**TASK 2 MODEL DESIGN AND ESTIMATION REPORT (UPDATED DRAFT)**

**SPR 788**

By

Liming Wang

Portland state University

Toulan School of Urban Studies and Planning

for

oregon department of transportation

Research Section

555 13th Street NE, Suite 1

Salem OR 97301

and

Federal Highway Administration

400 Seventh Street, SW

Washington, DC 20590-0003

**May, 2017**

Technical Report Documentation Page

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 1. 1. Report No. | 2. Government Accession No. | | | 3. Recipient’s Catalog No. | |
| 4. Title and Subtitle | | | | 1. Report Date   -month- -year- | |
| 1. Performing Organization Code | |
| 7. Author(s) | | | | 8. Performing Organization Report No. | |
| 9. Performing Organization Name and Address  Oregon Department of Transportation  Research Section  555 13th Street NE, Suite 1  Salem, OR 97301 | | | | 10. Work Unit No. (trais) | |
| 1. Contract or Grant No. | