(Claritas variable) was selected as a positively correlated predictor in the walking model. Note that this variable represents employees who *live* in the block group, not who have a workplace in the block group.

5.2.2 ACS Journey to Work

Workers by Means Of Transportation To Work

(ACS variable) Workers who worked at home had a positively correlated effect on walk trips.

Workers by Time Leaving Home To Work

(ACS variable) Workers who left home early had a positively correlated effect on walk trips, while workers who left later had a negative correlation.

Workers by Length of Time Traveling To Work

(ACS variable) Workers who traveled fewer than 10 minutes to work by taxi, motorcycle, bike, or walk had a lower level of walk trips.

5.2.3 Nielsen Employment Data

Note that employment data in this section is for the block group where the employee works, not where the employee resides. A large number of interaction terms among variables in this section led to some indiscernible correlations.

Total Worker Measures

(Nielsen variables) Total number of workers had a positively correlated effect on walk trips, with interactions having some smaller, but negative correlations. Overall worker density was similar, with an overall positive correlation and smaller, negative correlations in interaction terms. The number of workers as a percent of the total population was found to have a negative correlation with walk trips.

Retail Worker Measures

(Nielsen variables) The density of retail workers was found to have an overall positive correlation with walk trips, while the log transformation of the same term had the opposite effect. The number of retail workers as a percent of the total population was found to have a mixed correlation once interaction terms were taken into account. Finally, the percent of employees working in retail in a given block group was negatively correlated with walk trips, once a square root transformation was applied.

5.3 ACTIVITY MEASURES

5.3.1 CDC Obesity and Inactivity

Gradation of percentage of physically inactive adults

The *County-level Gradation of percentage of physically inactive adults by total number of physically inactive adults in 2008* (CDC data) was selected as a positively correlated

predictor in the models for walk trips per person, but with interaction effects the correlation is mixed.

5.3.2 Walkability and Transit Accessibility Measures

Walk Scores

The candidate predictor *Walk Score* was selected as a positively correlated predictor for all of the models of pedestrian trips and mileage in both the classification and quantification model stages.

5.4 DESTINATION OF TRAVEL

Percent Weighted Trips in Destination Block Group for the Purpose Home

(NHTS variable) was selected as a positively correlated predictor in all of the classification-stage models for walking.

Percent Weighted Trips in Destination Block Group for the Purpose Work

(NHTS variable) was also selected as a negatively correlated predictor in all of the classification-stage models for walking.

Percent Weighted Trips in Destination Block Group for the Purpose Other

(NHTS variable) was selected as a positively correlated predictor in all of the classification-stage models for walking, with interaction terms having a negative correlation. This “Other” category includes trips not to Home, Work, Retail, or School.

5.5 OTHER FACTORS

Number of Land Uses

The Shannon entropy, our other measure of diversity, excluding “Residential” was selected in the stepwise discriminant model, and was thus added to the final discriminant model.

Urban Indicator

(Claritas variable), was also selected in the stepwise discriminant model for walk trips per person, and was thus included to the final discriminant model.

6. EVALUATION OF THE DISCRIMINANT MODEL AND RESULTING DATABASE

Subjects discussed under this Section include, how to use the discriminant function resulting from the modeling efforts (described in Section 4 of this report) for classification purposes (e.g., assign individual block group into walk or not walk) and how to evaluate its performance in this regard.

6.1 CLASSIFICATION USING THE DISCRIMINANT FUNCTION

The set of predictor variables, selected from previously discussed modeling efforts, was used to generate the final discriminant model formulation for the walk trips. The “output” from this model, i.e., the dependent variable, is a probability measure with values between 0 and 1. For convenience, results from the discriminant function were normalized to values between 0 and 100. These values can be viewed as scores or indices; where a value of 100 is the highest, which means the highest likelihood of walking in a community (i.e., a block group).

Thus, each block group in the study region is associated with a score/index calculated based on values from individual’s predictor variables. In order to use these scores/indices for allocating individuals to one of the two groups (i.e., walk or not-walk), one would need to specify a critical score, or cutoff point, such that a block group is assigned to the walk group if its score exceeds this cutoff point, and to the not-walk group if it does not. Typically, if information from past experience on other similar data is available (i.e., a prior), the cutoff can be set based on that knowledge. Or, if an individual can be randomly selected from the population, the ratio of the two probabilities can be used to estimate the cutoff point. Otherwise, if the relative size of the two groups is about the same, a mid-point such as 0.5 could be a reasonable estimate for the cutoff point (similar to nearly symmetrical shape observed in Figures 42 and 43).

Under certain circumstances, important considerations may warrant a modification of the cutoff point to something other than the symmetrical one, however. This is particularly true when cost or seriousness of an incorrect assignment is group-dependent. For example, in health related studies where the error of assigning an individual with certain disease as not having that disease on the basis of diagnostic tests (i.e., false negative) would be much more serious than to treat a healthy individual unnecessarily (i.e., false positive).

6.2 EVALUATION OF PERFORMANCE

The assignment procedure determined by the use of a discriminant function, as described above, is a statistical decision making process. Thus, it is important to have a method to measure its performance in terms of the accuracy of assignments. In order to perform such a model evaluation, two similar data sets are typically involved; one is generally referred to as the training set and is used to formulate the discriminant function; the other is called the testing set and is used to evaluate the classification results.

6.2.1 Selections of Training and Testing Data Sets

Because the study dataset, which included about 7,100 block groups with useable data in all predictor variables and the dependent variable (i.e., walk trip per person), is the only data set available for the analysis, this study set was split into two parts for the model evaluation purpose. Such a split sample method, unfortunately, has an effect on reducing the usable sample size for modeling. To reduce this impact and allow sufficient data for an accurate model formulation, the NMT team decided on a 75-25 split on the data records. That is, data from 75% of the block groups are to be applied as the training set for generating the discriminant function; and the remaining block groups' records are to be used as the testing set for model evaluation.

The block groups that will form the testing data set were randomly selected, but with the consideration of being closely resemble the training set data in terms of several characteristics, including population density, employment density, median income within the block group, and their walk trip propensity. In addition, geographic region coverage for the selected block groups was also considered in the selection. The selection process resulted in a total of 1,662 block group records in the testing set; and the remaining over 5,400 block group records were used as the training set for developing the discriminant function.

For a more robust error estimate, the training-testing split process described in Section 6 could be conducted multiple times; i.e., applying different sets of randomly selected records each time to obtain multiple sets of error rates (misclassification rates). Typically, these can be performed using the approach commonly known as Monte Carlo simulation method. The mean (or average) of the resulting error rates could be used as a more unbiased error rate estimate from the classification model. Examination procedure beyond the single training-testing set procedure unfortunately was not conducted in this project mainly due to resource limitation.

6.2.2 Review of Error Rates

As pointed out before, the numbers of block groups with and without walk trips, based on the NHTS samples, are about the same. A cutoff point of 0.5 would be a good approximation. Furthermore, there is no prior knowledge available in terms of the probabilities of walk or not-walk in block groups. Thus, assuming an equal likelihood of walk and not-walk in the block group seems reasonable, which also leads to the cutoff point of 0.5.

Using the selected predictor variables as determined from the modeling effort, a discriminant function was generated based on the training data set. This formulation was then applied using the testing data set to generate their predicted classifications (or assignments). These predicted classifications were compared to the actual classifications (i.e., observations based on NHTS data for those block groups) to estimate misclassification rates, or error rates. Based on the test data selected, as Figure 48 shows, the false positive rate of this classification model is about 31% and the false negative rate is about 30%.

In statistical textbooks and literatures, a “Mahalanobis’ D” which represents a generalized measure of the distance between the two targeted groups (or populations) is commonly provided by statistical program software for the discriminant analysis. The quantity D^2 can be used to estimate misclassification rates. Based on the discriminant function generated from this study,

the D^2 value is 1.027; which resulted in an estimate of 0.289 for the error rate (i.e., ~29%)⁴¹. Error rates from the test set “observation” and the Mahalanobis’s statistics are consistent; both estimated the misclassification rate for this model at approximately 30%. This suggested the classification model (i.e., the discriminant function) developed based on data from the 2009 NHTS does perform reasonably well, given certain sampling limitations in the data for non-motorized travel.

<i>Observed*</i>	<i>Predicted</i>		
	<i>not walk</i>	<i>walk</i>	
<i>not walk</i>	595	264	<div style="border: 1px solid black; width: 100px; height: 15px; background-color: #d9ead3; margin-bottom: 5px;"></div> false positive
	69.27%	30.73%	
<i>walk</i>	243	560	<div style="border: 1px solid black; width: 100px; height: 15px; background-color: #fce4d6; margin-bottom: 5px;"></div> false negative
	30.26%	69.74%	

Figure 50. Classification table using the discriminant function developed from the NMT study and applied with data from test set of block groups (*observation based on NHTS data).

6.3 RESULTING SCORES/INDICES

As mentioned previously, the scores/indices estimated from the discriminant function were normalized to values between 0 and 100, with 100 being the highest likelihood of walking in a block group. Maps shown in Figures 50- 58 display the estimated walk indices in each of the 9 study regions, based on the model developed under this study. Thus, they are based on the same “national” scale, which makes comparisons between regions possible. The white borders shown in each of the maps are the block group boundaries. Walk indices are categorized into 5 categories, as defined in Table 8. Note that walk indices for a small number of block groups, about 60 of the total almost 14,000 block groups, could not be estimated. Some of these block groups are regions where no walking could be done (e.g., lakes) or areas with too much missing data in predictor variables that the indices were not estimated. Those were given an index of 0.

Based on the walk indices estimated from the discriminant function, a list of block groups with the highest 100 walk indices is provided in Appendix H; with the list of block groups with the lowest 100 walk indices follows in the same appendix. Using the categories defined in Table 8, walk index distributions of block groups in the study region as a whole i.e., (national) and each of the nine individual regions are shown in Figure 49.

⁴¹ Estimated error rate based on standard normal cumulative probabilities of the value from $-(\text{square root } (D^2))/2$.

Table 8. Definition of Walk Indices Category

Block Group Walk Indices	Category
10 & under	Very low likelihood walking
10.01-30	Low likelihood of walking
30.01-60	Moderate likelihood of walking
60.01-80	Considerable likelihood of walking
Over 80	High likelihood of walking

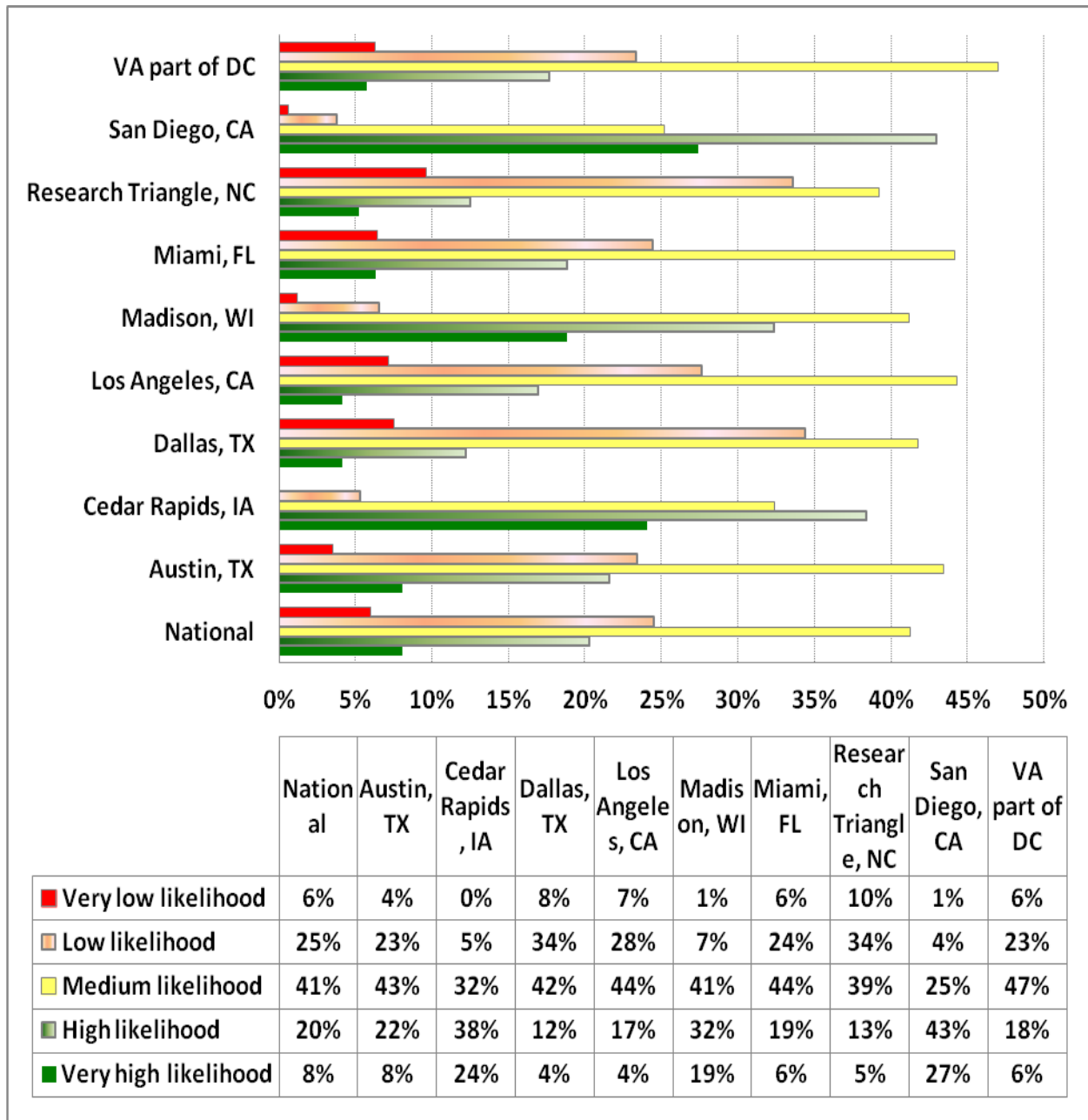


Figure 51. Distribution of block group walk index categories by each of the nine study regions, and with the combined study region as a whole (i.e., national)

At the national level, about 6% of block groups are classified as “very low likelihood” of walking. On a region by region base, this percentage varies, ranging from 0% in Cedar Rapid, IA to about 10% in Research Triangle, NC. Less than 1% of block groups in the San Diego region is classified in the “very low likelihood” category, while 7% of block groups in LA is in the same category. The two smallest regions (in terms of population, see Table 7 in Section 4), Cedar Rapid, IA and Madison, WI, along with San Diego (a NHTS over sampled area), all have larger percents of block groups within their regions that are categorized as “very high likelihood” of walking. Percentagewise, the Research Triangle region of the NC has the highest share of block group in “very low likelihood” of walking category. Part of this might be due to NHTS sampling limitations, however.

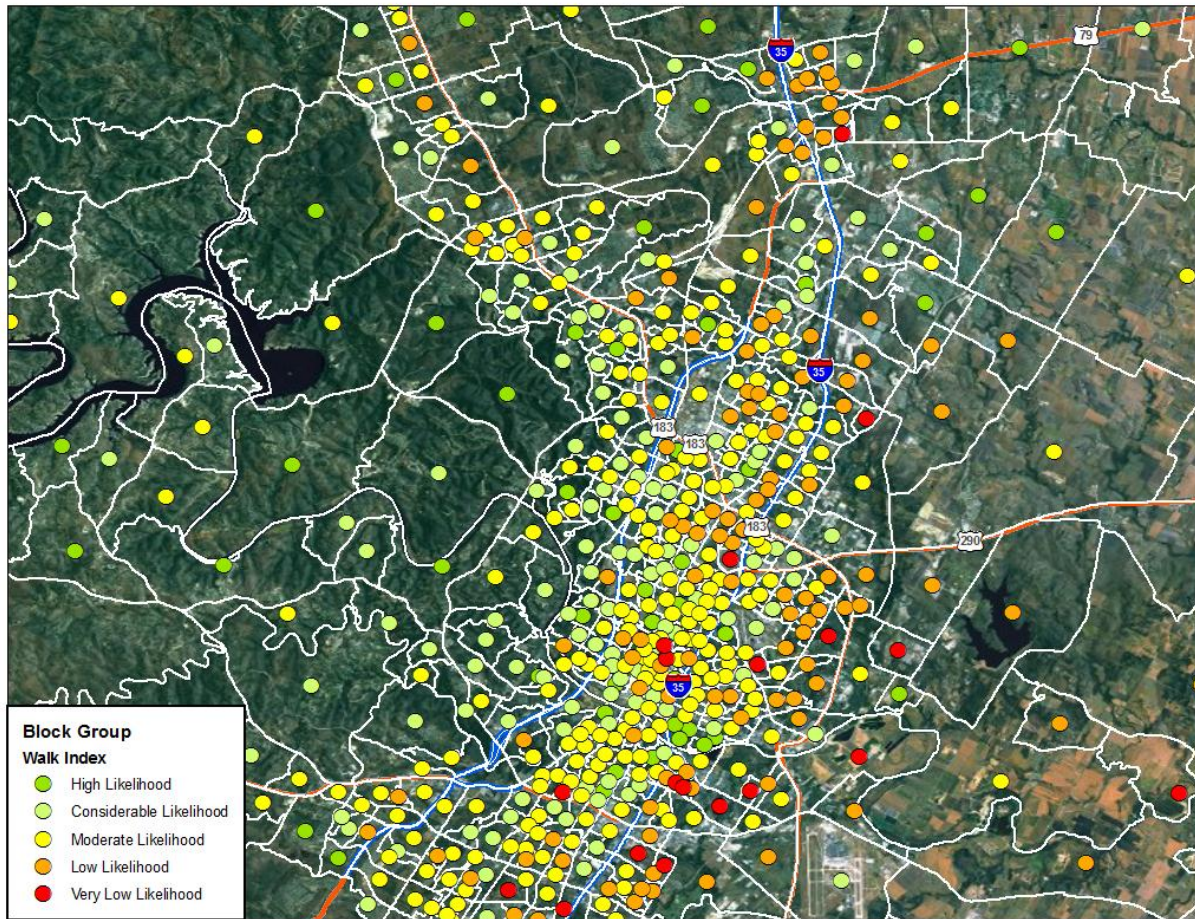


Figure 52. Estimated walk indices for block groups in Austin, TX.

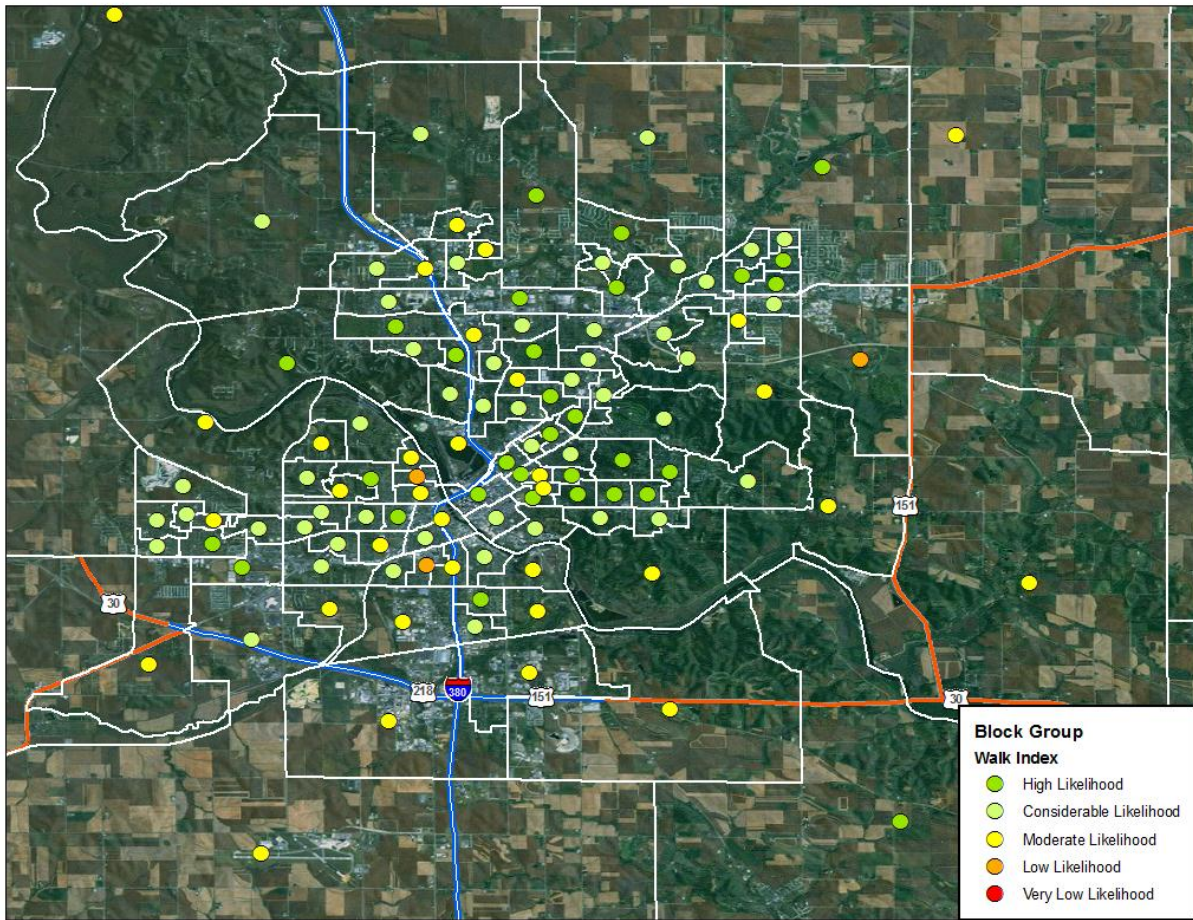


Figure 53. Estimated walk indices for block groups in Cedar Rapids, IA.

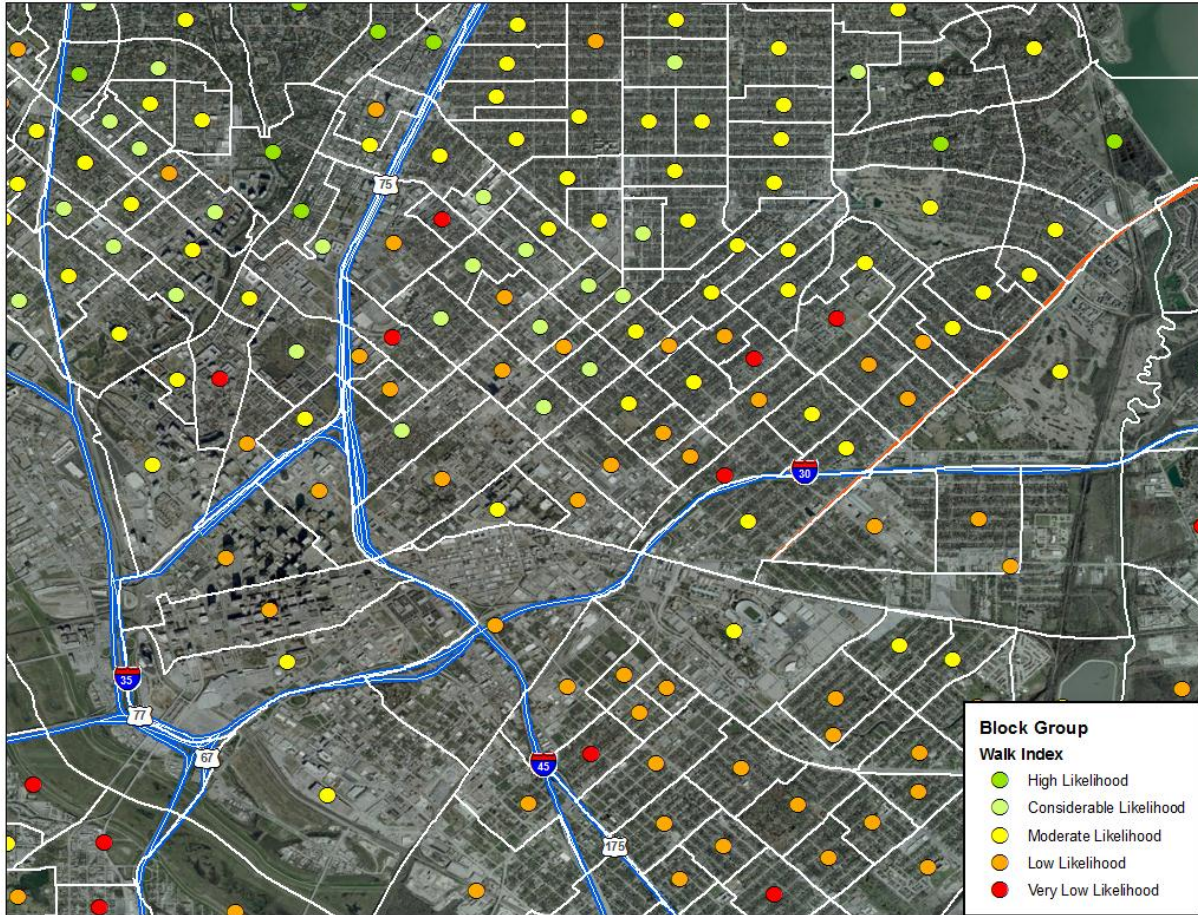


Figure 54. Estimated walk indices for block groups in Dallas, TX.

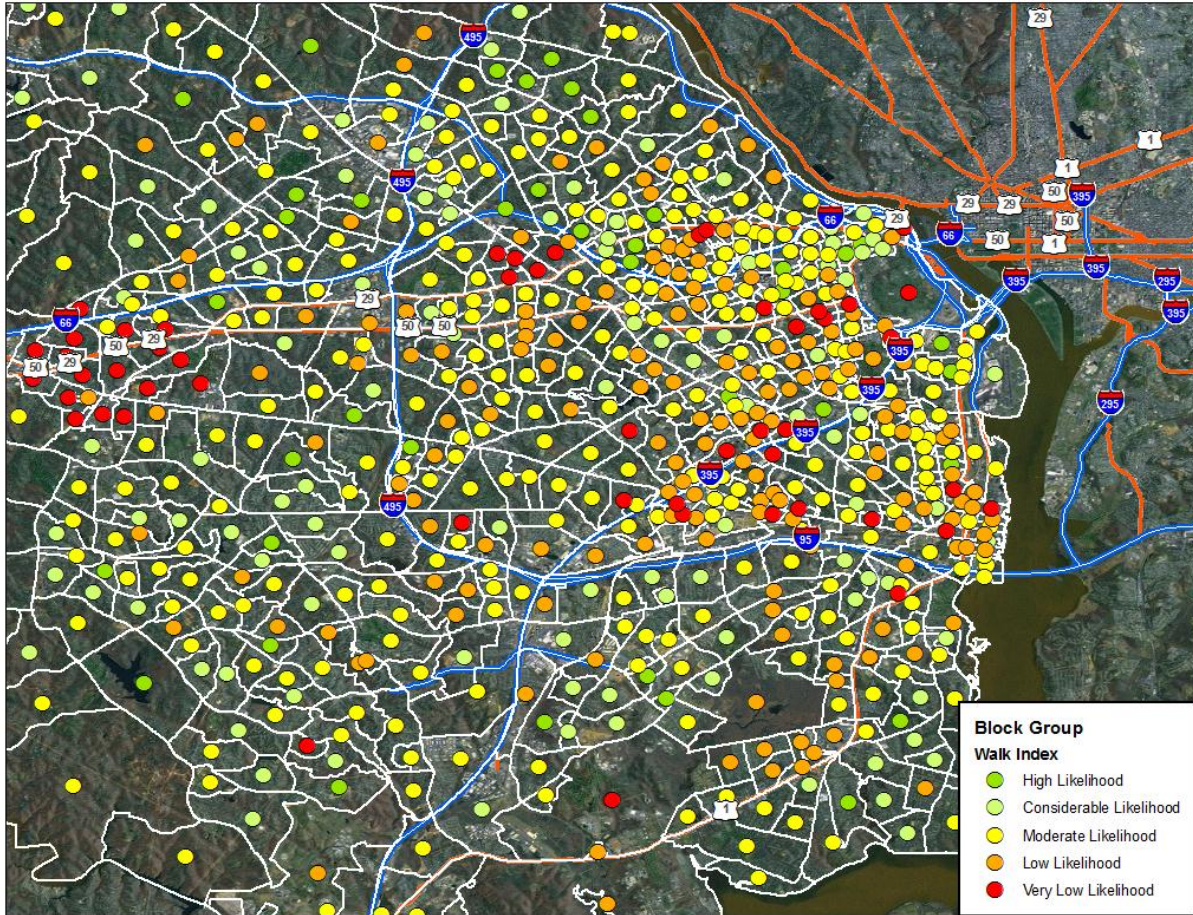


Figure 55. Estimated walk indices for block groups in Virginia portion of Washington, DC.

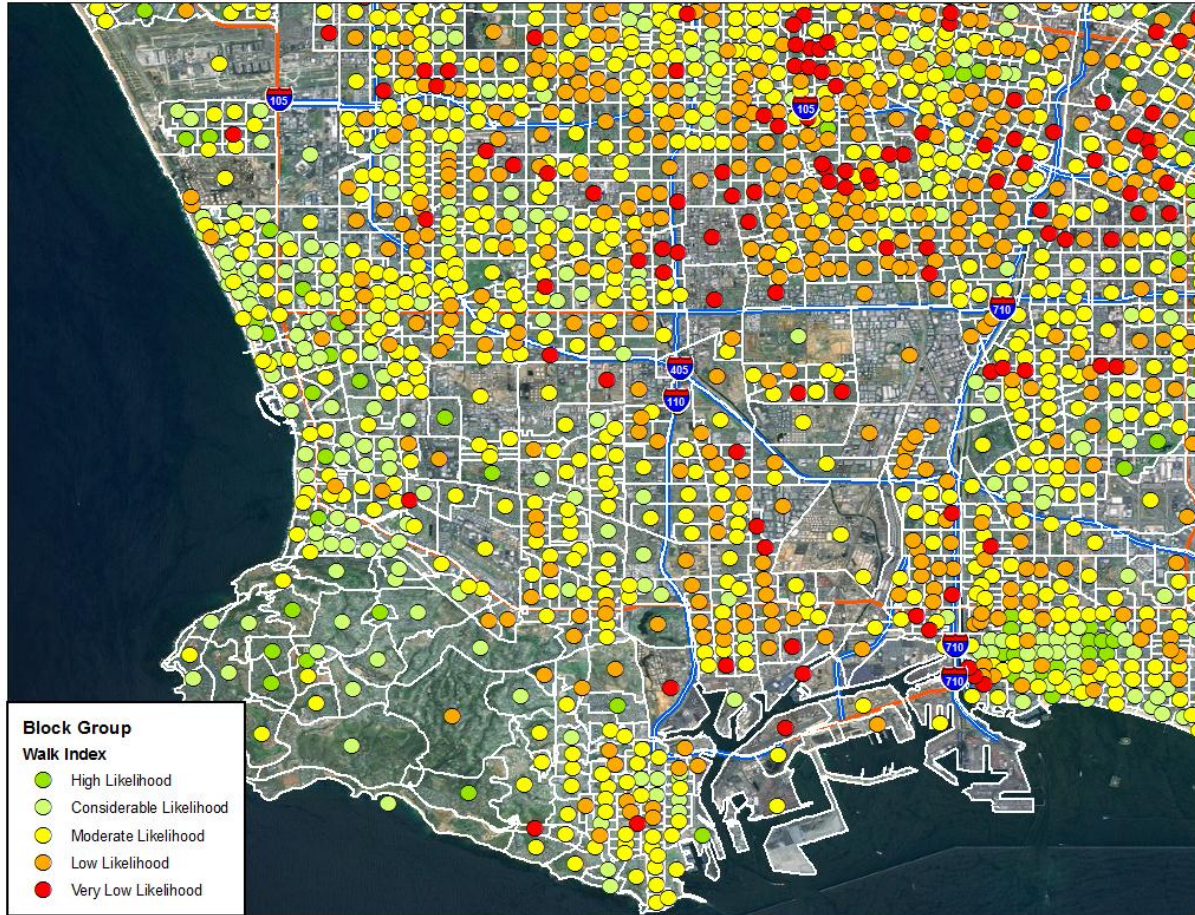


Figure 56. Estimated walk indices for block groups in Los Angeles, CA.

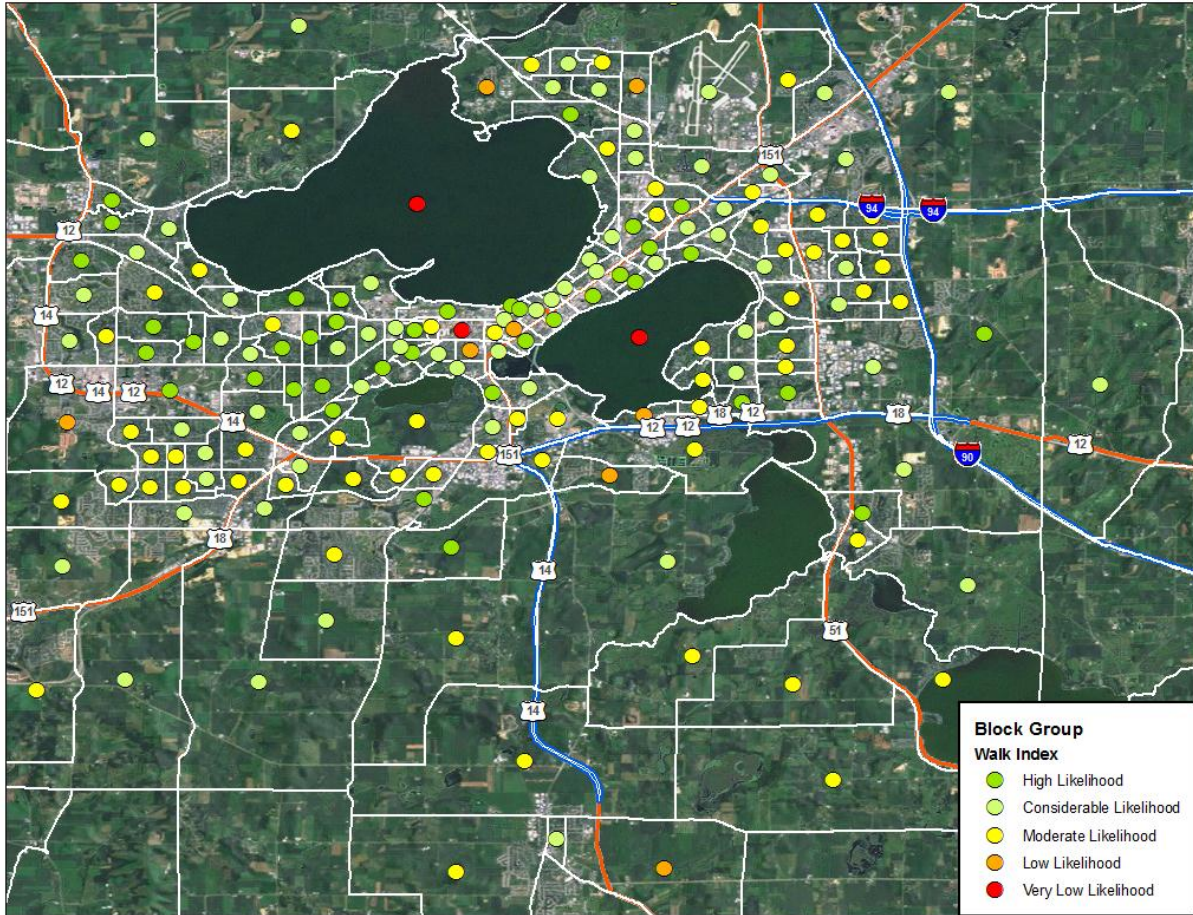


Figure 57. Estimated walk indices for block group in Madison, WI.

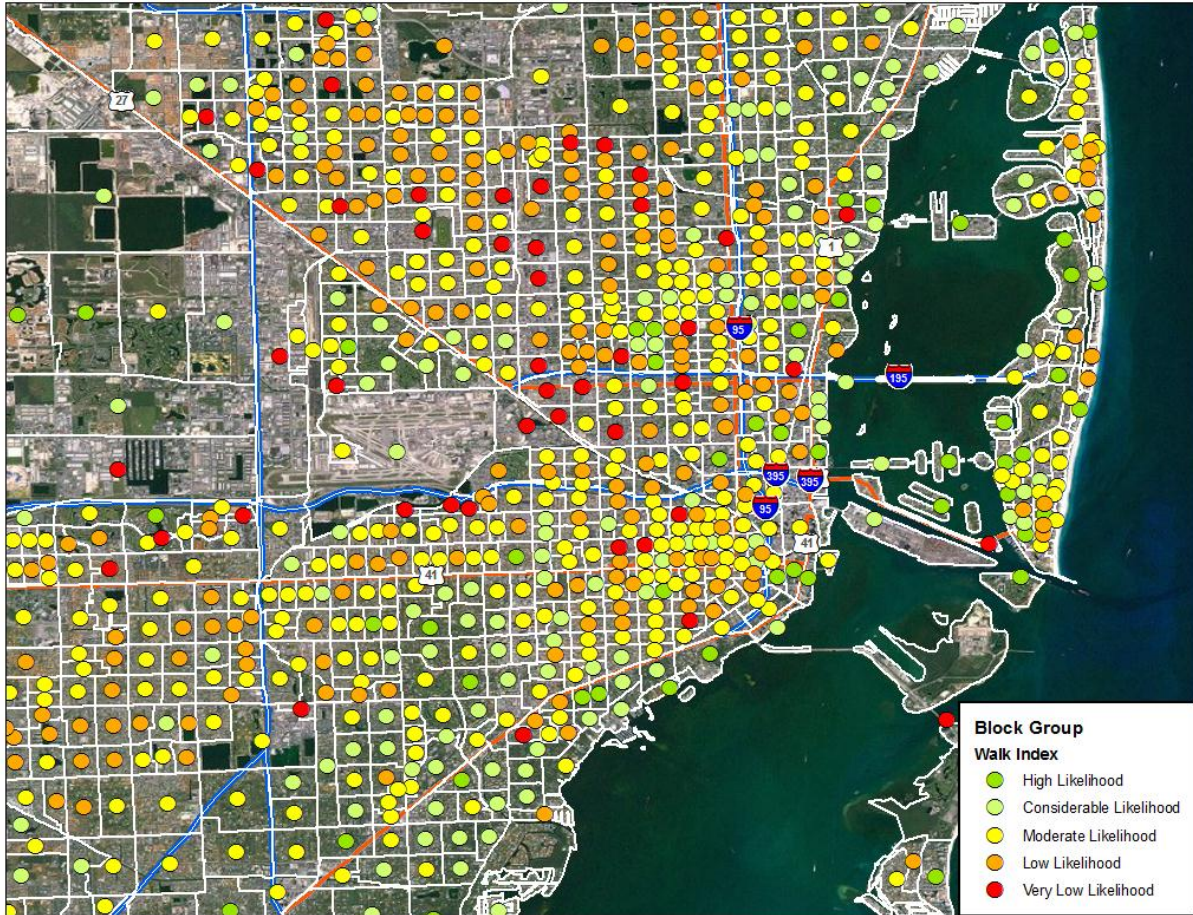


Figure 58. Estimated walk indices for block groups in Miami, FL.

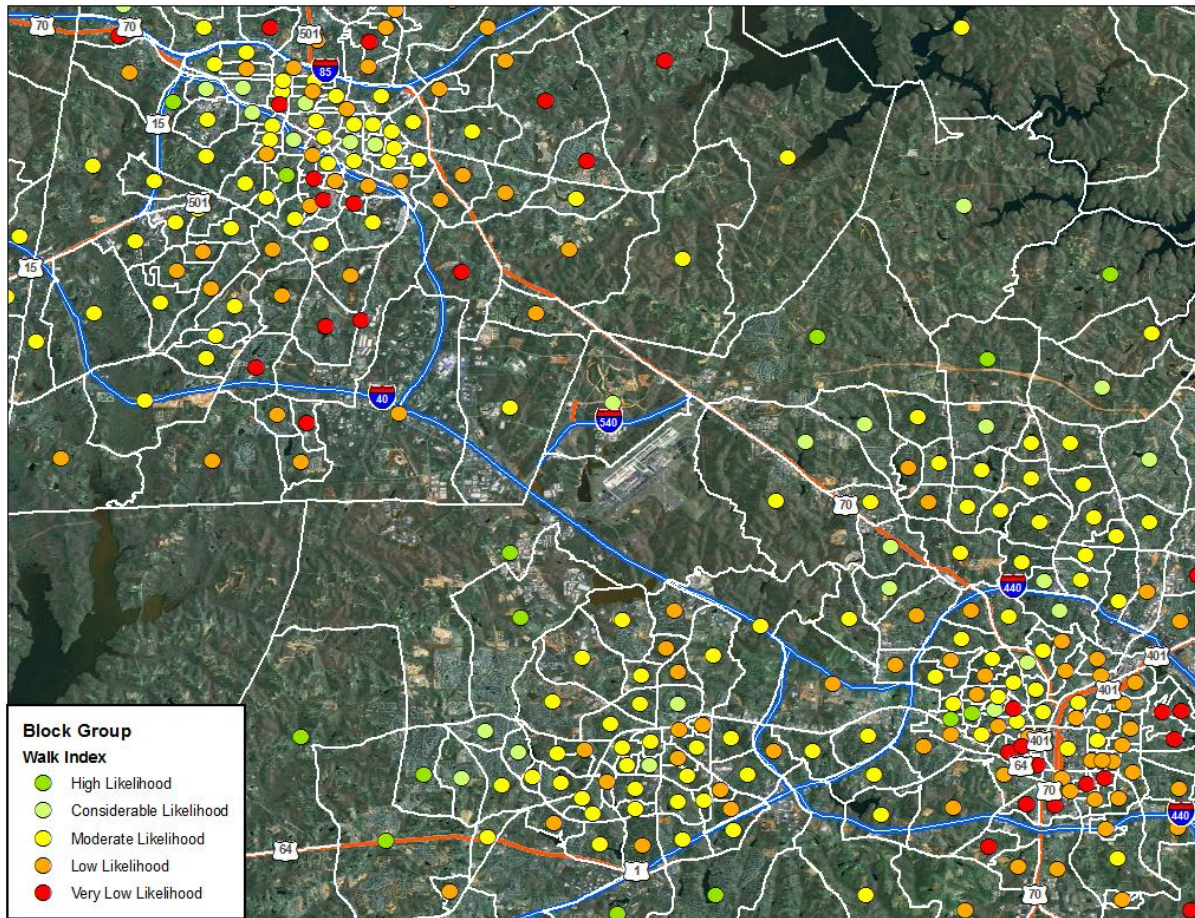


Figure 59. Estimated walk indices for block groups in Research Triangle, NC.

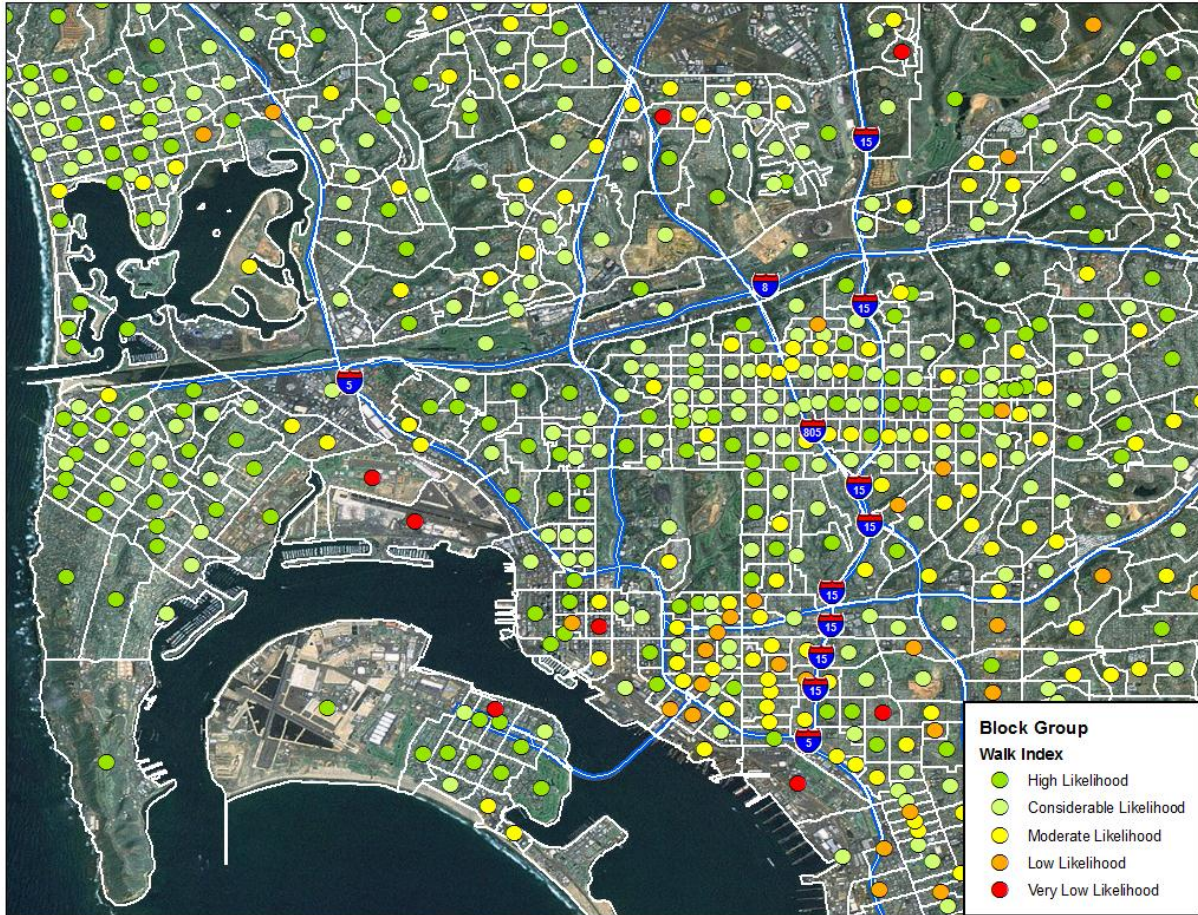


Figure 60. Estimated walk indices for block groups in San Diego, CA.

7. SUMMARY AND CONCLUSIONS

7.1 OVERVIEW

The goals of this NMT study were to identify major factors that influence communities to walk and bike; and to understand why, or why not, travelers walk and bike in their communities. This study relies on the 2009 NHTS data as the major source of data, and was supplemented with data from ACS, educational survey, health, employment, and others. All together, these data sets provided over 400 potential predictor variables for the study. Initial statistical screening methods were applied, and examined with various measures (e.g., walk trip per person, walk mileage per person, bike trip per person, bike mileage per person) as the dependent variables. The best geographic level of detail used in the modeling for this study was determined to be the block group level for walking and Census tract level for biking.

The need for additional supplemental private data (i.e., Walk Scores and Nielsen employment data), and geospatial information that reflects land use and physical environments, became evident after an examination of findings from the initial screening models. To be feasible, in terms of costs and time, the geographic scale of the study region was scaled down to nine selected NHTS add-on regions. These regions were chosen based on various criteria including transit availability, population size, and a mix of geographic locations across the nation. Given the similarities in modeling results from walk trips and walk mileages, additional modeling efforts conducted under the later part of this study were focused on walk trip per person.

Bike models were limited only with the stepwise logistic models using Census tracts in the selected regions. Due to NHTS sampling limitations, only about 12% of these tracts have bike trips recorded from NHTS sampled households. The modeling with NHTS bike data proved to be more challenging and time consuming than what was anticipated. Along with the late arrival of Nielsen employment data, the project team had to limit the modeling effort to focus on walking. Therefore, the final modeling and discriminant analysis was conducted only for walking trips.

The discriminating factors identified for walk trips per person include predictors such as driver and vehicle related variables, household and population counts, income, education, race and language, employment-related variables, obesity and walkability measures, trip purpose, and land use.

A final discriminant function with the final set of variables was developed using 75% of the records from the study data set. This function was then evaluated with the remaining 25% of the records from the study data. Examination of the performance of this discriminant function suggested promising results; with error rate on misclassifications at about 30%. Once the error rates for the testing dataset were determined to be acceptable, the entire dataset was used to compute the final scores.

7.2 CHALLENGES AND LESSONS LEARNED

Developing estimates on the propensity to travel non-motorized proved to be too large a task on a national scale. The reduction of geography to nine add-on areas with better NHTS sampling made outside data acquisition more affordable and walking models possible, although the scarcity of biking data was too difficult to overcome.

As mentioned previously, sampling limitations resulted in a lack of NHTS data (i.e., sampled-households) within many block groups in the study region. This data coverage issue impacts the lower income households living in inner-city/downtown areas especially. Without a proper representation in the sampled data, the model and the associated discriminant function developed based on data obtained from sampled block groups may not necessarily capture certain unique characteristics that exist among non-sampled block groups. For example, Figure 59 shows the NHTS coverage in the San Diego region, along with the estimated block group walk indices. Note that, San Diego was oversampled under the 2009 NHTS, thus its coverage is much better than most other NHTS-sampled areas.



Figure 61. NHTS sample coverage and walk index in the San Diego region; beige-shaded block groups contain one or more sampled households in the 2009 NHTS, those not shaded are non-sampled block groups.

As a comparison, sample coverage of block groups within Virginia, near Washington, DC, was much spottier (see Figure 60). Clearly, from the two maps shown in Figures 59 and 60, block groups with red dots (i.e., very low likelihood of walking) are more likely to be in un-shaded areas, signifying a data limitation in NHTS samples and, therefore, impacting results from the model. This also provides an indication that, with a better sample coverage, it is possible that the developed model would perform better. For a reference, maps showing NHTS coverage in other study regions are included in Appendix I of this report.

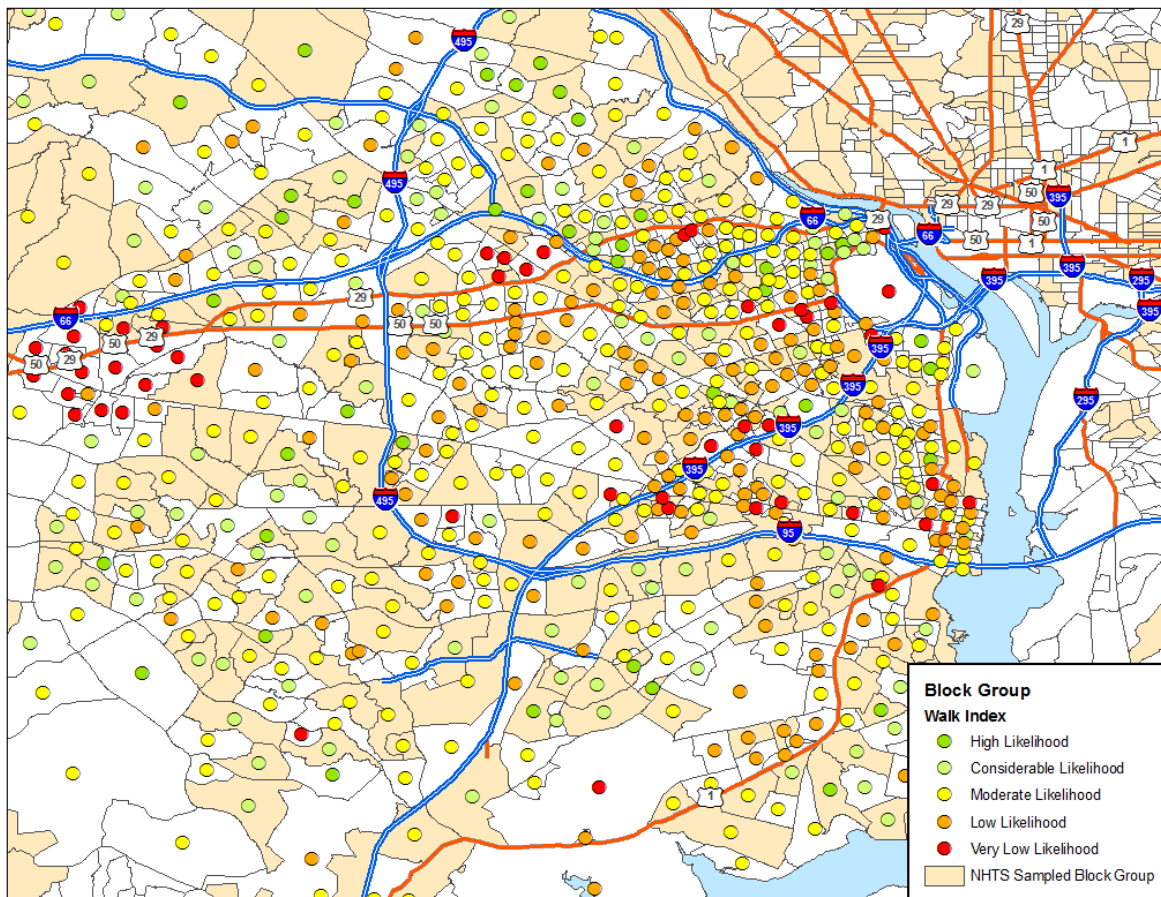


Figure 62. NHTS sample coverage and walk indices in Virginia near Washington DC; beige-shaded areas are block groups with NHTS sampled households and non-shaded areas are non-sampled block groups.

The sampling limitation also impacted the model's ability to accurately estimate characteristics in areas known for high walking activities, such as the City of Alexandria, Virginia. In fact, there are only 42 NHTS-sampled households in Alexandria, with 20 of these households reporting walk trips which resulted in a total of 57 walk trips. Using data from such a small sample, the model was unfortunately not able to reinforce the walking activities that should have been in this specific region. Furthermore, as discussed in Section 4.5.4, the majority of the missing block groups were imputed by using values from their corresponding Census tracts, which is much coarser in terms of representing local characteristics. As a result, a highly walked region, such as the City of Alexandria, Virginia, was not accurately captured by the current model due to the use of largely imputed block group level data. Actually, only 34 of the total 99

block groups in Alexandria had NHTS-sampled households. That is, values of NHTS variables in about two-thirds of the block groups within Alexandria were imputed. Because of this limitation, the model was not able to produce an index to accurately reflect the high walkability within this specific area. Greater sample coverage of the region, i.e., more data, could allow better calibration of the model and an improvement of accuracy would be expected.

In addition, transit accessibility is known to increase walking; higher walking activities would be expected in block groups around the transit stations. As shown in Figure 61, which overlays rail transit lines and their associated stations over a map similar to the one shown in Figure 59, low walk indices (red or orange dots) on or near the transit stations are clearly associated with non-sampled block groups. A similar map including rail transit information focusing on the Alexandria, VA area is shown in Figure 62. Again, this indicates a need for better data coverage so that models developed based on the data would be more accurate.



Figure 63. NHTS coverage, walk indices, and rail transit stations within the San Diego, CA area.

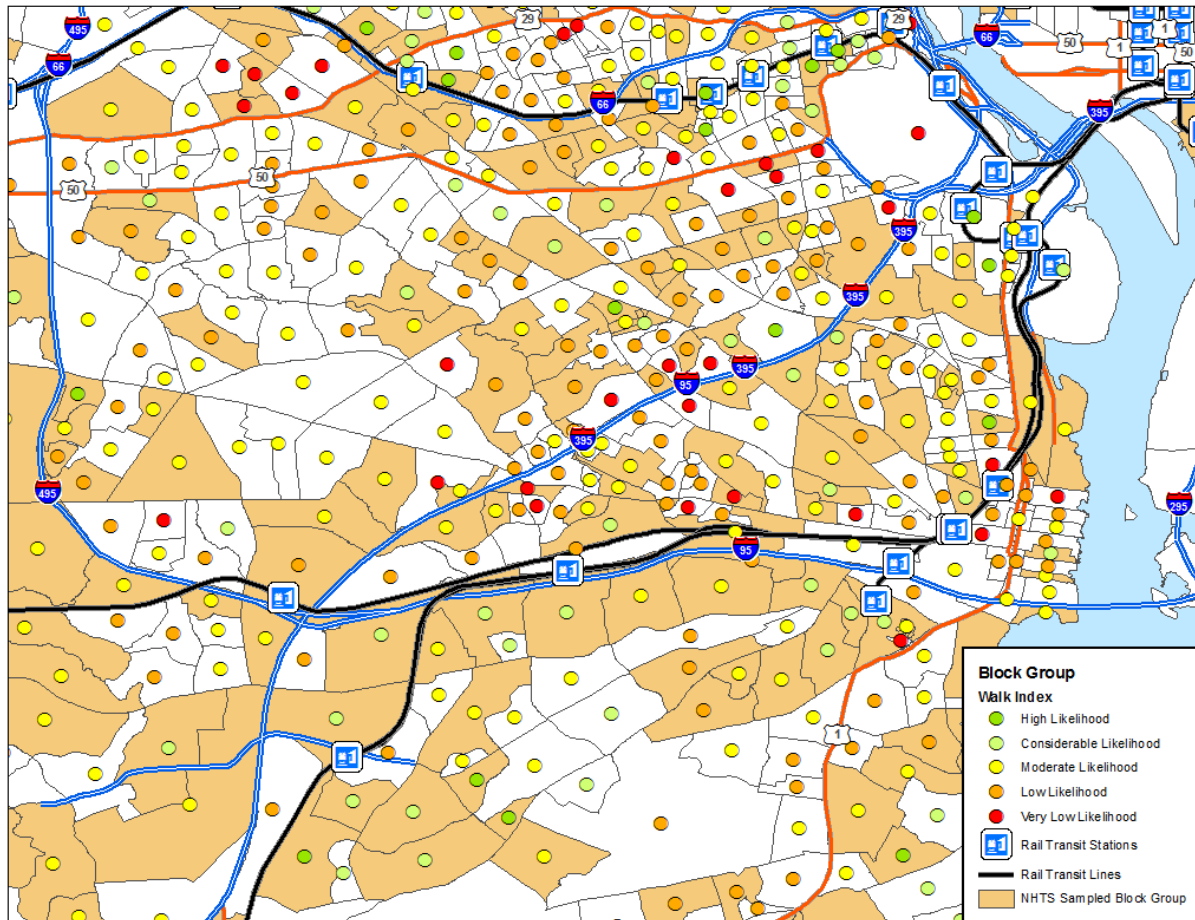


Figure 64. NHTS data coverage of block groups, walk indices, and accessibility for rail transit stations in the Alexandria, VA area.

Rail transit information was used in Figures 61 and 62 to illustrate the data coverage impact mainly because such data sets are readily available to the ORNL NMT team. This mode of transportation, unfortunately, is not common in most U.S. cities. On the other hand, transit buses have a wider coverage and availability to residents and workers in most U.S. cities. As discussed in Section 3.4.2, GTFS provides data on transit bus services (routes, schedules, and stops) that can be used to produce important measures that are relevant to NMT studies. This data source was not pursued mainly due to the scale of study region and the efforts necessary for obtaining information that could be used in the model development. Instead, Transit Scores at the block group centroids were used as the measure to associate walkability with transit accessibility.

7.3 FUTURE STUDIES OR NEXT STEPS

A national scale analysis such as that conducted in this study had the advantage of pulling all available data together to allow sufficient information in examining and identifying major factors that impact walking and biking. This data set, however, lacks the necessary small geographic detail that is needed for a model to echo localized activities such as walking and biking. Walk and bike activities are more likely to be associated with regional characteristics and local

environment. Moreover, as a national survey, the NHTS was not designed to provide coverage for tract level or block group level analyses. Thus, the determination of whether such a small geographic area is more likely to walk or not is difficult to make using this data set. The modeling framework used in this study, nevertheless, appeared to have reasonable performance (with about 30% error rate) when there is data. Although imputation methods were applied to produce estimates for missing variables so that discriminant functions could be applied to generate estimated walk indices for non-sampled block groups, these estimates (i.e., based on imputed variables) appeared to have some weaknesses.

Future studies may include the proposed Phase 2 of this research, which would involve developing a prototype of a visualization-based sketch planning tool for non-motorized travel. Specifically, to further examine how factors such as those identified from this current study, and the modeling framework developed under this study can be applied to local/regional level planning activities, a small geographic area with more detailed local data would be necessary. Although DC was not one of the 2009 NHTS add-ons, DC conducted a household travel survey of 11,000 households in 2007-2008. The National Capital Region Planning Board (TPB) at the Metropolitan Washington Council of Governments (MWCOC) conducted this household travel survey, which collected data from February 2007 through April 2008. Sample coverage of this survey is listed along with the NHTS 2009 survey in Table 9. Clearly, data coverage under the MWCOC survey, especially in the DC area (defined in Figure 63), is much higher than those in the NHTS.

Furthermore, the Office of Planning, Environment, & Realty in FHWA has information on DC's bus routes and timetables. With geo-coded information on household locations, the origin and destination of individual trips (where available), DC's bus routes, rail transit network, and walk and bike routes, such a database would provide detailed local information for conducting focused analysis that was not possible at the national level. In addition, the sidewalk inventory that FHWA has for the DC area, as well as specific data obtained by the ORNL NMT research team on schools, shopping malls, crime statistics, and pedestrian and bike fatalities for the region can all be integrated into a Geospatial Information System (GIS). With this GIS database, a visualization-based decision support tool that allows local planners to examine possible impacts in walking/biking behaviors under various scenarios of change in specific demographic, environmental, or other factors can be developed. The combination of travel data and transit service profile would also facilitate the future development of a "pedestrian-friendliness" index between origin and destination pairs. Such a visualization-based sketch planning tool could also be built with enhancements/capabilities to provide policy makers and planners a means to conduct informed investment decisions by predicting relative costs and benefits of potential investment scenarios.

Although transit access/egress walk trips were considered in the earlier part of this study as one of the potential dependent variables, the later part of the model development was focused only on the main walk trips per person, mainly due to the concern of additional missing variables at block group levels that already have data limitations. With a more focused region and increased data coverage, and improved model that takes into consideration all walk trips, including transit access/egress by walking, will also be more achievable.

Table 9. Comparisons of Total Number of Samples Captured in 2009 NHTS and 2007-2008 MWCOG Survey by Jurisdiction

JURISDICTION	MWCOG TOTAL COMPLETED SAMPLES	NHTS SAMPLED
District of Columbia	1,658	264
Montgomery County, MD	1,585	64
Prince George's County MD	1,482	34
Arlington County, VA	606	62
Alexandria City, VA	414	42
Fairfax County/Cities, VA	1,527	329
Loudon County, VA	383	73
Howard County, MD	540	23
Anne Arundel County, MD	764	31
Charles County, MD	231	8
Carroll County, MD	421	12
Calvert County, MD	104	3
St. Mary's County, MD	109	9
King George County, VA	56	94
Fredericksburg City, VA	46	56
Stafford County, VA	160	425
Spotsylvania County, VA	102	452
Fauquier County, VA	91	20
Clarke County, VA	56	83
Jefferson County, WV	113	8
Prince William County/Cities, VA	508	90
Frederick County/City, MD	373	16
Arlington+ (Columbia Pike)	97	
Arlington+ (Rosslyn-Ballston)	36	
Arlington+ (Shirlington)	10	
TOTAL	11,472	2,183

* Additional samples were collected under MWCOG survey for Arlington County (shaded cells).

APPENDIX A. VARIABLES OF INTERESTS BY DATA SOURCES IDENTIFIED DURING THE LITERATURE RESEARCH

Variable Description	ACS Summary	ACS PUMS	Claritas	NHTS	Census 2010	CTPP 2000	Other Sources
Demographic - Household							
Income	X	X	X	X		X	
Household size	X	X	X	X		X	
Number of Vehicles	X	X		X		X	
Number of Workers	X	X		X		X	
Number of Drivers				X			
Own/Rent Home	X	X		X			
Language	X	X		X			
MSA/Region	X	X	X	X	X	X	
Demographic - Person							
Age	X	X	X	X	X (18+)	X	
Gender	X	X		X		X	
Worker	X	X		X		X	
Class of Worker	X			X		X	
Driver				X			
Race/Hispanic Status	X	X	X	X	X	X	
Education	X	X	X	X			
Workplace				X		X	
College students	X						X
Land Use							
Population density			X	X	X		X (UIC)
Employment density			X	X		X	X (UIC)
Road density							X (UIC)
Intersection density							X (UIC)
Transit stops							X (UIC, ORNL)
Topography							X
Pavement quality							X
Bike lanes							(HPMS)
Sidewalk coverage							
Trip level data							
JTW Mode	X	X		X		X	

Variable Description	ACS Summary	ACS PUMS	Claritas	NHTS	Census 2010	CTPP 2000	Other Sources
							X (DVRPC)
JTW/Trip Time of Day	X	X		X		X	
JTW Minutes Travelled	X			X		X	
							X (DVRPC)
Bike trip information				X			
Walk trip information				X			
							X (Walk- score)
Walkability score							
Trip purpose	X (JTW only)	X (JTW only)		X		X (JTW only)	
Access/egress to transit				X			
Transit ridership				X			
VMT/Traffic counts				X			
Geocoded trip destinations				X			
Others							
							X (FARS)
Fatality data							
Weather/climate							X
Generator/attractor information							X
Crime rates							X
Traffic speed							X

APPENDIX B. VARIABLES USED OR DISCUSSED IN OTHER REFERENCES COMPILED DURING THE LITERATURE RESEARCH

Variable of Interests	References - Models that Utilize the Variable of Interests
Demographic - Household	
Income	Agrawal, A. W., and Schimek, P., "Extent and correlates of walking in the USA." <i>Transportation Research Part D</i> , Vol. 12, Issue 8 (December, 2007).
	Dill, J., and Carr, T., "Bicycle Commuting and Facilities in Major U.S. Cities: If You Build Them, Commuters Will Use Them." <i>Transportation Research Record</i> 1828 (2003).
	Dill, J., and Voros, K., "Factors Affecting Bicycling Demand: Initial Survey Findings from the Portland, Oregon, Region." <i>Transportation Research Record</i> 2031 (2007).
	Ginger et al, "Bicycle Planning, Best Practices, and Count Methodology" <i>The Transportation Studio</i> , Department of Urban Design and Planning, University of Washington, 2011.
	Krizek, K., et al. NCHRP 552: Guidelines for Analysis of Investment in Bicycle Facilities (2006).
	Kuzmyak, R. NCHRP Project No 08-78, "Estimating Bicycling and Walking for Planning and Project Development," Task 3 Interim Report (2011).
	Moudon, A. V., Lee, C., Cheadle, A. D., Collier, C. W., Johnson, D., Schmid, T. L., and Weathers, R. D., "Cycling and the built environment, a US perspective." <i>Transportation Research Part D</i> , Vol. 10, Issue 3 (May, 2005).
	Thunderhead Alliance, "Bicycling and Walking in the U.S.: Benchmarking Report 2007." Washington, DC. http://www.thunderheadalliance.org/benchmarking.htm (August, 2007).
	Wynne, G. A Study of Bicycle and Pedestrian Programs in European Countries. National Bicycling and Walking Study, Federal Highway Administration. FHWA Case Study Number xxxxx (1992).
Household Size	Frank, L., et al. "Development of a Walkability Index: Application to Neighborhood Quality of Life Study." <i>British Journal of Sports Medicine</i> , Vol. 44, (2010).
	Kuzmyak, J.R., Fregonese Associates, and Fehr & Peers. Local Sustainability Planning Model Development. Final Report. For Southern California Association of Governments (March 2010).
Number of Vehicles	Cervero, R., and Duncan, M., "Walking, Bicycling, and Urban Landscapes: Evidence From the San Francisco Bay Area." <i>American Journal of Public Health</i> , Vol. 93, No. 9 (September, 2003).
	Dill and Carr (2003)
	Dill and Voros (2007)
	Krizek et al (2006)
	Kuzmyak (2011)

Variable of Interests	References - Models that Utilize the Variable of Interests
	Moudon et al (2005)
	Pratt, R. NCHRP Project No 08-78, "Estimating Bicycling and Walking for Planning and Project Development," Task 1 Report: State of the Practice Review. (2010).
Number of Workers	Moudon et al (2005)
	Pratt (2010)
Number of Drivers	Dill and Voros (2007)
Own/Rent Home	
Language	
MSA/Region	Buehler, R. and J. Pucher. "Impacts of Bike Paths and Lanes on Cycling in Large American Cities." Paper presented at TRB Annual Meetings, January 2011
	Kuzmyak (2011)
Demographic - Person	
Age	Dill and Voros (2007)
	Ginger et al (2011)
	Moudon et al (2005)
	Pratt (2010)
Gender	Dill and Voros (2007)
	Krizek et al (2006)
	Moudon et al (2005)
	Pratt (2010)
Worker	Moudon et al (2005)
	Pratt (2010)
Class of Worker	Pratt (2010)
Driver	Dill and Voros (2007)
	Kuzmyak (2011)
Race/Hispanic Status	Cervero and Duncan (2003)
	Dill and Voros (2007)
	Pratt (2010)
Education	Krizek et al (2006)
Workplace	Baltes, M.R. Factors Influencing Nondiscretionary Work Trips by Bicycle Determined from 1990 US Census Metropolitan Statistical Area Data. Transportation Research Record 1538 (1996)
College Students	Dill and Carr (2003)
Land Use	
Population Density	Buehler and Pucher (2011)
	Dill and Carr (2003)
	Ginger et al (2011)
	Moudon et al (2005)

Variable of Interests	References - Models that Utilize the Variable of Interests
	Saelens, B. E., Sallis, J. F., Black, J. B. and Chen, D., "Neighborhood-Based Differences in Physical Activity: An Environmental Scale Evaluation." American Journal of Public Health, Vol. 93, No. 9 (September, 2003).
Employment Density	Ginger et al (2011)
Road Density	Dill and Voros (2007)
	Ewing, R., and Cervero, R., "Travel and the Built Environment — A Meta-Analysis." Journal of the American Planning Association, Vol. 6, No. 3 (Summer, 2010).
Intersection Density	Ewing and Cervero (2010)
Transit Stops	Alta Planning + Design, Seamless Travel Study. (2010)
	Desyllas, J., E. Duxbury, J. Ward, and A. Smith. PEDESTRIAN DEMAND MODELLING OF LARGE CITIES: AN APPLIED EXAMPLE FROM LONDON. Paper 62, Center for Advanced Spatial Analysis, University College of London (2003).
	Raford, N. and D. Ragland. SpaceSyntax: An innovative Pedestrian Modeling Tool for Pedestrian Safety. Safe Transportation Research & Education Center, UC Berkeley. Dec. 2003.
Topography	Broach, J., Gliebe, J., and Dill, J., "Development of a Multi-class Bicyclist Route Choice Model Using Revealed Preference Data." Paper submitted to the 12th International Conference on Travel Behavior Research, Jaipur, India (December 13-18, 2009).
	Cervero and Duncan (2003)
	Moudon et al (2005)
	NHTSA and BTS, "National Survey of Pedestrian & Bicyclist Attitudes and Behaviors: Highlights Report." Sponsored by the National Highway Traffic Safety Administration (NHTSA) and the Bureau of Transportation Statistics (BTS), U.S. Department of Transportation, and administered by The Gallup Organization. Washington, DC (2002).
Pavement Quality	Parkin, J., M. Wardman, and M. Page. "Estimation of the determinants of bicycle mode share for the journey to work using census data." Transportation (2008).
	Stinson, M. A. and C. R. Bhat. "A Comparison of the Route Preferences of Experienced and Inexperienced Bicycle Commuters." TRB 2005 Annual Meeting CD ROM (Nov. 2004).
Bike Lanes	Dill and Carr (2003)
	Moudon et al (2005)
Sidewalk Coverage	NHTSA and BTS (2002)
	Pratt, R. H., Evans, J. E., IV, Levinson, H. S., Turner, S., and Nabors, D., Chapter 16, "Pedestrian and Bicycle Facilities," TCRP Report 95 – Traveler Response to Transportation System Changes, Preliminary Draft (PDFR), Transportation Research Board, Washington, DC (September, 2010).

Variable of Interests	References - Models that Utilize the Variable of Interests
	Saelens, B. E., and Handy, S. L., "Built Environment Correlates of Walking: A Review." <i>Medicine & Science in Sports & Exercise</i> , Vol. 40, Issue 7, Supplement 1, (July, 2008).
Trip level data	
JTW Mode	Kuzmyak (2011)
	Krizek et al (2006)
JTW/Trip Time of Day	Ginger et al (2011)
JTW Minutes Travelled	Goldsmith, S. Estimating the Effect of Bicycle Facilities on VMT and Emissions. Seattle Engineering Dept. (1997).
	Kuzmyak (2011)
Bike Trip Information	Kuzmyak (2011)
Walk Trip Information	Kuzmyak (2011)
Walkability Scores	Frank et al (2010)
	Kuzmyak, J. R., Baber, C., and Savory, D., "Use of a Walk Opportunities Index to Quantify Local Accessibility." <i>Transportation Research Record 1977</i> (2006).
	Walkscore.com
Trip Purpose	NHTSA and BTS (2002)
Transit	Alta Planning + Design (2010)
	Desyllas, et al. (2003)
	Raford and Ragland (2003)
VMT/Traffic Counts	Krizek, K. J., "Residential Relocation and Changes in Urban Travel: Does Neighborhood-Scale Urban Form Matter?" <i>Journal of the American Planning Association</i> , Vol. 69, No. 3 (Summer, 2003).
Geocoded Trip Destinations	NHTSA and BTS (2002)
Others	
Fatality Data	Buehler and Pucher (2011)
Fuel Cost	Kuzmyak (2011)
Weather/Climate	Cervero and Duncan (2003)
	Dill and Carr (2003)
	Ginger et al (2011)
Generator/ Attractor Information	Matlick, J.M. If We Built It, Will They Come? Bicycle and Pedestrian Federations of America, Pro Walk/Pro-Bike '96 (1996)
	Raford and Ragland (2003)
Crime Rates	Agrawal and Schimek (2007)
	Pratt et al (2010)

Variable of Interests	References - Models that Utilize the Variable of Interests
	Targa, F., and Clifton, K. J., Built Environment and Trip Generation for Non-Motorized Travel. Paper summary submitted for presentation. National Household Travel Survey Data Conference: Data for Understanding Our Nation's Travel, Washington, DC. http://www.trb.org/conferences/nhts/Clifton.pdf (November 1-2, 2004).
Traffic Speed	Hall, R.E. HPE's Walkability Index -- Quantifying the Pedestrian Experience. 2010 ITE Technical Conference.
	Raford and Ragland (2003)
	Wynne (1992)

APPENDIX C. INITIAL NATIONAL SCALED MODELS SUMMARY RESULT TABLES

C.1 Estimated Walk Trips Per Person Stepwise-Selected Logistic Regression Model Number of Observations Used: 46,354, Cox & Snell R-Square: 22.2

Predictor Variable	Parameter Estimate	Standard Error	Nom. Sig. Level
Percent Weighted Trips in Destination Block Group for the Purpose Home-Based Social-Recreational, All Modes (NHTS+: HBSOCREC_DestPct)	1.6486	0.0542	4.E-203
Percent Persons in BG with Educ of HS or Less (NHTS: PctHSLess)	-0.6153	0.0459	5.1E-41
Sample HHs in BG (NHTS: NHHs)	0.2452	0.0199	9.7E-35
Average Worker Count of HHs in BG (NHTS: AvgWorkers)	0.1961	0.0190	5.1E-25
County-level Percentage of Adults Who Are Obese (Other: PctObese)	-0.0369	0.0037	1.4E-23
Average Vehicle Count of HHs in BG (NHTS: AvgVehCount)	-0.1546	0.0161	6.7E-22
Sample Persons in BG (NHTS: NPers)	0.0922	0.0097	1.6E-21
Percent Range CY Pop 25+, College Graduate (Claritas: HBPCOLGRD)	0.0073	0.0008	1.6E-21
Percent Drivers in BG (NHTS: PctDrivers)	-0.5156	0.0595	4.2E-18
Percent Weighted Trips in Destination Block Group for the Purpose Home-Based Other, All Modes (NHTS+: HBO_DestPct)	0.3637	0.0495	1.9E-13
Average Driver Count of HHs in BG (NHTS: AvgDrivers)	0.1901	0.0272	2.6E-12
CY Median Year HU Structure Built (Claritas: HBHMEDYR)	-0.0055	0.0008	7.8E-12
Percent Range CY Housing Units, Renter Occ (Claritas: HBHTNRNT)	0.0039	0.0006	3.2E-11
Intercept	10.3434	1.5786	5.7E-11
Percent Persons in BG with Some College/Bachelor's (NHTS: PctCollege)	-0.2772	0.0440	2.9E-10
Wtd Pct HHs in BG with English Interviews (NHTS: PctEngHH)	-0.3568	0.0599	2.6E-09
CY Own Occ HUs Median Value (Claritas: HBHMEDHS)	0.0000	0.0000	3.5E-09

Predictor Variable	Parameter Estimate	Standard Error	Nom. Sig. Level
Wtd Pct HHs in BG Owning HH (NHTS: PctOwnHH)	-0.1918	0.0349	3.9E-08
Households by Occupancy Status, Vacant (ACS: B25002e3)	0.0003	0.0001	3.8E-07
Employee Range CY Employees per square mile (Claritas: HBEEMPDN)	0.0000	0.0000	4.4E-07
Percent Range CY Pop, Asian (Claritas: HBPRCASN)	-0.0067	0.0013	4.6E-07
Percent Weighted Trips in Destination Block Group for the Purpose Home-Based Work, All Modes (NHTS+: HBW_DestPct)	-0.3245	0.0661	9.2E-07
Percent Males in BG (NHTS: PctMales)	0.1986	0.0418	2.0E-06
Workers by Means Of Transportation To Work, Public transportation (excluding taxicab): Subway or elevated (ACS: B08301e13)	0.0017	0.0004	4.0E-06
Town & Country Indicator (Claritas: HBHUR_TC)	-0.1197	0.0277	1.5E-05

**C.2 Estimated Walk Trips Per Person (When Positive)
Stepwise-Selected Least Squares Regression Model
Number of Observations Used: 18,388, R-Square: 43.8**

Predictor Variable	Parameter Estimate	Standard Error	Nom. Sig. Level
Intercept	16.2957	0.6283	1.E-145
CY Median Year HU Structure Built (Claritas: HBHMEDYR)	-0.0073	0.0003	2.E-115
Households in BG (NHTS: TotHHs)	0.0001	0.0000	4.E-105
Percent Weighted Trips in Destination Block Group for the Purpose Home-Based Social-Recreational, All Modes (NHTS+: HBSOCREC_DestPct)	0.3644	0.0214	8.6E-65
Wtd Pct HHs in BG Owning HH (NHTS: PctOwnHH)	-0.1895	0.0125	8.8E-52
Average HH Size of HHs in BG (NHTS: AvgHHSIZE)	0.0756	0.0051	7.9E-49
Sample HHs in BG (NHTS: NHHs)	-0.0579	0.0042	1.7E-43
Persons by Sex By School Enrollment By Level Of School For The Population 3 Years And Over, Total (ACS: B14002e1)	-0.0001	0.0000	4.7E-38
CY Own Occ HUs Median Value (Claritas: HBHMEDHS)	0.0000	0.0000	1.2E-36
Median Age, Total (ACS: B01002e1)	0.0061	0.0005	1.2E-33
County-level Gradation of percentage of physically inactive adults by total number of physically inactive adults in 2008 (Other: LIPERTOT08)	-0.0204	0.0019	4.0E-26
Average Vehicle Count of HHs in BG (NHTS: AvgVehCount)	-0.0517	0.0052	2.0E-23
Sample Persons in BG (NHTS: NPers)	0.0160	0.0019	9.4E-17
Wtd Pct Afr Am. HHs in BG (NHTS: PctAfAmHH)	0.1537	0.0187	2.0E-16
Employee Range CY Employees per square mile (Claritas: HBEEMPDN)	0.0000	0.0000	2.0E-16
Percent Weighted Trips in Destination Block Group for the Purpose Home-Based Other, All Modes (NHTS+: HBO_DestPct)	0.1491	0.0200	1.0E-13
Wtd Pct Hisp HHs in BG (NHTS: PctHispHH)	0.1067	0.0151	1.4E-12
Persons in BG (NHTS: TotPers)	0.0000	0.0000	3.1E-12
Workers by Means Of Transportation To Work By Travel Time To Work, Car, truck, or van - drove alone: 20 to 24 minutes (ACS: C08134e15)	-0.0002	0.0000	1.3E-11

Predictor Variable	Parameter Estimate	Standard Error	Nom. Sig. Level
Percent Weighted Trips in Destination Block Group for the Purpose Not Home-Based, All Modes (NHTS+: NHB_DestPct)	-0.1303	0.0201	1.0E-10
Percent Range CY Pop, White (Claritas: HBPRCCAU)	-0.0016	0.0003	3.0E-10
Percent Range CY Pop, Foreign born (Claritas: HBPFORBN)	0.0022	0.0004	9.7E-10
Percent Workers in BG (NHTS: PctWorkers)	0.1400	0.0229	1.0E-09
Household Income In The Past 12 Months, \$100,000 to \$124,999 (ACS: B19001e14)	0.0004	0.0001	2.4E-09
Percent Drivers in BG (NHTS: PctDrivers)	-0.1150	0.0196	4.1E-09
Percent Range CY Families Below Poverty (Claritas: HBPLTPOV)	0.0026	0.0005	6.1E-09
County-level Percentage of Adults Who Are Obese (Other: PctObese)	0.0070	0.0012	1.7E-08
Persons by Sex By Educational Attainment For The Population 25 Years And Over, Female: 7th and 8th grade (ACS: B15002e23)	-0.0009	0.0002	5.8E-08
Persons by Race, American Indian and Alaska Native alone (ACS: B02001e4)	-0.0004	0.0001	1.0E-07
Average Worker Count of HHs in BG (NHTS: AvgWorkers)	0.0549	0.0112	9.6E-07
Percent Persons in BG with Grad/Prof Degree (NHTS: PctGrad)	0.0780	0.0163	1.7E-06
Persons by Sex & Age, Male: 50 to 54 years (ACS: B01001e16)	-0.0004	0.0001	1.8E-06
Persons by Sex & Age, Female: 30 to 34 years (ACS: B01001e36)	0.0003	0.0001	4.4E-06
Persons by Race, Black or African American alone (ACS: B02001e3)	-0.0001	0.0000	4.5E-06
Percent Range CY Housing Units, Owner Occ (Claritas: HBHTNOWN)	-0.0010	0.0002	1.8E-05
Percent Males in BG (NHTS: PctMales)	0.0659	0.0155	2.1E-05
Percent Range CY Pop 20-64 (Claritas: HBP20TO64)	0.0021	0.0005	2.2E-05
Percent Range CY HUs, Built 2000 or Later (Claritas: HBHRCENT)	0.0013	0.0003	3.7E-05
Workers by Travel Time To Work, 90 or more minutes (ACS: B08303e13)	0.0004	0.0001	3.9E-05

Predictor Variable	Parameter Estimate	Standard Error	Nom. Sig. Level
Household Income In The Past 12 Months, Less than \$10,000 (ACS: B19001e2)	-0.0003	0.0001	4.7E-05
Town & Country Indicator (Claritas: HBHUR_TC)	0.0388	0.0095	4.8E-05

C.3 Estimated Walk Mileage Per Person
Stepwise-Selected Logistic Regression Model
Number of Observations Used: 46,354, Cox & Snell R-Square: 21.8

Predictor Variable	Parameter Estimate	Standard Error	Nom. Sig. Level
Percent Weighted Trips in Destination Block Group for the Purpose Home-Based Social-Recreational, All Modes (NHTS+: HBSOCREC_DestPct)	1.6361	0.0542	4.E-200
Percent Persons in BG with Educ of HS or Less (NHTS: PctHSLess)	-0.6097	0.0459	3.5E-40
Sample HHs in BG (NHTS: NHHs)	0.2328	0.0198	6.3E-32
Average Worker Count of HHs in BG (NHTS: AvgWorkers)	0.1986	0.0190	1.5E-25
Average Vehicle Count of HHs in BG (NHTS: AvgVehCount)	-0.1579	0.0161	1.1E-22
County-level Percentage of Adults Who Are Obese (Other: PctObese)	-0.0361	0.0037	1.1E-22
Sample Persons in BG (NHTS: NPers)	0.0937	0.0096	1.8E-22
Percent Range CY Pop 25+, College Graduate (Claritas: HBPCOLGRD)	0.0072	0.0008	5.4E-21
Percent Drivers in BG (NHTS: PctDrivers)	-0.4906	0.0595	1.7E-16
Percent Weighted Trips in Destination Block Group for the Purpose Home-Based Other, All Modes (NHTS+: HBO_DestPct)	0.3612	0.0495	3.1E-13
CY Median Year HU Structure Built (Claritas: HBHMEDYR)	-0.0058	0.0008	4.3E-13
Average Driver Count of HHs in BG (NHTS: AvgDrivers)	0.1877	0.0272	4.9E-12
Intercept	10.8604	1.5787	6.0E-12
Percent Range CY Housing Units, Renter Occ (Claritas: HBHTNRNT)	0.0038	0.0006	1.6E-10
CY Own Occ HUs Median Value (Claritas: HBHMEDHS)	0.0000	0.0000	9.1E-10
Percent Persons in BG with Some College/Bachelor's (NHTS: PctCollege)	-0.2686	0.0440	1.0E-09
Wtd Pct HHs in BG Owning HH (NHTS: PctOwnHH)	-0.1910	0.0349	4.6E-08
Percent Range CY Pop, Asian (Claritas: HBPRCASN)	-0.0069	0.0013	2.1E-07
Households by Occupancy Status, Vacant (ACS: B25002e3)	0.0003	0.0001	6.6E-07
Percent Males in BG (NHTS: PctMales)	0.2081	0.0419	6.6E-07

Predictor Variable	Parameter Estimate	Standard Error	Nom. Sig. Level
Employee Range CY Employees per square mile (Claritas: HBEEMPDN)	0.0000	0.0000	8.0E-07
Percent Weighted Trips in Destination Block Group for the Purpose Home-Based Work, All Modes (NHTS+: HBW_DestPct)	-0.3131	0.0662	2.2E-06
Workers by Means Of Transportation To Work, Public transportation (excluding taxicab): Subway or elevated (ACS: B08301e13)	0.0017	0.0004	2.3E-06
Wtd Pct HHs in BG with English Interviews (NHTS: PctEngHH)	-0.2793	0.0602	3.5E-06
Town & Country Indicator (Claritas: HBHUR_TC)	-0.1166	0.0277	2.5E-05

**C.4 Estimated Walk Mileage Per Person (When Positive)
Stepwise-Selected Least Squares Regression Model
Number of Observations Used: 18,151, R-Square: 32.8**

Predictor Variable	Parameter Estimate	Standard Error	Nom. Sig. Level
Intercept	14.2656	0.6427	1.E-107
Percent Weighted Trips in Destination Block Group for the Purpose Home-Based Social-Recreational, All Modes (NHTS+: HBSOCREC_DestPct)	0.5470	0.0254	2.E-101
CY Median Year HU Structure Built (Claritas: HBHMEDYR)	-0.0064	0.0003	6.0E-87
Households in BG (NHTS: TotHHs)	0.0001	0.0000	1.1E-67
Wtd Pct HHs in BG Owning HH (NHTS: PctOwnHH)	-0.1947	0.0145	6.6E-41
Average HH Size of HHs in BG (NHTS: AvgHHSIZE)	0.0777	0.0060	4.2E-38
CY Own Occ HUs Median Value (Claritas: HBHMEDHS)	0.0000	0.0000	3.2E-36
Sample HHs in BG (NHTS: NHHs)	-0.0629	0.0051	2.4E-35
County-level Gradation of percentage of physically inactive adults by total number of physically inactive adults in 2008 (Other: LIPERTOT08)	-0.0249	0.0022	5.2E-29
Persons by Sex By School Enrollment By Level Of School For The Population 3 Years And Over, Total (ACS: B14002e1)	-0.0001	0.0000	4.0E-27
Wtd Pct White HHs in BG (NHTS: PctWhiteHH)	-0.1572	0.0152	6.3E-25
Percent Weighted Trips in Destination Block Group for the Purpose Home-Based Other, All Modes (NHTS+: HBO_DestPct)	0.2351	0.0247	2.0E-21
Sample Persons in BG (NHTS: NPers)	0.0202	0.0023	4.6E-18
Average Vehicle Count of HHs in BG (NHTS: AvgVehCount)	-0.0479	0.0060	1.5E-15
Median Age, Female (ACS: B01002e3)	0.0038	0.0005	1.3E-12
Persons by Sex By Educational Attainment For The Population 25 Years And Over, Male: Bachelor's degree (ACS: B15002e15)	0.0003	0.0000	1.7E-11
Percent Range CY Families Below Poverty (Claritas: HBPLTPOV)	0.0034	0.0005	5.2E-11
Employee Range CY Employees per square mile (Claritas: HBEEMPDN)	0.0000	0.0000	6.6E-11
Percent Range CY Pop, Foreign born (Claritas: HBPFORBN)	0.0025	0.0004	1.2E-10

Predictor Variable	Parameter Estimate	Standard Error	Nom. Sig. Level
Workers by Means Of Transportation To Work By Travel Time To Work, Car, truck, or van - drove alone: 20 to 24 minutes (ACS: C08134e15)	-0.0002	0.0000	1.3E-09
Percent Persons in BG with Grad/Prof Degree (NHTS: PctGrad)	0.1151	0.0198	6.3E-09
Persons by Sex & Age, Male: 50 to 54 years (ACS: B01001e16)	-0.0006	0.0001	8.5E-09
Workers by Travel Time To Work, 90 or more minutes (ACS: B08303e13)	0.0006	0.0001	2.5E-08
Persons in BG (NHTS: TotPers)	0.0000	0.0000	1.2E-07
Percent Males in BG (NHTS: PctMales)	0.1003	0.0190	1.3E-07
Percent Workers in BG (NHTS: PctWorkers)	0.1380	0.0275	5.5E-07
Household Income In The Past 12 Months, Less than \$10,000 (ACS: B19001e2)	-0.0004	0.0001	1.0E-06
Percent Weighted Trips in Destination Block Group for the Purpose Home-Based Shopping, All Modes (NHTS+: HBSHOP_DestPct)	0.1149	0.0244	2.4E-06
Average Worker Count of HHs in BG (NHTS: AvgWorkers)	0.0639	0.0137	3.0E-06
Percent Range CY Pop 20-64 (Claritas: HBP20TO64)	0.0027	0.0006	5.3E-06
Wtd Pct Asian Am. HHs in BG (NHTS: PctAsAmHH)	-0.1556	0.0348	7.9E-06
Persons by Sex By Educational Attainment For The Population 25 Years And Over, Female: 7th and 8th grade (ACS: B15002e23)	-0.0009	0.0002	1.0E-05

C.5 Estimated Bike Trips Per Person
Stepwise-Selected Logistic Regression Model
Number of Observations Used: 46,354, Cox & Snell R-Square: 6.3

Predictor Variable	Parameter Estimate	Standard Error	Nom. Sig. Level
Sample Persons in BG (NHTS: NPers)	0.0670	0.0022	3.E-197
Percent Weighted Trips in Destination Block Group for the Purpose Home-Based Social-Recreational, All Modes (NHTS+: HBSOCREC_DestPct)	1.4395	0.0929	4.2E-54
Percent Males in BG (NHTS: PctMales)	1.1667	0.0927	2.7E-36
Average Worker Count of HHs in BG (NHTS: AvgWorkers)	0.3676	0.0341	4.1E-27
Percent Persons in BG with Educ of HS or Less (NHTS: PctHSLess)	-1.4320	0.1371	1.6E-25
Percent Persons in BG with Some College/Bachelor's (NHTS: PctCollege)	-1.2326	0.1407	2.0E-18
County-level Percentage of Adults Who Are Obese (Other: PctObese)	-0.0534	0.0061	2.7E-18
Percent Range CY Pop 25+, College Graduate (Claritas: HBPCOLGRD)	0.0102	0.0012	1.3E-17
Percent Drivers in BG (NHTS: PctDrivers)	-0.8802	0.1195	1.8E-13
Intercept	-1.4983	0.2043	2.2E-13
Percent Persons in BG with Grad/Prof Degree (NHTS: PctGrad)	-1.0646	0.1623	5.4E-11
Workers by Means Of Transportation To Work, Bicycle (ACS: B08301e18)	0.0058	0.0009	7.4E-10
Average Vehicle Count of HHs in BG (NHTS: AvgVehCount)	-0.1720	0.0282	1.1E-09
Percent Range CY Pop, Asian (Claritas: HBPRCASN)	-0.0140	0.0027	3.0E-07
Households by Occupancy Status, Vacant (ACS: B25002e3)	0.0004	0.0001	1.3E-05
Persons by Sex & Age, Female: 75 to 79 years (ACS: B01001e47)	0.0020	0.0005	1.3E-05

**C.6 Estimated Bike Trips Per Person (When Positive)
Stepwise-Selected Least Squares Regression Model
Number of Observations Used: 2,760, R-Square: 50.6**

Predictor Variable	Parameter Estimate	Standard Error	Nom. Sig. Level
Sample HHs in BG (NHTS: NHHs)	-0.0315	0.0016	1.2E-80
Intercept	19.6156	1.3268	1.2E-47
CY Median Year HU Structure Built (Claritas: HBHMEDYR)	-0.0088	0.0007	3.0E-38
Persons by Sex By School Enrollment By Level Of School For The Population 3 Years And Over, Total (ACS: B14002e1)	-0.0001	0.0000	1.6E-26
Percent Weighted Trips in Destination Block Group for the Purpose Home-Based Social-Recreational, All Modes (NHTS+: HBSOCREC_DestPct)	0.5649	0.0526	2.3E-26
Wtd Pct HHs in BG Owning HH (NHTS: PctOwnHH)	-0.2406	0.0301	1.9E-15
Percent Males in BG (NHTS: PctMales)	0.3391	0.0467	5.0E-13
County-level Gradation of percentage of physically inactive adults by total number of physically inactive adults in 2008 (Other: LIPERTOT08)	-0.0270	0.0042	1.9E-10
Average HH Size of HHs in BG (NHTS: AvgHHSIZE)	0.0669	0.0108	6.7E-10
Employee Range CY Employees per square mile (Claritas: HBEEMPDN)	0.0000	0.0000	2.4E-09
Percent Workers in BG (NHTS: PctWorkers)	0.2054	0.0373	4.1E-08
Households in BG (NHTS: TotHHs)	0.0001	0.0000	2.8E-07
Percent Drivers in BG (NHTS: PctDrivers)	-0.2420	0.0505	1.8E-06
Persons in BG (NHTS: TotPers)	0.0000	0.0000	3.5E-05
Persons by Sex By School Enrollment By Level Of School For The Population 3 Years And Over, Male: Enrolled in grade 5 to grade 8 (ACS: B14002e13)	0.0007	0.0002	6.7E-05

C.7 Estimated Bike Mileage Per Person
Stepwise-Selected Logistic Regression Model
Number of Observations Used: 46,354, Cox & Snell R-Square: 6.1

Predictor Variable	Parameter Estimate	Standard Error	Nom. Sig. Level
Sample Persons in BG (NHTS: NPers)	0.0660	0.0022	6.E-191
Percent Weighted Trips in Destination Block Group for the Purpose Home-Based Social-Recreational, All Modes (NHTS+: HBSOCREC_DestPct)	1.4315	0.0938	1.4E-52
Percent Males in BG (NHTS: PctMales)	1.1618	0.0935	2.0E-35
Average Worker Count of HHs in BG (NHTS: AvgWorkers)	0.3680	0.0344	1.0E-26
Percent Persons in BG with Educ of HS or Less (NHTS: PctHSLess)	-1.4182	0.1390	1.9E-24
Percent Range CY Pop 25+, College Graduate (Claritas: HBPCOLGRD)	0.0106	0.0012	1.5E-18
County-level Percentage of Adults Who Are Obese (Other: PctObese)	-0.0533	0.0062	5.6E-18
Percent Persons in BG with Some College/Bachelor's (NHTS: PctCollege)	-1.1795	0.1423	1.1E-16
Intercept	-1.5561	0.2064	4.8E-14
Percent Drivers in BG (NHTS: PctDrivers)	-0.8744	0.1209	4.8E-13
Percent Persons in BG with Grad/Prof Degree (NHTS: PctGrad)	-1.0230	0.1638	4.2E-10
Workers by Means Of Transportation To Work, Bicycle (ACS: B08301e18)	0.0057	0.0010	1.5E-09
Average Vehicle Count of HHs in BG (NHTS: AvgVehCount)	-0.1720	0.0285	1.6E-09
Percent Range CY Pop, Asian (Claritas: HBPRCASN)	-0.0140	0.0027	3.4E-07
Persons by Sex & Age, Female: 75 to 79 years (ACS: B01001e47)	0.0019	0.0005	3.4E-05
Households by Occupancy Status, Vacant (ACS: B25002e3)	0.0003	0.0001	8.6E-05

C.8 Estimated Bike Mileage Per Person (When Positive)
Stepwise-Selected Least Squares Regression Model
Number of Observations Used: 2,698, R-Square: 33.4

Predictor Variable	Parameter Estimate	Standard Error	Nom. Sig. Level
Persons in BG (NHTS: TotPers)	0.0000	0.0000	4.2E-42
Sample HHs in BG (NHTS: NHHs)	-0.0286	0.0022	1.8E-37
Intercept	21.1407	1.8639	3.7E-29
Persons by Sex & Age, Female (ACS: B01001e26)	-0.0001	0.0000	2.6E-28
CY Median Year HU Structure Built (Claritas: HBHMEDYR)	-0.0096	0.0009	5.6E-24
Percent Workers in BG (NHTS: PctWorkers)	0.3796	0.0497	3.1E-14
CY Own Occ HUs Median Value (Claritas: HBHMEDHS)	0.0000	0.0000	7.5E-14
Wtd Pct HHs in BG Owning HH (NHTS: PctOwnHH)	-0.2688	0.0408	5.4E-11
Percent Weighted Trips in Destination Block Group for the Purpose Home-Based Social-Recreational, All Modes (NHTS+: HBSOCREC_DestPct)	0.4326	0.0744	6.7E-09
Percent Males in BG (NHTS: PctMales)	0.3360	0.0668	5.2E-07

APPENDIX D. STATISTICAL SUMMARY TABLES WITH THE RESCALED GEOGRAPHIC REGIONS PRIOR TO NIELSEN EMPLOYMENT DATA

D.1 Estimated Walk Trips Per Person Stepwise-Selected Logistic Regression Model Number of Observations Used: 6,911, Cox & Snell R-Square: 11.0

Predictor Variable	Parameter Estimate	Standard Error
Percent Range CY Pop, White (Claritas: HBPRCCAU)	0.0124	0.0012
Households in BG (NHTS: TotHHs)	0.0002	0.0000
Percent Drivers in BG (NHTS: PctDrivers)	-0.6812	0.1161
County-level Percentage of Adults Who Are Obese (Other: PctObese)	-0.0570	0.0104
Workers by Time Leaving Home To Go To Work, 6:30 a.m. to 6:59 a.m. (ACS: B08302e6)	0.0021	0.0004
Percent Persons in BG with Educ of HS or Less (NHTS: PctHSLess)	-0.4270	0.0832
Household Income In The Past 12 Months, \$150,000 to \$199,999 (ACS: B19001e16)	0.0027	0.0005
Percent Weighted Trips in Destination Block Group for the Purpose Work & Related, All Modes (NHTS+: Work_DestPct)	-0.9205	0.1832
Percent Weighted Trips in Destination Block Group for the Purpose Retail (Shopping/Meals/Services), All Modes (NHTS+: Retail_DestPct)	-0.5663	0.1134
Average HH Size of HHs in BG (NHTS: AvgHHSIZE)	0.1370	0.0278
Average Worker Count of HHs in BG (NHTS: AvgWorkers)	0.1919	0.0412
County-level Gradation of percentage of physically inactive adults by total number of physically inactive adults in 2008 (Other: LIPERTOT08)	0.0640	0.0150
Percent Range CY Pop, Foreign born (Claritas: HBPFORBN)	-0.0068	0.0016
Persons by Race, Two races excluding Some other race, and three or more races (ACS: B02001e10)	0.0021	0.0005
Walk Score (Other: WalkScore)	0.0054	0.0013
Workers by Means Of Transportation To Work By Travel Time To Work, Car, truck, or van - drove alone: 60 or more minutes (ACS: C08134e20)	-0.0016	0.0005
Wtd Pct HHs in BG Owning HH (NHTS: PctOwnHH)	-0.2600	0.0783
Intercept	-0.0952	0.3412

D.2 Estimated Walk Trips Per Person (When Positive)
Stepwise-Selected Least Squares Regression Model
Number of Observations Used: 3,315, R-Square: 23.1

Predictor Variable	Parameter Estimate	Standard Error
Intercept	2.2146	0.0729
Households in BG (NHTS: TotHHs)	-0.0001	0.0000
CY Own Occ HUs Median Value (Claritas: HBHMEDHS)	0.0000	0.0000
Average Worker Count of HHs in BG (NHTS: AvgWorkers)	-0.1155	0.0164
Walk Score (Other: WalkScore)	0.0026	0.0004
Percent Range CY Pop, White (Claritas: HBPRCCAU)	-0.0023	0.0004
Percent Range CY Pop, Hisp/Lat (Claritas: HBPHISP)	0.0020	0.0003
Average Vehicle Count of HHs in BG (NHTS: AvgVehCount)	-0.0577	0.0096
Percent Persons in BG with Grad/Prof Degree (NHTS: PctGrad)	0.1768	0.0304
Percent Range CY Pop 20-64 (Claritas: HBP20TO64)	0.0050	0.0010
Percent Workers in BG (NHTS: PctWorkers)	0.1753	0.0349
County-level Gradation of percentage of physically inactive adults by total number of physically inactive adults in 2008 (Other: LIPERTOT08)	-0.0213	0.0043
Persons by Sex By Educational Attainment For The Population 25 Years And Over, Female: High school graduate, GED, or alternative (ACS: B15002e28)	-0.0002	0.0000
ACSDENSITY	0.0000	0.0000

D.3 Estimated Walk Mileage Per Person
Stepwise-Selected Logistic Regression Model
Number of Observations Used: 6,911, Cox & Snell R-Square: 10.9

Predictor Variable	Parameter Estimate	Standard Error
Percent Range CY Pop, White (Claritas: HBPRCCAU)	0.0111	0.0013
Percent Drivers in BG (NHTS: PctDrivers)	-0.6588	0.1163
Households in BG (NHTS: TotHHs)	0.0001	0.0000
Average HH Size of HHs in BG (NHTS: AvgHHSIZE)	0.1464	0.0278
Average Worker Count of HHs in BG (NHTS: AvgWorkers)	0.2167	0.0414
County-level Percentage of Adults Who Are Obese (Other: PctObese)	-0.0532	0.0104
Percent Weighted Trips in Destination Block Group for the Purpose Work & Related, All Modes (NHTS+: Work_DestPct)	-0.9358	0.1838
Percent Persons in BG with Educ of HS or Less (NHTS: PctHSLess)	-0.4238	0.0835
Percent Weighted Trips in Destination Block Group for the Purpose Retail (Shopping/Meals/Services), All Modes (NHTS+: Retail_DestPct)	-0.5544	0.1132
Persons by Race, Two races excluding Some other race, and three or more races (ACS: B02001e10)	0.0022	0.0005
Walk Score (Other: WalkScore)	0.0058	0.0013
Household Income In The Past 12 Months, \$150,000 to \$199,999 (ACS: B19001e16)	0.0023	0.0005
Workers by Means Of Transportation To Work By Travel Time To Work, Car, truck, or van - drove alone: 15 to 19 minutes (ACS: C08134e14)	0.0013	0.0003
Wtd Pct HHs in BG Owning HH (NHTS: PctOwnHH)	-0.3027	0.0785
Percent Range CY Pop, Foreign born (Claritas: HBPFORBN)	-0.0060	0.0016
Percent Range CY Pop 65+ (Claritas: HBP65P)	0.0108	0.0029
Intercept	0.1006	0.3253

**D.4 Estimated Walk Mileage Per Person (When Positive)
Stepwise-Selected Least Squares Regression Model
Number of Observations Used: 3,269, R-Square: 13.2**

Predictor Variable	Parameter Estimate	Standard Error
Intercept	1.5376	0.0787
Walk Score (Other: WalkScore)	0.0034	0.0004
CY Own Occ HUs Median Value (Claritas: HBHMEDHS)	0.0000	0.0000
Households in BG (NHTS: TotHHs)	-0.0000	0.0000
Percent Range CY Pop, Some Other Race (Claritas: HBPRCOTH)	0.0040	0.0007
Percent Persons in BG with Grad/Prof Degree (NHTS: PctGrad)	0.2270	0.0392
Average Worker Count of HHs in BG (NHTS: AvgWorkers)	-0.1152	0.0212
Percent Range CY Pop 20-64 (Claritas: HBP20TO64)	0.0061	0.0012
Percent Workers in BG (NHTS: PctWorkers)	0.1787	0.0449
Percent Weighted Trips in Destination Block Group for the Purpose Retail (Shopping/Meals/Services), All Modes (NHTS+: Retail_DestPct)	-0.1595	0.0460
Average Vehicle Count of HHs in BG (NHTS: AvgVehCount)	-0.0409	0.0123

**D.5 Estimated Walk Trips Per Person Including Transit Access/Egress
Stepwise-Selected Logistic Regression Model
Number of Observations Used: 6,911, Cox & Snell R-Square: 12.3**

Predictor Variable	Parameter Estimate	Standard Error
Percent Range CY Pop, White (Claritas: HBPRCCAU)	0.0142	0.0014
Households in BG (NHTS: TotHHs)	0.0002	0.0000
Average Worker Count of HHs in BG (NHTS: AvgWorkers)	0.5416	0.0717
Average Vehicle Count of HHs in BG (NHTS: AvgVehCount)	-0.2604	0.0348
Percent Range CY Pop, Hisp/Lat (Claritas: HBPHISP)	-0.0117	0.0016
Percent Weighted Trips in Destination Block Group for the Purpose Work & Related, All Modes (NHTS+: Work_DestPct)	-1.1283	0.1892
Percent Weighted Trips in Destination Block Group for the Purpose Retail (Shopping/Meals/Services), All Modes (NHTS+: Retail_DestPct)	-0.6516	0.1179
County-level Percentage of Adults Who Are Obese (Other: PctObese)	-0.0533	0.0098
County-level Gradation of percentage of physically inactive adults by total number of physically inactive adults in 2008 (Other: LIPERTOT08)	0.0804	0.0152
Average HH Size of HHs in BG (NHTS: AvgHHSIZE)	0.1697	0.0322
Percent Range CY Pop, Some Other Race (Claritas: HBPRCOTH)	0.0134	0.0030
Workers by Time Leaving Home To Go To Work, 8:30 a.m. to 8:59 a.m. (ACS: B08302e10)	-0.0027	0.0006
Persons by Sex By Educational Attainment For The Population 25 Years And Over, Male: Associate's degree (ACS: B15002e14)	0.0027	0.0006
Percent Workers in BG (NHTS: PctWorkers)	-0.5784	0.1328
Workers by Travel Time To Work, 15 to 19 minutes (ACS: B08303e5)	0.0013	0.0003
Workers by Means Of Transportation To Work, Worked at home (ACS: B08301e21)	0.0025	0.0006
Percent Persons in BG with Educ of HS or Less (NHTS: PctHSLess)	-0.3488	0.0858
Wtd Pct HHs in BG with English Interviews (NHTS: PctEngHH)	-0.4792	0.1184
Transit Score (Other: TransitScore)	0.0046	0.0012
Employee Range CY Employees per square mile (Claritas: HBEEMPDN)	0.0001	0.0000
Persons by Race, Two races excluding Some other race, and three or more races (ACS: B02001e10)	0.0018	0.0005
Intercept	-0.1674	0.3316

**D.6 Estimated Walk Trips Per Person Including Transit Access/Egress (When Positive)
Stepwise-Selected Least Squares Regression Model
Number of Observations Used: 3,512, R-Square: 27.0**

Predictor Variable	Parameter Estimate	Standard Error
Intercept	2.1614	0.0830
Average Vehicle Count of HHs in BG (NHTS: AvgVehCount)	-0.0908	0.0092
CY Own Occ HUs Median Value (Claritas: HBHMEDHS)	0.0000	0.0000
Households in BG (NHTS: TotHHs)	-0.0000	0.0000
Persons by Sex By Educational Attainment For The Population 25 Years And Over, Female: High school graduate, GED, or alternative (ACS: B15002e28)	-0.0004	0.0001
Walk Score (Other: WalkScore)	0.0024	0.0004
Percent Range CY Pop, Some Other Race (Claritas: HBPRCOTH)	0.0032	0.0005
Percent Range CY Pop 20-64 (Claritas: HBP20TO64)	0.0054	0.0009
Percent Persons in BG with Grad/Prof Degree (NHTS: PctGrad)	0.1531	0.0293
Persons by Race, Black or African American alone (ACS: B02001e3)	0.0001	0.0000
Percent Low Income Households (Claritas: HBHHINC1)	0.0034	0.0007
County-level Gradation of percentage of physically inactive adults by total number of physically inactive adults in 2008 (Other: LIPERTOT08)	-0.0197	0.0041
Workers by Means Of Transportation To Work By Travel Time To Work, Car, truck, or van - drove alone: 10 to 14 minutes (ACS: C08134e13)	-0.0003	0.0001
Percent Drivers in BG (NHTS: PctDrivers)	0.1221	0.0304
Household Income In The Past 12 Months, \$100,000 to \$124,999 (ACS: B19001e14)	0.0004	0.0001
Average Worker Count of HHs in BG (NHTS: AvgWorkers)	-0.0440	0.0111
Urban Indicator (Claritas: HBHUR_U)	0.0685	0.0178
Number of Land Uses (Other: NumberLandUseTypes)	-0.0248	0.0067
Persons by Sex By Educational Attainment For The Population 25 Years And Over, Male: 7th and 8th grade (ACS: B15002e6)	0.0011	0.0003
Percent Range CY Pop, Foreign born (Claritas: HBPFORBN)	0.0017	0.0005

**D.7 Estimated Walk Trips Per Person
Variables Selected by Stepwise Discriminant Analysis**

Predictor Variable
Percent Range CY Pop, White (Claritas: HBPRCCAU)
Percent Drivers in BG (NHTS: PctDrivers)
County-level Percentage of Adults Who Are Obese (Other: PctObese)
Households in BG (NHTS: TotHHs)
Average HH Size of HHs in BG (NHTS: AvgHHSize)
Percent Weighted Trips in Destination Block Group for the Purpose Work & Related, All Modes (NHTS+: Work_DestPct)
Percent Persons in BG with Educ of HS or Less (NHTS: PctHSLess)
Workers by Means Of Transportation To Work By Travel Time To Work, Car, truck, or van - drove alone: 15 to 19 minutes (ACS: C08134e14)
Average Worker Count of HHs in BG (NHTS: AvgWorkers)
Percent Weighted Trips in Destination Block Group for the Purpose Retail (Shopping/Meals/Services), All Modes (NHTS+: Retail_DestPct)
CY Avg HH Size (Claritas: HBHHSIZE)
Household Income In The Past 12 Months, \$150,000 to \$199,999 (ACS: B19001e16)
Persons by Race, Two races excluding Some other race, and three or more races (ACS: B02001e10)
County-level Gradation of percentage of physically inactive adults by total number of physically inactive adults in 2008 (Other: LIPERTOT08)
Workers by Time Leaving Home To Go To Work, 8:30 a.m. to 8:59 a.m. (ACS: B08302e10)
Wtd Pct HHs in BG Owning HH (NHTS: PctOwnHH)

**D.8 Estimated Walk Mileage Per Person
Variables Selected by Stepwise Discriminant Analysis**

Predictor Variable
Percent Range CY Pop, White (Claritas: HBPRCCAU)
Households in BG (NHTS: TotHHs)
County-level Percentage of Adults Who Are Obese (Other: PctObese)
Percent Drivers in BG (NHTS: PctDrivers)
Average HH Size of HHs in BG (NHTS: AvgHHSize)
Percent Persons in BG with Educ of HS or Less (NHTS: PctHSLess)
Persons by Race, Two races excluding Some other race, and three or more races (ACS: B02001e10)
Workers by Means Of Transportation To Work By Travel Time To Work, Car, truck, or van - drove alone: 15 to 19 minutes (ACS: C08134e14)
Percent Weighted Trips in Destination Block Group for the Purpose Work & Related, All Modes (NHTS+: Work_DestPct)
Average Worker Count of HHs in BG (NHTS: AvgWorkers)
Percent Weighted Trips in Destination Block Group for the Purpose Retail (Shopping/Meals/Services), All Modes (NHTS+: Retail_DestPct)
CY Avg HH Size (Claritas: HBHHSIZE)
County-level Gradation of percentage of physically inactive adults by total number of physically inactive adults in 2008 (Other: LIPERTOT08)

D.9 Estimated Walk Trips Per Person Including Transit Access/Egress Variables Selected by Stepwise Discriminant Analysis

Predictor Variable
Percent Range CY Pop, White (Claritas: HBPRCCAU)
Average Worker Count of HHs in BG (NHTS: AvgWorkers)
Percent Range CY Pop, Hisp/Lat (Claritas: HBPHISP)
Average Vehicle Count of HHs in BG (NHTS: AvgVehCount)
Households in BG (NHTS: TotHHs)
County-level Percentage of Adults Who Are Obese (Other: PctObese)
Average HH Size of HHs in BG (NHTS: AvgHHSize)
Percent Weighted Trips in Destination Block Group for the Purpose Work & Related, All Modes (NHTS+: Work_DestPct)
Persons by Sex By Educational Attainment For The Population 25 Years And Over, Male: Associate's degree (ACS: B15002e14)
Percent Weighted Trips in Destination Block Group for the Purpose Retail (Shopping/Meals/Services), All Modes (NHTS+: Retail_DestPct)
County-level Gradation of percentage of physically inactive adults by total number of physically inactive adults in 2008 (Other: LIPERTOT08)
Percent Range CY Pop, Some Other Race (Claritas: HBPRCOTH)
Percent Persons in BG with Educ of HS or Less (NHTS: PctHSLess)
Percent Workers in BG (NHTS: PctWorkers)
Workers by Travel Time To Work, 15 to 19 minutes (ACS: B08303e5)
Wtd Pct HHs in BG with English Interviews (NHTS: PctEngHH)
Transit Score (Other: TransitScore)
Wtd Pct HHs in BG Owning HH (NHTS: PctOwnHH)

APPENDIX E. STATISTICAL SUMMARY TABLES WITH THE NIELSEN EMPLOYMENT DATA

**E.1 Estimated Walk Trips Per Person
Stepwise-Selected Logistic Regression Model
Number of Observations Used: 6,650, Cox & Snell R-Square: 16.4**

Predictor Variable	Parameter Estimate	Standard Error
Retail Employees per Square Mile, County-level, Log Transformation (Nielsen: RETDEN_LOG)	-1.4741	0.1211
Households (000s) in BG (NHTS: TotHHsA)	0.3054	0.0303
County-level Gradation of percentage of physically inactive adults by total number of physically inactive adults in 2008 (Other: LIPERTOT08)	0.2515	0.0250
Average Driver Count of HHs in BG (NHTS: AvgDrivers)	0.4525	0.0568
Percent Weighted Trips in Destination Block Group for the Purpose Home-Residential, All Modes (NHTS+: Home_DestPct)	0.8906	0.1198
Percent Drivers in BG (NHTS: PctDrivers)	-0.8753	0.1491
Total Employees per Square Mile (000s), County-level (Nielsen: NIELDEN_CNTYA)	3.4171	0.6025
Number of Persons 25+ with Graduate or higher Education (ACS: ACSPctGrad)	1.8031	0.3398
Percent Persons in BG with Educ of HS or Less (NHTS: PctHSLess)	-0.9546	0.1849
Retail Employees as a Percent of Persons, County-level (Nielsen: RETPCT_CNTY)	73.4773	14.2714
Retail Employees per Square Mile, County-level (Nielsen: RETDEN_CNTY)	-0.0153	0.0032
Percent Weighted Trips in Destination Block Group for the Purpose Other, All Modes (NHTS+: Other_DestPct)	0.8336	0.1832
Percent Range CY Pop, Some Other Race (Claritas: HBPRCOTH)	0.0146	0.0032
Workers by Time Leaving Home To Go To Work, 8:30 a.m. to 8:59 a.m. (ACS: B08302e10)	-0.0027	0.0006
Percent Range CY Pop, Hisp/Lat (Claritas: HBPHISP)	-0.0080	0.0018
Intercept	12.0027	2.8863
Percent of Employees that are retail, County-level, Square Root Transformation (Nielsen: PCTRET_SQRT)	-23.4198	6.1454
Employees as a Percent of Persons, County-level (Nielsen: NIELPCT_CNTY)	-9.7307	2.6436
Average Vehicle Count of HHs in BG (NHTS: AvgVehCount)	-0.1458	0.0402
Wtd Pct HHs in BG with English Interviews (NHTS: PctEngHH)	-0.4901	0.1362

Predictor Variable	Parameter Estimate	Standard Error
Wtd Pct Hisp HHs in BG (NHTS: PctHispHH)	-0.3783	0.1068
Percent Persons in BG with Some College/Bachelor's (NHTS: PctCollege)	-0.6260	0.1907
Workers by Time Leaving Home To Go To Work, 6:30 a.m. to 6:59 a.m. (ACS: B08302e6)	0.0015	0.0005
Total Employees (000s), County-level (Nielsen: employ_cntyA)	0.0001	0.0000
Workers by Sex By Class Of Worker For The Civilian Employed Population 16 Years And Over, Female: State government workers (ACS: B24080e18)	-0.0029	0.0009
Employee Range (000s) CY Employees per square mile (Claritas: HBEEMPDNA)	0.0635	0.0212
Household Income In The Past 12 Months, \$150,000 to \$199,999 (ACS: B19001e16)	0.0021	0.0007
Persons by Race, Two races excluding Some other race, and three or more races (ACS: B02001e10)	0.0015	0.0006
Walk Score (Other: WalkScore)	0.0046	0.0018
Persons (000s) by Race, Black or African American alone (ACS: B02001e3A)	-0.2400	0.0924
Workers by Means Of Transportation To Work, Worked at home (ACS: B08301e21)	0.0017	0.0007
Percent Range CY Pop, White (Claritas: HBPRCCAU)	0.0045	0.0019
Wtd Pct White HHs in BG (NHTS: PctWhiteHH)	0.1921	0.0810
Percent Persons in BG with Grad/Prof Degree (NHTS: PctGrad)	-0.4824	0.2114
Household Income In The Past 12 Months, \$100,000 to \$124,999 (000s of HHs) (ACS: B19001e14A)	1.4397	0.6581
Household Income In The Past 12 Months, \$75,000 to \$99,999 (000s of HHs) (ACS: B19001e13A)	1.0312	0.4867
Percent Weighted Trips in Destination Block Group for the Purpose Work & Related, All Modes (NHTS+: Work_DestPct)	-0.4720	0.2243
Workers (000s)by Means Of Transportation To Work By Travel Time To Work, Taxicab, motorcycle, bicycle, walked, or other means: 10 to 14 minutes (ACS: C08134e43A)	3.2634	1.6066

**E.2 Estimated Walk Trips Per Person (When Positive)
Stepwise-Selected Least Squares Regression Model
Number of Observations Used: 3,190, R-Square: 31.4**

Predictor Variable	Parameter Estimate	Standard Error
Retail Employees per Square Mile, County-level, Square Root Transformation (Nielsen: RETDEN_SQRT)	0.0736	0.0058
Households (000s) in BG (NHTS: TotHHsA)	-0.0793	0.0063
Intercept	5.9466	0.6626
Total Employees per Square Mile (000s), County-level (Nielsen: NIELDEN_CNTYA)	-0.2931	0.0336
County-level Gradation of percentage of physically inactive adults by total number of physically inactive adults in 2008 (Other: LIPERTOT08)	-0.0406	0.0051
Persons by Sex & Age, Total, Log Transformation (ACS: ACSPOP_LOG)	-0.1385	0.0181
Retail Employees as a Percent of Persons, County-level (Nielsen: RETPCT_CNTY)	-4.5451	0.7446
Number of Persons 25+ with College Education (ACS: ACSPctCollege)	-0.5536	0.0934
Average Vehicle Count of HHs in BG (NHTS: AvgVehCount)	-0.0549	0.0094
Average HH Size of HHs in BG (NHTS: AvgHHSIZE)	-0.0396	0.0076
Workers (000s) by Sex By Class Of Worker For The Civilian Employed Population 16 Years And Over, Male: Employee of private company workers (ACS: B24080e4A)	0.1819	0.0358
Number of Persons 25+ with High School or Less Education (ACS: ACSPctHSLess)	-0.4039	0.0798
Walk Score (Other: WalkScore)	0.0021	0.0004
Percent Weighted Trips in Destination Block Group for the Purpose Other, All Modes (NHTS+: Other_DestPct)	0.2112	0.0454
Percent Persons in BG with Grad/Prof Degree (NHTS: PctGrad)	0.1624	0.0356
Percent Range CY Pop 20-64 (Claritas: HBP20TO64)	0.0048	0.0011
County-level Percentage of Adults Who Are Obese (Other: PctObese)	-0.0155	0.0036
Wtd Pct HHs in BG with English Interviews (NHTS: PctEngHH)	-0.1187	0.0325
Wtd Pct HHs in BG Owning HH (NHTS: PctOwnHH)	-0.0750	0.0225
Percent Low Income Households (Claritas: HBHHINC1)	0.0025	0.0008
Median Age, Total, Log Transformation (ACS: ACSAGE_LOG)	-1.1257	0.4058

Predictor Variable	Parameter Estimate	Standard Error
Household Income In The Past 12 Months, \$150,000 to \$199,999 (ACS: B19001e16)	-0.0004	0.0001
CY Avg HH Size (Claritas: HBHHSIZE)	0.0446	0.0164
Workers by Sex By Class Of Worker For The Civilian Employed Population 16 Years And Over, Male: Private not-for-profit wage and salary workers (ACS: B24080e6)	0.0006	0.0002
Household Income In The Past 12 Months, \$100,000 to \$124,999 (000s of HHs) (ACS: B19001e14A)	0.3416	0.1323
Median Age, Total, Square Root Transformation (ACS: ACSAGE_SQRT)	0.3460	0.1341
Persons (000s) by Race, Two races including Some other race (ACS: B02001e9A)	-0.5271	0.2076
Population Density (Persons per Square Mile, in 000s) (ACS: ACSDENSITYA)	0.0026	0.0011
Workers (000s) by Means Of Transportation To Work By Travel Time To Work, Taxicab, motorcycle, bicycle, walked, or other means: 10 to 14 minutes (ACS: C08134e43A)	-0.9051	0.3981
Workers (000s) by Sex By Class Of Worker For The Civilian Employed Population 16 Years And Over, Female: Local government workers (ACS: B24080e17A)	-0.3765	0.1719
Percent Persons in BG with Some College/Bachelor's (NHTS: PctCollege)	0.0607	0.0278
Wtd Pct Afr Am. HHs in BG (NHTS: PctAfAmHH)	0.0678	0.0323
Workers by Sex By Class Of Worker For The Civilian Employed Population 16 Years And Over, Male: Unpaid family workers (ACS: B24080e11)	0.0032	0.0015

**E.3 Estimated Walk Mileage Per Person
Stepwise-Selected Logistic Regression Model
Number of Observations Used: 6,650, Cox & Snell R-Square: 15.9**

Predictor Variable	Parameter Estimate	Standard Error
Total Employees per Square Mile, County-level, Square Root Transformation (Nielsen: NIELDEN_SQRT)	-0.1696	0.0148
Households (000s) in BG (NHTS: TotHHsA)	0.3038	0.0298
County-level Gradation of percentage of physically inactive adults by total number of physically inactive adults in 2008 (Other: LIPERTOT08)	0.2341	0.0243
Percent Drivers in BG (NHTS: PctDrivers)	-1.0970	0.1199
Average Driver Count of HHs in BG (NHTS: AvgDrivers)	0.4889	0.0545
Percent Weighted Trips in Destination Block Group for the Purpose Home-Residential, All Modes (NHTS+: Home_DestPct)	0.9524	0.1094
Total Employees per Square Mile (000s), County-level (Nielsen: NIELDEN_CNTYA)	4.2348	0.5975
Percent Range CY Pop, Some Other Race (Claritas: HBPRCOTH)	0.0171	0.0033
Number of Persons 25+ with High School or Less Education (ACS: ACSPctHSLess)	-1.7326	0.3343
Percent Persons in BG with Educ of HS or Less (NHTS: PctHSLess)	-0.4605	0.0925
Percent Weighted Trips in Destination Block Group for the Purpose Other, All Modes (NHTS+: Other_DestPct)	0.8368	0.1756
Retail Employees per Square Mile, County-level (Nielsen: RETDEN_CNTY)	-0.0138	0.0030
Percent Range CY Pop, Hisp/Lat (Claritas: HBPHISP)	-0.0085	0.0019
Retail Employees as a Percent of Persons, County-level (Nielsen: RETPCT_CNTY)	67.7285	15.6850
Workers by Time Leaving Home To Go To Work, 8:30 a.m. to 8:59 a.m. (ACS: B08302e10)	-0.0026	0.0006
Percent of Employees that are retail, County-level, Log Transformation (Nielsen: PCTRET_LOG)	-6.6062	1.6320
Intercept	-10.4186	2.6500
Number of Persons 25+ with College Education (ACS: ACSPctCollege)	-1.3908	0.3884
Wtd Pct Hisp HHs in BG (NHTS: PctHispHH)	-0.3662	0.1066
Average Vehicle Count of HHs in BG (NHTS: AvgVehCount)	-0.1367	0.0401
Wtd Pct HHs in BG with English Interviews (NHTS: PctEngHH)	-0.4432	0.1363
Employees as a Percent of Persons, County-level (Nielsen: NIELPCT_CNTY)	-9.6141	3.0476

Predictor Variable	Parameter Estimate	Standard Error
Workers by Time Leaving Home To Go To Work, 6:30 a.m. to 6:59 a.m. (ACS: B08302e6)	0.0015	0.0005
Persons by Race, Two races excluding Some other race, and three or more races (ACS: B02001e10)	0.0017	0.0005
Workers by Sex By Class Of Worker For The Civilian Employed Population 16 Years And Over, Female: State government workers (ACS: B24080e18)	-0.0027	0.0009
Percent Range CY Pop, White (Claritas: HBPRCCAU)	0.0056	0.0019
Employee Range (000s) CY Employees per square mile (Claritas: HBEEMPDNA)	0.0602	0.0211
Walk Score (Other: WalkScore)	0.0050	0.0018
Household Income In The Past 12 Months, \$100,000 to \$124,999 (000s of HHs) (ACS: B19001e14A)	1.5836	0.5906
Workers by Means Of Transportation To Work, Worked at home (ACS: B08301e21)	0.0017	0.0007
Wtd Pct White HHs in BG (NHTS: PctWhiteHH)	0.2012	0.0808
Household Income In The Past 12 Months, \$150,000 to \$199,999 (ACS: B19001e16)	0.0017	0.0007
Total Employees (000s), County-level (Nielsen: employ_cntyA)	0.0001	0.0000
Percent Range CY Pop 20-64 (Claritas: HBP20TO64)	-0.0089	0.0040
Persons (000s) by Race, Black or African American alone (ACS: B02001e3A)	-0.1996	0.0910
Workers by Travel Time To Work, 15 to 19 minutes (ACS: B08303e5)	0.0007	0.0004

**E.4 Estimated Walk Mileage Per Person (When Positive)
Stepwise-Selected Least Squares Regression Model
Number of Observations Used: 3,144, R-Square: 18.3**

Predictor Variable	Parameter Estimate	Standard Error
Intercept	2.7613	0.2902
Households (000s) in BG (NHTS: TotHHsA)	-0.0658	0.0079
Retail Employees per Square Mile, County-level, Log Transformation (Nielsen: RETDEN_LOG)	0.2406	0.0307
County-level Gradation of percentage of physically inactive adults by total number of physically inactive adults in 2008 (Other: LIPERTOT08)	-0.0486	0.0070
Retail Employees as a Percent of Persons, County-level (Nielsen: RETPCT_CNTY)	-5.6530	1.0393
Walk Score (Other: WalkScore)	0.0023	0.0005
Percent Persons in BG with Grad/Prof Degree (NHTS: PctGrad)	0.2268	0.0475
Number of Persons 25+ with College Education (ACS: ACSPctCollege)	-0.5670	0.1286
Persons by Sex & Age, Total, Log Transformation (ACS: ACSPOP_LOG)	-0.0875	0.0230
Average Vehicle Count of HHs in BG (NHTS: AvgVehCount)	-0.0440	0.0121
Household Income In The Past 12 Months, \$100,000 to \$124,999 (000s of HHs) (ACS: B19001e14A)	0.4695	0.1306
Percent Range CY Pop 20-64 (Claritas: HBP20TO64)	0.0045	0.0013
Total Employees per Square Mile (000s), County-level (Nielsen: NIELDEN_CNTYA)	-0.3959	0.1146
CY Own Occ HUs Median Value (000s) (Claritas: HBHMEDHSA)	0.0002	0.0001
Average HH Size of HHs in BG (NHTS: AvgHHSIZE)	-0.0322	0.0100
Percent Weighted Trips in Destination Block Group for the Purpose Other, All Modes (NHTS+: Other_DestPct)	0.1797	0.0614
Total enrollment of K-12 Public schools in the blockgroup (000s) (DWilson: PUB_ENROLLA)	-0.0353	0.0125
Number of Males (000s), 22-34 years old (ACS: MALE22_34A)	0.1578	0.0561
Wtd Pct HHs in BG with English Interviews (NHTS: PctEngHH)	-0.1169	0.0430
Number of Persons 25+ with High School or Less Education (ACS: ACSPctHSLess)	-0.3030	0.1117
Percent Persons in BG with Some College/Bachelor's (NHTS: PctCollege)	0.1001	0.0370
Percent Range CY Families Below Poverty (Claritas: HBPLTPOV)	0.0030	0.0012
Number of Land Uses (Other: NumberLandUseTypes)	-0.0234	0.0093

Predictor Variable	Parameter Estimate	Standard Error
Retail Employees per Square Mile, County-level (Nielsen: RETDEN_CNTY)	0.0016	0.0007
Wtd Pct Asian Am. HHs in BG (NHTS: PctAsAmHH)	-0.1084	0.0521
Percent Range CY Pop, Hisp/Lat (Claritas: HBPHISP)	0.0011	0.0005

**E.5 Estimated Walk Trips Per Person Including Transit Access/Egress
Stepwise-Selected Logistic Regression Model
Number of Observations Used: 6,650, Cox & Snell R-Square: 16.6**

Predictor Variable	Parameter Estimate	Standard Error
Retail Employees per Square Mile, County-level, Square Root Transformation (Nielsen: RETDEN_SQRT)	-0.3586	0.0248
Total Employees per Square Mile (000s), County-level (Nielsen: NIELDEN_CNTYA)	1.5561	0.1460
County-level Gradation of percentage of physically inactive adults by total number of physically inactive adults in 2008 (Other: LIPERTOT08)	0.2251	0.0214
Average Driver Count of HHs in BG (NHTS: AvgDrivers)	0.6014	0.0584
Households (000s) in BG (NHTS: TotHHsA)	0.3138	0.0315
Percent Weighted Trips in Destination Block Group for the Purpose Home-Residential, All Modes (NHTS+: Home_DestPct)	0.9495	0.1188
Average Vehicle Count of HHs in BG (NHTS: AvgVehCount)	-0.2851	0.0414
Percent Persons in BG with Educ of HS or Less (NHTS: PctHSLess)	-0.9597	0.1867
Retail Employees as a Percent of Persons, County-level (Nielsen: RETPCT_CNTY)	11.8942	2.4290
Percent Drivers in BG (NHTS: PctDrivers)	-0.7130	0.1481
Household Language By Linguistic Isolation, English only (000s of HHs) (ACS: B16002e2A)	0.6623	0.1432
Percent Weighted Trips in Destination Block Group for the Purpose Other, All Modes (NHTS+: Other_DestPct)	0.7827	0.1819
Persons (000s) by Race, Black or African American alone (ACS: B02001e3A)	-0.3642	0.0888
Wtd Pct HHs in BG with English Interviews (NHTS: PctEngHH)	-0.5555	0.1365
Workers by Sex By Class Of Worker For The Civilian Employed Population 16 Years And Over, Female: Private for-profit wage and salary workers (ACS: B24080e13)	-0.0009	0.0002
Workers by Time Leaving Home To Go To Work, 8:30 a.m. to 8:59 a.m. (ACS: B08302e10)	-0.0024	0.0006
Wtd Pct Hisp HHs in BG (NHTS: PctHispHH)	-0.4040	0.1027
Percent Persons in BG with Some College/Bachelor's (NHTS: PctCollege)	-0.7256	0.1927
Workers by Sex By Class Of Worker For The Civilian Employed Population 16 Years And Over, Female: State government workers (ACS: B24080e18)	-0.0036	0.0010
Percent of Employees that are retail, County-level (Nielsen: PCTRETAIL_CNTY)	-5.3174	1.4676

Predictor Variable	Parameter Estimate	Standard Error
CY Own Occ HUs Median Value (000s) (Claritas: HBHMEDHSA)	0.0006	0.0002
Number of Persons 25+ with Graduate or higher Education (ACS: ACSPctGrad)	1.1618	0.3379
Employee Range (000s) CY Employees per square mile (Claritas: HBEEMPDNA)	0.0723	0.0213
Workers by Time Leaving Home To Go To Work, 6:30 a.m. to 6:59 a.m. (ACS: B08302e6)	0.0015	0.0005
Workers by Means Of Transportation To Work, Worked at home (ACS: B08301e21)	0.0021	0.0007
Workers by Means Of Transportation To Work, Public transportation (excluding taxicab): Bus or trolley bus (ACS: B08301e11)	0.0016	0.0006
Percent Persons in BG with Grad/Prof Degree (NHTS: PctGrad)	-0.6052	0.2129
Household Income In The Past 12 Months, \$100,000 to \$124,999 (000s of HHs) (ACS: B19001e14A)	1.7831	0.6679
Number of Males, 0-14 years old (ACS: MALE0_14)	0.0006	0.0002
Walk Score (Other: WalkScore)	0.0044	0.0017
Total enrollment of colleges in the blockgroup (000s) (DWilson: COL_ENROLLA)	0.0374	0.0161
Percent Weighted Trips in Destination Block Group for the Purpose Work & Related, All Modes (NHTS+: Work_DestPct)	-0.4678	0.2197
Household Income In The Past 12 Months, \$75,000 to \$99,999 (000s of HHs) (ACS: B19001e13A)	1.1619	0.5565
Wtd Pct White HHs in BG (NHTS: PctWhiteHH)	0.1546	0.0766
Intercept	0.7008	0.4971

**E.6 Estimated Walk Trips Per Person Including Transit Access/Egress (When Positive)
Stepwise-Selected Least Squares Regression Model
Number of Observations Used: 3,385, R-Square: 32.9**

Predictor Variable	Parameter Estimate	Standard Error
Intercept	4.2636	0.2008
Retail Employees per Square Mile, County-level, Square Root Transformation (Nielsen: RETDEN_SQRT)	0.0694	0.0055
Households (000s) in BG (NHTS: TotHHsA)	-0.0678	0.0061
Total Employees per Square Mile (000s), County-level (Nielsen: NIELDEN_CNTYA)	-0.2732	0.0321
County-level Gradation of percentage of physically inactive adults by total number of physically inactive adults in 2008 (Other: LIPERTOT08)	-0.0396	0.0049
Average Vehicle Count of HHs in BG (NHTS: AvgVehCount)	-0.0748	0.0105
Number of Persons 25+ with College Education (ACS: ACSPctCollege)	-0.5435	0.0914
Persons by Sex & Age, Total, Log Transformation (ACS: ACSPOP_LOG)	-0.1118	0.0189
Retail Employees as a Percent of Persons, County-level (Nielsen: RETPCT_CNTY)	-4.1600	0.7271
Percent Drivers in BG (NHTS: PctDrivers)	0.1682	0.0313
County-level Percentage of Adults Who Are Obese (Other: PctObese)	-0.0186	0.0035
Walk Score (Other: WalkScore)	0.0020	0.0004
Workers by Sex By Class Of Worker For The Civilian Employed Population 16 Years And Over, Male (ACS: B24080e2)	0.0002	0.0000
Number of Persons 25+ with High School or Less Education (ACS: ACSPctHSLess)	-0.3399	0.0724
Wtd Pct HHs in BG Owning HH (NHTS: PctOwnHH)	-0.0942	0.0214
Percent Weighted Trips in Destination Block Group for the Purpose Other, All Modes (NHTS+: Other_DestPct)	0.1842	0.0442
Wtd Pct HHs in BG with English Interviews (NHTS: PctEngHH)	-0.1210	0.0292
Household Income In The Past 12 Months, \$150,000 to \$199,999 (ACS: B19001e16)	-0.0005	0.0001
Percent Low Income Households (Claritas: HBHHINC1)	0.0026	0.0007
Average Driver Count of HHs in BG (NHTS: AvgDrivers)	-0.0451	0.0131
Wtd Pct Afr Am. HHs in BG (NHTS: PctAfAmHH)	0.1005	0.0303
Percent Persons in BG with Grad/Prof Degree (NHTS: PctGrad)	0.0946	0.0296
Percent Range CY Pop 20-64 (Claritas: HBP20TO64)	0.0029	0.0010
Median Age, Female (ACS: B01002e3)	-0.0024	0.0009

Predictor Variable	Parameter Estimate	Standard Error
Number of Land Uses (Other: NumberLandUseTypes)	-0.0185	0.0067
Persons (000s) by Race, Two races including Some other race (ACS: B02001e9A)	-0.5182	0.1971
Household Income In The Past 12 Months, \$100,000 to \$124,999 (000s of HHs) (ACS: B19001e14A)	0.3577	0.1373
Household Income In The Past 12 Months, \$75,000 to \$99,999 (000s of HHs) (ACS: B19001e13A)	-0.2702	0.1136
Transit Score (Other: TransitScore)	0.0007	0.0003
Workers by Sex By Class Of Worker For The Civilian Employed Population 16 Years And Over, Male: Private not-for-profit wage and salary workers (ACS: B24080e6)	0.0005	0.0002
Workers by Travel Time To Work, 10 to 14 minutes (ACS: B08303e4)	-0.0001	0.0001

**E.7 Estimated Walk Trips Per Person
Variables Selected by Stepwise Discriminant Analysis**

Predictor Variable
Households (000s) in BG (NHTS: TotHHsA)
Percent Weighted Trips in Destination Block Group for the Purpose Home-Residential, All Modes (NHTS+: Home_DestPct)
County-level Gradation of percentage of physically inactive adults by total number of physically inactive adults in 2008 (Other: LIPERTOT08)
Average Driver Count of HHs in BG (NHTS: AvgDrivers)
Percent of Employees that are retail, County-level, Log Transformation (Nielsen: PCTRET_LOG)
Percent of Employees that are retail, County-level, Square Root Transformation (Nielsen: PCTRET_SQRT)
Percent of Employees that are retail, County-level (Nielsen: PCTRETAIL_CNTY)
Percent Drivers in BG (NHTS: PctDrivers)
Percent Persons in BG with Educ of HS or Less (NHTS: PctHSLess)
Total Employees per Square Mile (000s), County-level (Nielsen: NIELDEN_CNTYA)
Percent Weighted Trips in Destination Block Group for the Purpose Other, All Modes (NHTS+: Other_DestPct)
Percent Range CY Pop, Some Other Race (Claritas: HBPRCOTH)
Number of Persons 25+ with High School or Less Education (ACS: ACSPctHSLess)
Retail Employees per Square Mile, County-level (Nielsen: RETDEN_CNTY)
Total Employees per Square Mile, County-level, Square Root Transformation (Nielsen: NIELDEN_SQRT)
Average Vehicle Count of HHs in BG (NHTS: AvgVehCount)
Percent Range CY Pop, Hisp/Lat (Claritas: HBPHISP)
Persons by Sex & Age, Total, Log Transformation (ACS: ACSPOP_LOG)
Percent Persons in BG with Some College/Bachelor's (NHTS: PctCollege)
Wtd Pct HHs in BG with English Interviews (NHTS: PctEngHH)
Retail Employees per Square Mile, County-level, Square Root Transformation (Nielsen: RETDEN_SQRT)
Retail Employees as a Percent of Persons, County-level (Nielsen: RETPCT_CNTY)
Wtd Pct Hisp HHs in BG (NHTS: PctHispHH)
Workers by Time Leaving Home To Go To Work, 8:30 a.m. to 8:59 a.m. (ACS: B08302e10)
Employee Range (000s) CY Employees per square mile (Claritas: HBEEMPDNA)
Number of Persons 25+ with College Education (ACS: ACSPctCollege)
Workers by Time Leaving Home To Go To Work, 6:30 a.m. to 6:59 a.m. (ACS: B08302e6)
Number of Females, 22-34 years old (ACS: FEM22_34)
Walk Score (Other: WalkScore)

Predictor Variable
Percent Persons in BG with Grad/Prof Degree (NHTS: PctGrad)
Retail Employees per Square Mile, County-level, Log Transformation (Nielsen: RETDEN_LOG)
Persons by Race, Two races excluding Some other race, and three or more races (ACS: B02001e10)
Persons (000s) by Race, Black or African American alone (ACS: B02001e3A)
Household Income In The Past 12 Months, \$75,000 to \$99,999 (000s of HHs) (ACS: B19001e13A)
CY Own Occ HUs Median Value (000s) (Claritas: HBHMEDHSA)
Percent Range CY Pop, White (Claritas: HBPRCCAU)
Wtd Pct White HHs in BG (NHTS: PctWhiteHH)
Household Income In The Past 12 Months, \$100,000 to \$124,999 (000s of HHs) (ACS: B19001e14A)

**E. 8 Estimated Walk Mileage Per Person
Variables Selected by Stepwise Discriminant Analysis**

Predictor Variable
County-level Gradation of percentage of physically inactive adults by total number of physically inactive adults in 2008 (Other: LIPERTOT08)
Households (000s) in BG (NHTS: TotHHsA)
Average Driver Count of HHs in BG (NHTS: AvgDrivers)
Retail Employees per Square Mile, County-level, Log Transformation (Nielsen: RETDEN_LOG)
Percent Weighted Trips in Destination Block Group for the Purpose Home-Residential, All Modes (NHTS+: Home_DestPct)
Total Employees per Square Mile, County-level, Log Transformation (Nielsen: NIELDEN_LOG)
Percent Drivers in BG (NHTS: PctDrivers)
Percent Persons in BG with Educ of HS or Less (NHTS: PctHSLess)
Total Employees per Square Mile (000s), County-level (Nielsen: NIELDEN_CNTYA)
Persons (000s) by Race, Black or African American alone (ACS: B02001e3A)
Percent Range CY Pop, Some Other Race (Claritas: HBPRCOTH)
Employee Range (000s) CY Employees per square mile (Claritas: HBEEMPDNA)
Number of Persons 25+ with Graduate or higher Education (ACS: ACSPctGrad)
Percent Weighted Trips in Destination Block Group for the Purpose Other, All Modes (NHTS+: Other_DestPct)
Percent Range CY Pop, Hisp/Lat (Claritas: HBPHISP)
Total Employees per Square Mile, County-level, Square Root Transformation (Nielsen: NIELDEN_SQRT)
Percent Persons in BG with Some College/Bachelor's (NHTS: PctCollege)
Average Vehicle Count of HHs in BG (NHTS: AvgVehCount)
Workers by Time Leaving Home To Go To Work, 8:30 a.m. to 8:59 a.m. (ACS: B08302e10)
Wtd Pct Hisp HHs in BG (NHTS: PctHispHH)
Wtd Pct HHs in BG with English Interviews (NHTS: PctEngHH)
Workers by Time Leaving Home To Go To Work, 6:30 a.m. to 6:59 a.m. (ACS: B08302e6)
Persons by Sex & Age, Total, Log Transformation (ACS: ACSPOP_LOG)
Household Language By Linguistic Isolation, English only (000s of HHs) (ACS: B16002e2A)
Percent Persons in BG with Grad/Prof Degree (NHTS: PctGrad)
Persons by Race, Two races excluding Some other race, and three or more races (ACS: B02001e10)
Percent Range CY Pop, Asian (Claritas: HBPRCASN)
Wtd Pct White HHs in BG (NHTS: PctWhiteHH)

Predictor Variable
Workers by Sex By Class Of Worker For The Civilian Employed Population 16 Years And Over, Female: State government workers (ACS: B24080e18)
Number of Persons 25+ with High School or Less Education (ACS: ACSPctHSLess)
Percent Weighted Trips in Destination Block Group for the Purpose Work & Related, All Modes (NHTS+: Work_DestPct)

**E. 9 Estimated Walk Trips Per Person Including Transit Access/Egress
Variables Selected by Stepwise Discriminant Analysis**

Predictor Variable
Average Driver Count of HHs in BG (NHTS: AvgDrivers)
Households (000s) in BG (NHTS: TotHHsA)
Percent Weighted Trips in Destination Block Group for the Purpose Home-Residential, All Modes (NHTS+: Home_DestPct)
County-level Gradation of percentage of physically inactive adults by total number of physically inactive adults in 2008 (Other: LIPERTOT08)
Average Vehicle Count of HHs in BG (NHTS: AvgVehCount)
Percent of Employees that are retail, County-level, Log Transformation (Nielsen: PCTRET_LOG)
Percent of Employees that are retail, County-level, Square Root Transformation (Nielsen: PCTRET_SQRT)
Percent of Employees that are retail, County-level (Nielsen: PCTRETAIL_CNTY)
Percent Persons in BG with Educ of HS or Less (NHTS: PctHSLess)
Total Employees per Square Mile (000s), County-level (Nielsen: NIELDEN_CNTYA)
Percent Drivers in BG (NHTS: PctDrivers)
Household Language By Linguistic Isolation, English only (000s of HHs) (ACS: B16002e2A)
Wtd Pct HHs in BG with English Interviews (NHTS: PctEngHH)
Percent Weighted Trips in Destination Block Group for the Purpose Other, All Modes (NHTS+: Other_DestPct)
Persons (000s) by Race, Black or African American alone (ACS: B02001e3A)
Percent Persons in BG with Some College/Bachelor's (NHTS: PctCollege)
Employee Range (000s) CY Employees per square mile (Claritas: HBEEMPDNA)
Retail Employees per Square Mile, County-level (Nielsen: RETDEN_CNTY)
Persons by Sex & Age, Total, Log Transformation (ACS: ACSPOP_LOG)
Total Employees per Square Mile, County-level, Square Root Transformation (Nielsen: NIELDEN_SQRT)
Workers by Time Leaving Home To Go To Work, 8:30 a.m. to 8:59 a.m. (ACS: B08302e10)
Number of Persons 25+ with Graduate or higher Education (ACS: ACSPctGrad)
Wtd Pct Hisp HHs in BG (NHTS: PctHispHH)
Workers by Time Leaving Home To Go To Work, 6:30 a.m. to 6:59 a.m. (ACS: B08302e6)
Retail Employees per Square Mile, County-level, Square Root Transformation (Nielsen: RETDEN_SQRT)
CY Own Occ HUs Median Value (000s) (Claritas: HBHMEDHSA)
Percent Persons in BG with Grad/Prof Degree (NHTS: PctGrad)
Retail Employees as a Percent of Persons, County-level (Nielsen: RETPCT_CNTY)

Predictor Variable
Workers by Means Of Transportation To Work, Public transportation (excluding taxicab): Bus or trolley bus (ACS: B08301e11)
Households with Income \$25-49,999 (ACS: INC25_50K)
Workers by Sex By Class Of Worker For The Civilian Employed Population 16 Years And Over, Female: State government workers (ACS: B24080e18)
Wtd Pct White HHs in BG (NHTS: PctWhiteHH)
Walk Score (Other: WalkScore)
Percent Range CY Pop, Foreign born (Claritas: HBPFORBN)
Percent Weighted Trips in Destination Block Group for the Purpose Work & Related, All Modes (NHTS+: Work_DestPct)

APPENDIX F. STATISTICAL SUMMARY TABLES WITH INTERACTION TERMS INCORPORATED

Stepwise Logistic procedure was carried out with the NMT data set, which includes a total number of 13,927 block-group records. The procedure used 7,086 non-missing records in the estimation process. Among those records (7,806), 3,690 block groups reported no walking activities (i.e., no walk trips) and the remaining 3396 block groups reported with at least one walk trip. This process was used to identify major factors that can be used in the discriminant analysis.

The LOGISTIC Procedure

Number of Observations Read	13927
Number of Observations Used	7086

Response Profile		
Ordered Value	Wt_WkTrpPC_I	Total Frequency
1	0	3690
2	1	3396

Note 6841 observations were deleted due to missing values for the response or explanatory variables.

R-Square	0.2129	Max-rescaled R-Square	0.2840
-----------------	--------	------------------------------	--------

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	1696.5154	88	<.0001
Score	1483.0578	88	<.0001
Wald	1187.4833	88	<.0001

The following table provides a list of selected variables that are significant at 0.5 levels.

Parameter	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	3.1116	1.2455	6.2413	0.0125
Home_Dest*Other_Dest	8.2510	0.7074	136.0622	<.0001
AvgDrivers	0.8444	0.1074	61.7693	<.0001
TotHHs	0.000381	0.000106	12.9188	0.0003
PctCollege*PctGrad	2.6649	0.4574	33.9375	<.0001
PctDrivers	-1.4999	0.3216	21.7451	<.0001
B08301e21*PctWhiteHH	0.00669	0.00157	18.2216	<.0001
PctCollege	-1.4797	0.3046	23.5970	<.0001
PctGrad*ACSPctGrad	3.0870	0.8347	13.6769	0.0002
Home_DestPct*TotHHs	-0.00034	0.000105	10.5033	0.0012
HBEEMPDN*EMPLOYMENT	-2.55E-8	1.311E-8	3.7926	0.0515
PctCollege*PctHSLess	2.1628	0.4113	27.6580	<.0001
B24080e18	-0.0138	0.00230	35.9293	<.0001
Home_Dest*Work_DestP	3.9853	1.0122	15.5019	<.0001
RET_DEN	0.000093	0.000047	3.9335	0.0473
RETDEN_LOG	-0.5957	0.0548	118.1711	<.0001
TotHHs*AvgVehCount	0.000119	0.000033	12.7962	0.0003
AvgVehCoun*B19001e14	0.00263	0.000716	13.4485	0.0002
PctHispHH	-2.5565	0.6566	15.1576	<.0001
Work_DestPct	-0.9306	0.3924	5.6244	0.0177
Home_DestP*PctHispHH	0.6603	0.2557	6.6692	0.0098
PctEngHH	-2.4153	0.6415	14.1768	0.0002
RET_DEN*NIEL_DEN	-1.65E-9	8.57E-10	3.6908	0.0547
AvgDrivers*B19001e14	-0.00321	0.000930	11.9036	0.0006
ACSPctGrad	1.2590	0.3936	10.2307	0.0014
PctHispHH*PctEngHH	2.2415	0.6026	13.8356	0.0002
B08301e21*B08302e6	-2.57E-6	8.749E-7	8.6585	0.0033
PctHSLess*PctEngHH	-1.1747	0.3162	13.8023	0.0002
AvgVehCount	-0.2802	0.1063	6.9445	0.0084
HBPRCOTH	0.0198	0.00459	18.5904	<.0001

Parameter	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
B24080e18*LIPERTOT08	0.00161	0.000379	18.0640	<.0001
AvgDrivers*PctHispHH	-0.2424	0.1014	5.7210	0.0168
PctCollege*HBPRCCA	0.00991	0.00344	8.3103	0.0039
Other_DestP*B02001e3	-0.00148	0.000501	8.7281	0.0031
Other_Dest*B08302e10	-0.00707	0.00310	5.1943	0.0227
HBPHISP*B08302e10	0.000060	0.000022	7.3155	0.0068
Other_Des*LIPERTOT08	-0.2598	0.0891	8.4957	0.0036
AvgDriver*AvgVehCoun	-0.0885	0.0287	9.5514	0.0020
B19001e14	0.00455	0.00169	7.2680	0.0070
PCTRET_SQR*WalkScore	-0.2138	0.0535	15.9386	<.0001
B24080e18*B08302e6	0.000020	6.689E-6	9.0071	0.0027
PctWhiteHH*HBPRCCA	-0.00749	0.00307	5.9471	0.0147
B08302e6*LIPERTOT08	0.000430	0.000204	4.4382	0.0351
PctDrivers*B08301e21	-0.00588	0.00233	6.3430	0.0118
TotHHs*Work_DestPct	-0.00076	0.000269	8.0154	0.0046
PctDrivers*PctEngHH	0.8974	0.3360	7.1330	0.0076
WalkScore	0.1121	0.0244	21.1725	<.0001
PctGrad*PctHSLess	2.5808	0.8581	9.0453	0.0026
AvgVehCount*HBPRCOTH	-0.00512	0.00188	7.4314	0.0064
PctEngHH*C08134e42	0.0112	0.00448	6.2277	0.0126
B08302e10	-0.00285	0.000878	10.5645	0.0012
Home_DestP*B19001e14	-0.00317	0.00163	3.7651	0.0523
EMPLOYMENT	0.000452	0.000150	9.1144	0.0025
RETDEN_LO*EMPLOYMENT	-0.00006	0.000030	3.9967	0.0456
C08134e42	-0.00935	0.00426	4.8089	0.0283
Home_DestPct	0.6544	0.1847	12.5460	0.0004
PctGrad	-1.5878	0.3029	27.4739	<.0001
LIPERTOT08	0.1530	0.0418	13.3962	0.0003

Using estimated results as produced from the model and comparing them to the observed walked/not-walked classes from the data (coded as 1 = walked and 0 = not), the following statistic summary shows 77% of agreement and about 23% of misclassifications.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	77.0	Somers' D	0.542
Percent Discordant	22.8	Gamma	0.543
Percent Tied	0.2	Tau-a	0.271
Pairs	12531240	c	0.771

APPENDIX G. FINAL RESULTS AFTER FINAL DISCRIMINANT ANALYSIS FACTORS APPLIED

This section shows outputs from the Linear Discriminant Function procedure, which uses “walk trips per capita” measure as the independent variable. A total of 7,082 block group records were used (non-missing records) in this procedure, among them 3,688 block groups with no walking trips (group score0) and the remaining 3,394 records are block groups with at least one reported walk trips (group score1).

The DISCRIM Procedure

Total Sample Size	7082	DF Total	7081
Variables	94	DF Within Classes	7080
Classes	2	DF Between Classes	1

Number of Observations Read	7255
Number of Observations Used	7082

Class Level Information					
Wt_WkTrpPC_C	Variable Name	Frequency	Weight	Proportion	Prior Probability
Score0	Score0	3688	3688	0.520757	0.500000
Score1	Score1	3394	3394	0.479243	0.500000

The table below shows parameters for two functions (the two groups that we are discriminating – walked or not-walked). Each is used in a similar way as in a simple regression model.

Linear Discriminant Function for Wt_WkTrpPC_C		
Variable	Score0	Score1
Constant	-476826	-476882
Home_DestPctXOther_DestPct	269.48302	277.79793
AvgDrivers	-6.17268	-5.10981
TotHHs	-0.03724	-0.03687
RETDEN_LOGXRET_PCT	-65.62880	-65.57460
PctCollegeXPctGrad	126.76167	129.13092
HBEE MPDN	0.02601	0.02604
PctDrivers	128.20127	126.32829
B08301e21XPctWhiteHH	-0.01137	-0.00553
PctCollege	61.75502	61.54362
ACSPctGradXPctGrad	-206.45447	-203.83828
TotHHsXHome_DestPct	0.00191	0.00160
employmentXHBEE MPDN	8.03661E-6	8.01122E-6
RET_PCT	358.15181	357.30043
B24080e18	-5.80979	-5.82124
Home_DestPctXWork_DestPct	-168.90703	-165.29677
TotHHsXNIEL_PCT	0.00897	0.00896
RET_DEN	0.02067	0.02073
Home_DestPctXNIEL_DEN	0.00767	0.00771
RETDEN_LOG	88.04305	87.39729
TotHHsXAvgVehCount	0.00944	0.00954
AvgVehCountXB19001e14	0.29399	0.29628
PctHis HH	289.44026	286.84087
employmentXB19001e13	0.0000131	0.0000129
Work_DestPct	255.98128	255.20723
Home_DestPctXPctHis HH	69.50895	70.14805
PctEng HH	425.75104	422.68401
NIEL_DEN	0.02954	0.02958
NIEL_DENXRET_DEN	-4.6626E-7	-4.6765E-7
AvgDriversXB19001e14	-0.35430	-0.35748

Linear Discriminant Function for Wt_WkTrpPC_C		
Variable	Score0	Score1
ACSPctGrad	240.66015	242.04301
PctEngHHXPctHispHH	-376.38498	-373.93680
B08302e6XB08301e21	0.00112	0.00111
RETDEN_LOGXNIEL_DEN	-0.00406	-0.00407
AvgVehCount	-56.16034	-56.31789
HBPRCOTH	-10.19861	-10.17796
LIPERTOT08XB24080e18	1.40853	1.40997
AvgDriversXPctHispHH	23.74587	23.45484
PctCollegeXHBPRCCAUX	-0.42814	-0.42198
Other_DestPctXB02001e3	-0.04001	-0.04090
TotHHsXHBPHISP	0.0003575	0.0003571
RET_PCTXPctGrad	50.78360	51.33936
NIEL_DENXHBPHISP	0.0000201	0.0000201
Other_DestPctXB08302e10	0.29546	0.28877
B08302e10XHBPHISP	0.00588	0.00594
LIPERTOT08XOther_DestPct	-30.97013	-31.20456
AvgDriversXAvgVehCount	10.52574	10.39429
B19001e14	0.94334	0.94751
PCTRET_SQRTXWalkScore	57.06949	56.87703
employmentXWork_DestPct	-0.0004605	-0.0004486
B08302e6XB24080e18	0.0009653	0.0009760
HBPRCCAUXPctWhiteHH	-4.18430	-4.19095
TotHHsXB02001e3	0.0000298	0.0000298
LIPERTOT08XB08302e6	0.64933	0.64967
PctDriversXB08301e21	0.15115	0.14613
TotHHsXWork_DestPct	-0.11847	-0.11904
PctDriversXPctEngHH	-167.85141	-167.18782
WalkScore	-32.88617	-32.78294
RET_PCTXPctWhiteHH	23.15187	23.57367
PCTRET_SQRT	1151331	1151401
LIPERTOT08XWalkScore	1.25129	1.24989

Linear Discriminant Function for Wt_WkTrpPC_C		
Variable	Score0	Score1
HBPRCOTHXAvgVehCount	0.45422	0.44965
AvgDriversXNIEL_DEN	-0.00363	-0.00364
PctEngHHXC08134e42	0.04016	0.05167
LIPERTOT08XHBPHISP	-6.45155	-6.45200
B08302e10	-1.02290	-1.02566
Home_DestPctXB19001e14	-0.08187	-0.08417
employmentXB24080e18	-0.0002380	-0.0002381
EMPLOYMENT	-0.11670	-0.11633
LIPERTOT08XB19001e13	-0.80261	-0.80276
PctWhiteHH	200.19271	200.30769
RETDEN_LOGXemployment	0.01616	0.01613
AvgVehCountXPctWhiteHH	1.02756	1.14949
B08302e6	-5.95705	-5.95849
C08134e42	-1.16153	-1.17100
B08301e21	-0.39249	-0.39090
B02001e3	0.04468	0.04483
Home_DestPct	-5.31184	-4.66897
HBPRCCAU	-1.08435	-1.07808
HBPHISP	52.41991	52.41678
PctGrad	53.03126	52.57992
LIPERTOT08	480.14400	480.34891
Other_DestPct	89.80016	90.42762
NIEL_PCT	-11.52166	-11.51743
B19001e13	5.25102	5.25336
Entropy_R	-44.71498	-44.72460
HBHUR_S	-27.73812	-27.72784
ACSPctCollege	4.69691	5.16798
ACSPOP_LOG	182.54596	182.57122
FEM22_34	-0.09951	-0.09976
HBHMEDHSA	0.94501	0.94516
NIELDEN_SQRTBG	-2.31038	-2.31305

Linear Discriminant Function for Wt_WkTrpPC_C		
Variable	Score0	Score1
PCTRET_LOG	-269012	-269028
PCTRETAIL	46.34613	46.46520
RETDEN_SQRTBG	-1.26012	-1.25647

The DISCRIM Procedure
Classification Summary for Calibration Data
Re-substitution Summary using Linear Discriminant Function

Number of Observations and Percent Classified into Wt_WkTrpPC_C			
From Wt_WkTrpPC_C	Score0	Score1	Total
Score0	2602 70.55	1086 29.45	3688 100.00
Score1	1030 30.35	2364 69.65	3394 100.00
Total	3632 51.28	3450 48.72	7082 100.00
Priors	0.5	0.5	

Error Count Estimates for Wt_WkTrpPC_C			
	Score0	Score1	Total
Rate	0.2945	0.3035	0.2990
Priors	0.5000	0.5000	

The above two tables provide summary statistics on the estimated error rate using the discriminant function based on the 7,082 “non-missing” data records. Specifically, it indicates that the error rate for misclassification of “score0” is 29.5% and for “score1” is 30.4%; with an overall error rate measured at about 29.9%.

APPENDIX H. A LIST OF BLOCK GROUPS WITH THE HIGHEST 100 WALK INDICES

Table H.1 Block Groups in Study Region with Highest Walk Index

Region	County	Block group ID	Walk index	Walk Index Category
Austin	Hays	482090102002	95.4	81 - 100
	Travis	484530007001	97.5	81 - 100
		484530012004	96.7	81 - 100
		484530017321	96.0	81 - 100
		484530017324	97.4	81 - 100
		Williamson	484910201011	98.5
Cedar Rapids	Linn	191130001001	96.8	81 - 100
		191130002032	94.5	81 - 100
		191130019001	94.1	81 - 100
Dallas	Collin	480850305011	99.9	81 - 100
		480850305021	99.1	81 - 100
		480850305032	96.7	81 - 100
		480850313062	96.0	81 - 100
		480850314012	96.9	81 - 100
	Dallas	481130100003	99.7	81 - 100
		481130193024	96.9	81 - 100
	Denton	481210201021	98.1	81 - 100
		481210201023	97.0	81 - 100
		481210203051	97.5	81 - 100
		481210213023	97.0	81 - 100
		481210214023	98.9	81 - 100
		481210215041	96.9	81 - 100
		481210217052	96.4	81 - 100
		481210217103	94.8	81 - 100
		481210217141	99.8	81 - 100
		481210217142	96.1	81 - 100
		Kaufman	482570502023	95.1
	Los Angeles	Los Angeles	060371396001	95.0
060371914102			94.4	81 - 100
060372060101			96.0	81 - 100
060372119202			94.5	81 - 100
060372612009			95.7	81 - 100
060372971201			96.2	81 - 100

Region	County	Block group ID	Walk index	Walk Index Category
		060373019001	96.0	81 - 100
		060374019011	97.1	81 - 100
		060375751021	100.0	81 - 100
		060375758032	95.2	81 - 100
		060375762002	94.2	81 - 100
		060377001008	94.9	81 - 100
		060377011009	99.3	81 - 100
		060379201052	97.3	81 - 100
		060379203312	95.1	81 - 100
Madison	Dane	550250004041	95.6	81 - 100
		550250010001	95.5	81 - 100
		550250011001	95.2	81 - 100
		550250020003	95.0	81 - 100
		550250101001	98.4	81 - 100
		550250101002	98.4	81 - 100
		550250120011	97.3	81 - 100
Miami	Miami-Dade	120860001155	95.2	81 - 100
		120860027015	95.5	81 - 100
		120860071003	94.8	81 - 100
		120860101461	97.4	81 - 100
		120860101532	95.9	81 - 100
Research Triangle	Orange	371350114002	96.1	81 - 100
	Wake	371830534034	99.7	81 - 100
		371830536002	98.7	81 - 100
		371830537031	99.0	81 - 100
San Diego	San Diego	060730002004	95.3	81 - 100
		060730019001	96.1	81 - 100
		060730020012	95.3	81 - 100
		060730027074	97.4	81 - 100
		060730053004	96.4	81 - 100
		060730056001	94.9	81 - 100
		060730072001	98.0	81 - 100
		060730076002	94.5	81 - 100
		060730076005	95.7	81 - 100
		060730082003	94.9	81 - 100
		060730083032	97.3	81 - 100
		060730083103	98.1	81 - 100
		060730083301	98.0	81 - 100
060730083331	98.7	81 - 100		

Region	County	Block group ID	Walk index	Walk Index Category
		060730083352	96.5	81 - 100
		060730083381	97.8	81 - 100
		060730083421	95.6	81 - 100
		060730083431	95.8	81 - 100
		060730094001	100.0	81 - 100
		060730095052	97.5	81 - 100
		060730100141	99.4	81 - 100
		060730106029	96.0	81 - 100
		060730113009	99.6	81 - 100
		060730133131	99.9	81 - 100
		060730166101	95.4	81 - 100
		060730167011	94.4	81 - 100
		060730170291	96.0	81 - 100
		060730170301	95.3	81 - 100
		060730170322	97.1	81 - 100
		060730170352	96.9	81 - 100
		060730170381	98.4	81 - 100
		060730170421	97.6	81 - 100
		060730171051	98.6	81 - 100
		060730178121	97.5	81 - 100
		060730185142	94.8	81 - 100
		060730193011	95.8	81 - 100
		060730198061	98.7	81 - 100
		060730200104	98.3	81 - 100
		060730200131	97.4	81 - 100
		060730200165	95.4	81 - 100
Virginia - DC Area	Fairfax	510594154002	94.8	81 - 100
		510594822002	94.7	81 - 100

**Table H.2 Block Groups in the Study Region with Lowest Walk Index
(Non-zero and Non-missing)**

Region	County	Block group ID	Walk index	Walk Index Category
Austin	Travis	484530024132	3.0	1 - 10
Dallas	Collin	480850316485	2.1	1 - 10
	Dallas	481130018002	1.8	1 - 10
		481130063011	3.2	1 - 10
		481130068004	2.8	1 - 10
		481130078162	3.0	1 - 10
		481130093042	1.4	1 - 10
		481130096102	2.1	1 - 10
		481130106021	3.5	1 - 10
		481130109022	2.2	1 - 10
		481130111043	1.5	1 - 10
		481130111044	3.6	1 - 10
		481130111051	1.5	1 - 10
		481130141163	3.4	1 - 10
		481130141264	0.1	1 - 10
		481130159002	3.5	1 - 10
		481130167011	1.0	1 - 10
		481130167012	0.7	1 - 10
		Los Angeles	Los Angeles	060371048102
060371066043	3.4			1 - 10
060371219001	3.4			1 - 10
060371992013	2.5			1 - 10
060371998001	3.5			1 - 10
060372016003	1.9			1 - 10
060372075003	0.7			1 - 10
060372077102	0.0			1 - 10
060372149001	1.3			1 - 10
060372189002	2.9			1 - 10
060372227003	2.8			1 - 10
060372240101	3.4			1 - 10
060372242001	2.8			1 - 10
060372260004	3.5			1 - 10
060372260008	2.0			1 - 10
060372671001	2.6			1 - 10
060372947003	2.8			1 - 10
060372947005	3.0			1 - 10
060373202003	2.1	1 - 10		
060373202004	1.0	1 - 10		

Region	County	Block group ID	Walk index	Walk Index Category
		060373202005	1.0	1 - 10
		060373202006	1.2	1 - 10
		060373202007	2.0	1 - 10
		060374021011	3.0	1 - 10
		060374021012	1.0	1 - 10
		060374021013	2.3	1 - 10
		060374046009	0.8	1 - 10
		060374050013	1.9	1 - 10
		060374075003	1.6	1 - 10
		060375003001	0.0	1 - 10
		060375027005	2.5	1 - 10
		060375027006	3.1	1 - 10
		060375306021	1.7	1 - 10
		060375306022	2.5	1 - 10
		060375310006	2.9	1 - 10
		060375313011	1.6	1 - 10
		060375320011	2.4	1 - 10
		060375322003	1.8	1 - 10
		060375323031	3.2	1 - 10
		060375324004	2.2	1 - 10
		060375333002	1.1	1 - 10
		060375334031	3.0	1 - 10
		060375352002	2.5	1 - 10
		060375413005	3.6	1 - 10
		060375424013	3.6	1 - 10
		060375431005	2.6	1 - 10
		060375433042	0.9	1 - 10
		060375755004	0.0	1 - 10
		060377008006	2.3	1 - 10
Madison	Dane	550250011002	0.4	1 - 10
		120860004032	1.6	1 - 10
		120860004034	0.0	1 - 10
		120860010024	2.5	1 - 10
		120860017033	3.4	1 - 10
		120860022013	0.1	1 - 10
		120860025005	3.1	1 - 10
		120860054024	3.2	1 - 10
		120860057033	3.4	1 - 10
		120860072003	1.6	1 - 10
		120860077021	2.0	1 - 10
		120860083062	0.8	1 - 10
Miami	Miami-Dade			

Region	County	Block group ID	Walk index	Walk Index Category
		120860091009	1.1	1 - 10
		120860094001	3.4	1 - 10
		120860099028	1.4	1 - 10
Research Triangle	Durham	370630013012	1.3	1 - 10
		370630018043	0.2	1 - 10
		370630018051	3.4	1 - 10
		370630020141	1.7	1 - 10
	Orange	371350107041	2.7	1 - 10
	Wake	371830519001	3.4	1 - 10
		371830528032	3.3	1 - 10
		371830528052	3.0	1 - 10
371830541062		3.4	1 - 10	
San Diego	San Diego	060730063009	0.0	1 - 10
		060730187009	1.5	1 - 10
		060730203072	3.0	1 - 10
Virginia - DC Area	Alexandria City	515102001041	2.2	1 - 10
	Fairfax City	516003001001	1.8	1 - 10
		516003004001	2.0	1 - 10
		516003004002	3.4	1 - 10
	Falls Church City	516105002001	3.5	1 - 10
		516105002002	1.1	1 - 10

APPENDIX I. NHTS SAMPLE COVERAGE AND WALK INDICES IN SELECTED STUDY REGIONS

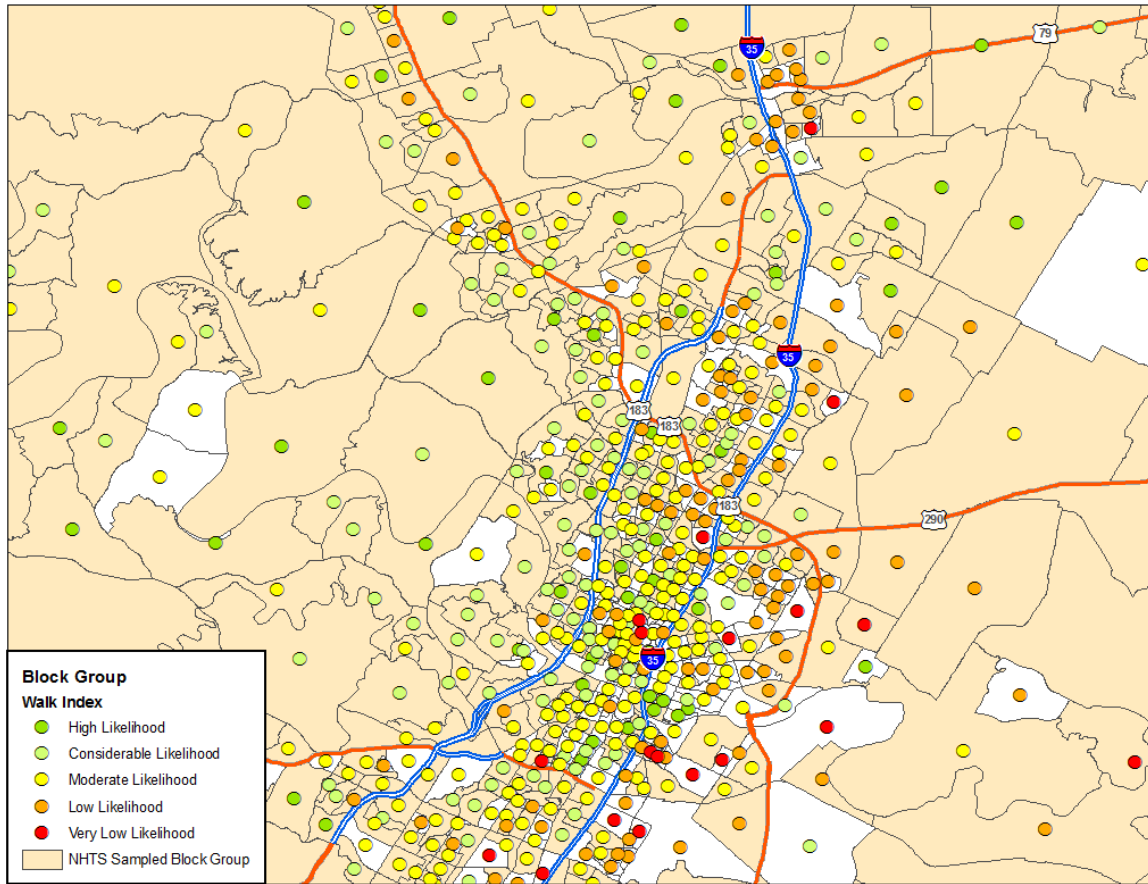


Figure I.1 NHTS sample coverage and walk indices in block groups within the Austin TX region

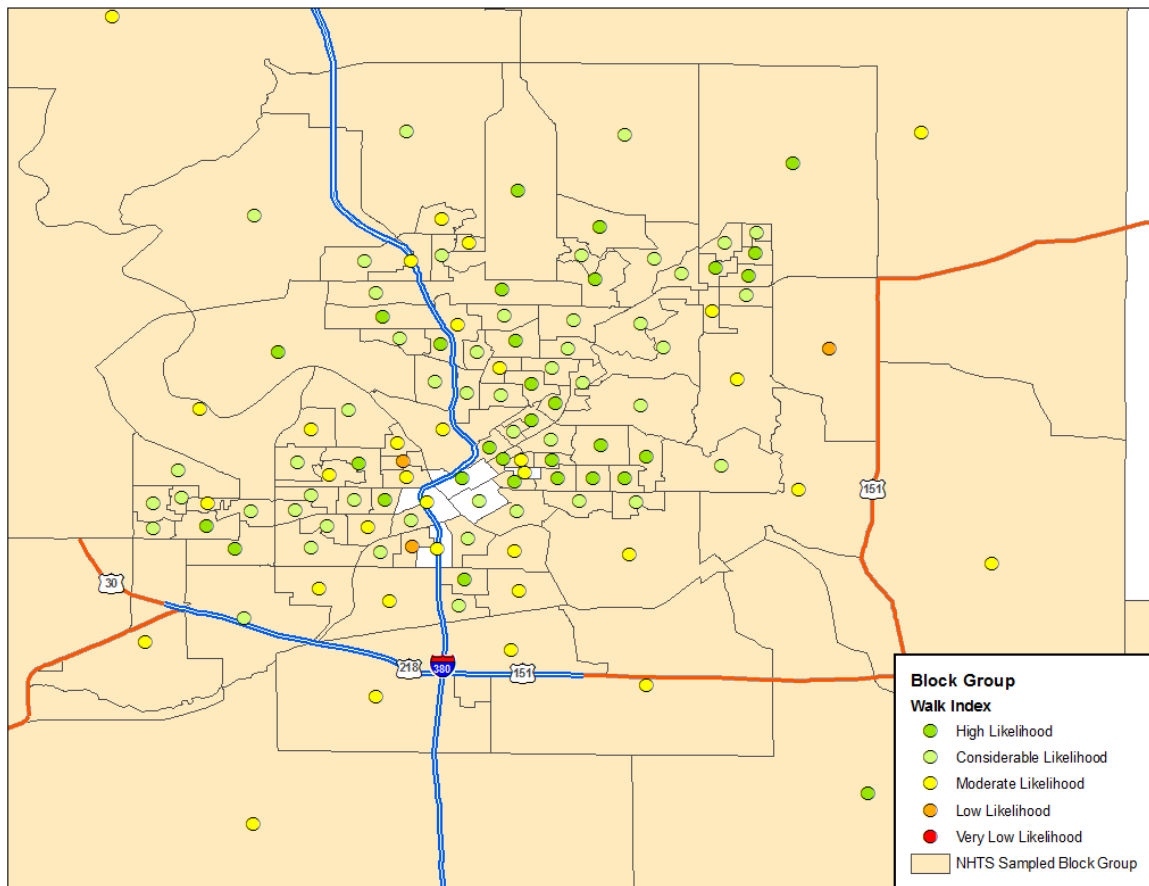


Figure I.2 NHTS sample coverage and walk indices in block groups within the Cedar Rapids, IA region

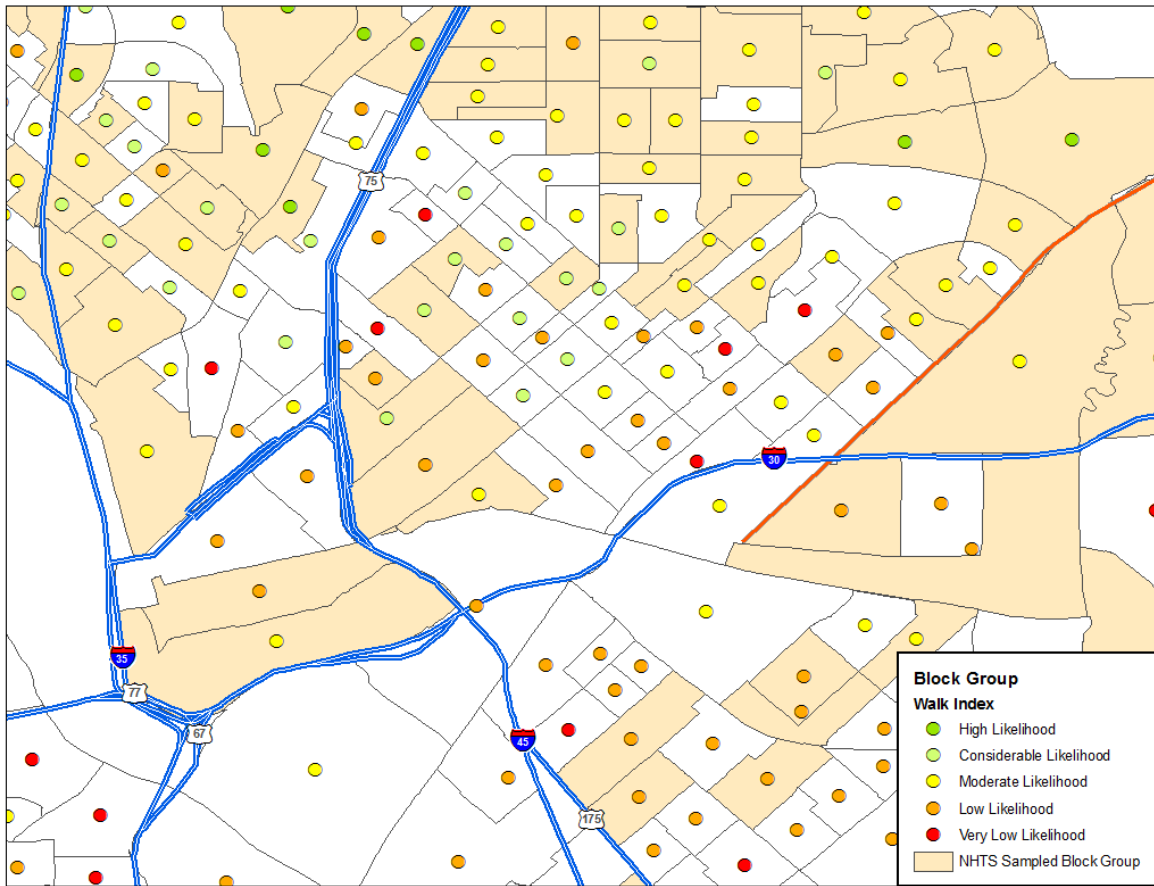


Figure I.3 NHTS sample coverage and walk indices in block groups within the Dallas, TX region

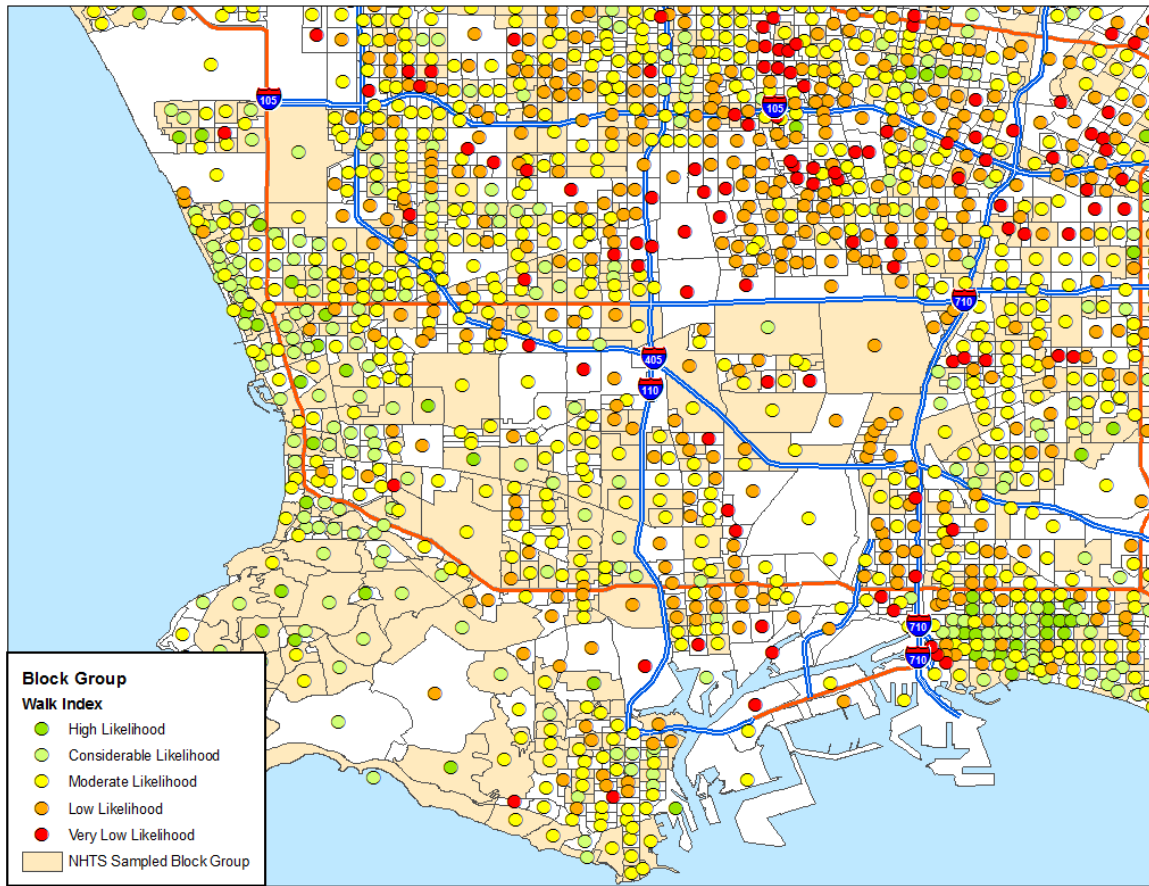


Figure I.4 NHTS sample coverage and walk indices in block groups within the Los Angeles, CA region

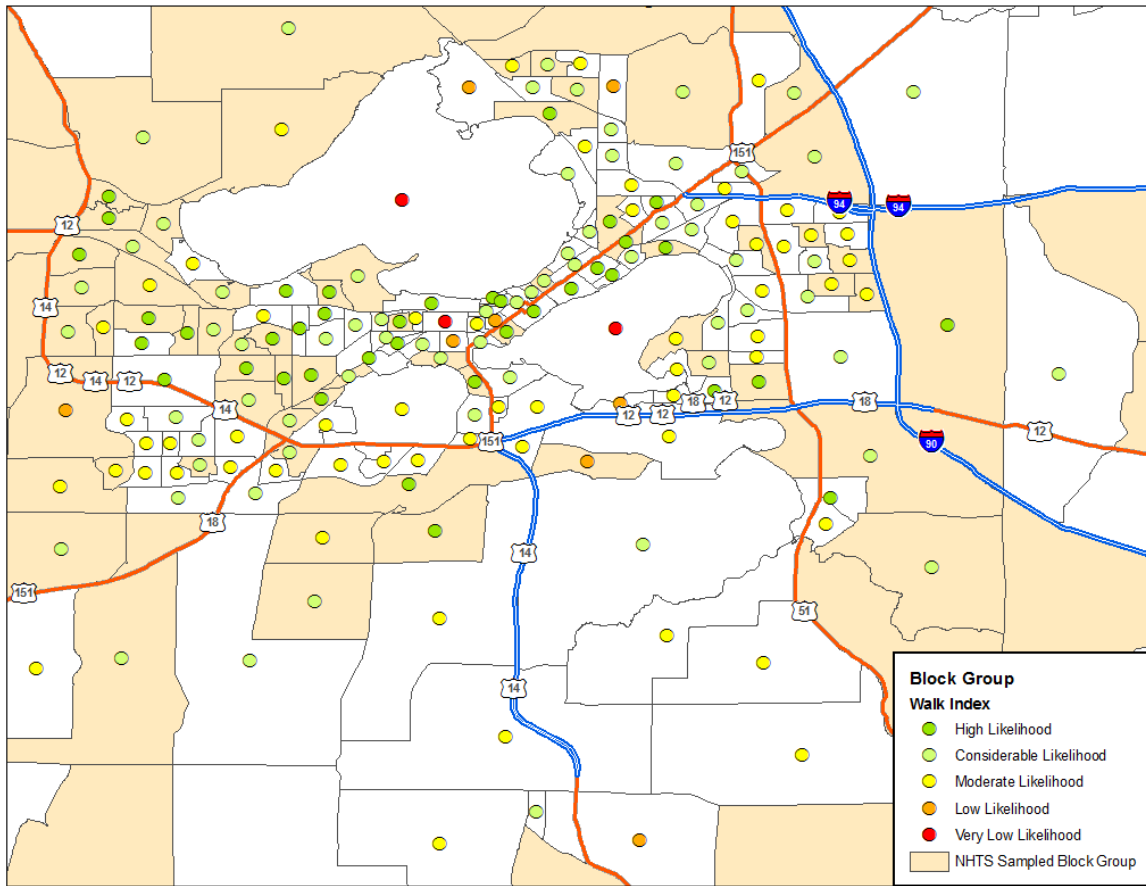


Figure I.5 NHTS sample coverage and walk indices in block groups within the Madison, WI region

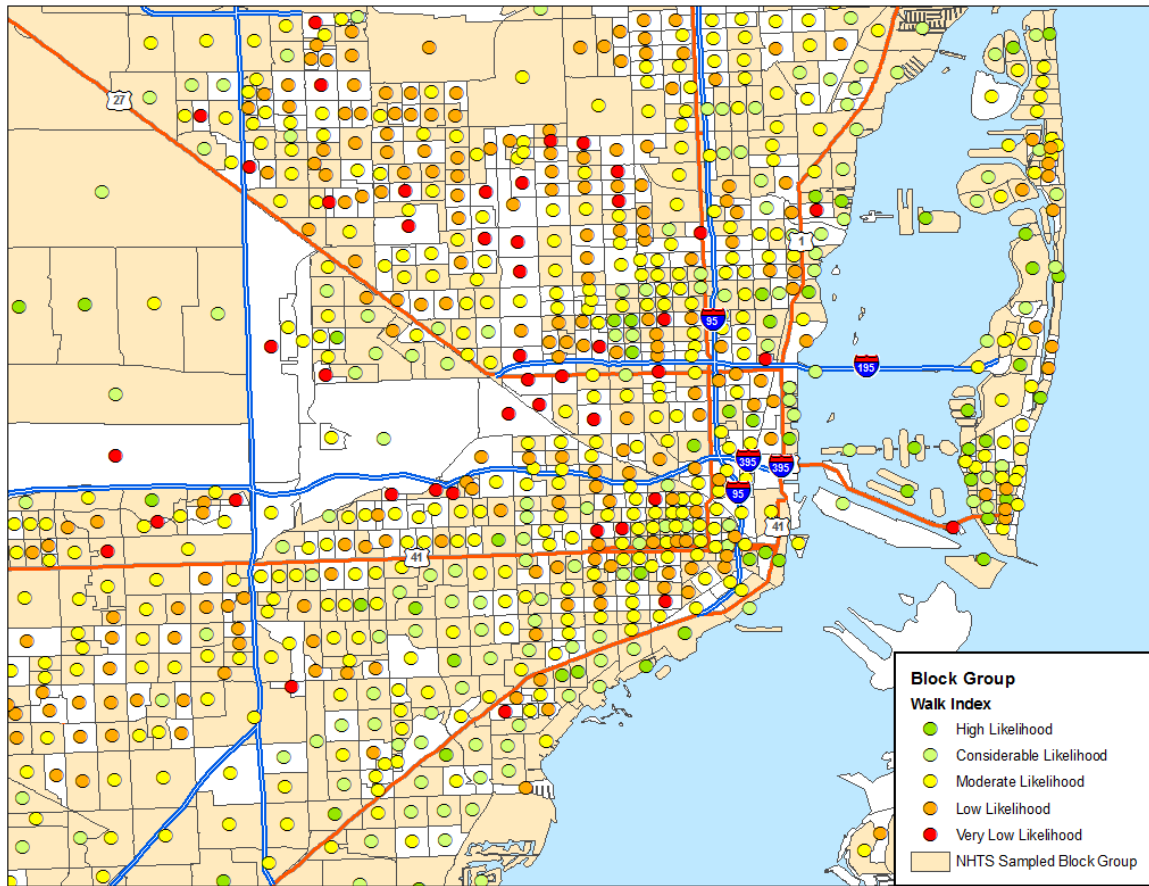


Figure I.6 NHTS sample coverage and walk indices in block groups within the Miami, FL region

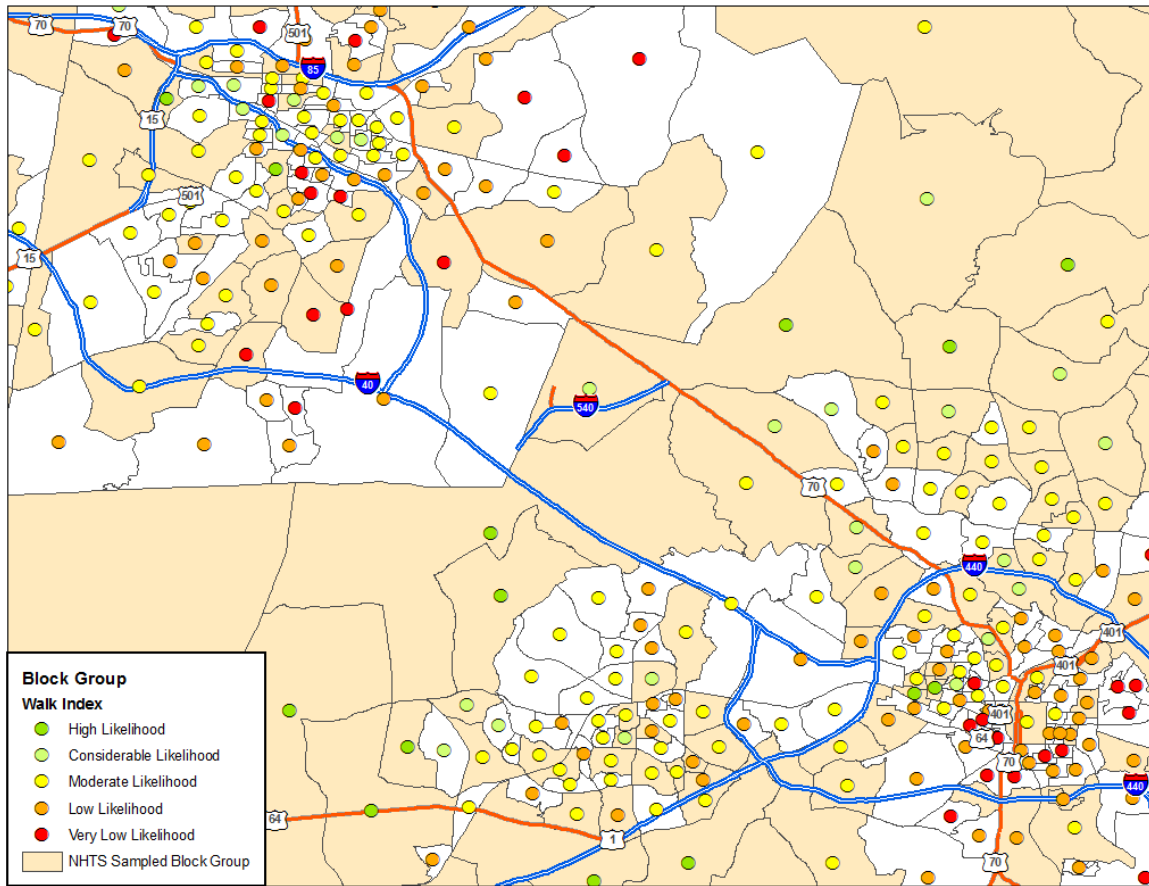


Figure I.7 NHTS sample coverage and walk indices in block groups within the Research Triangle, NC region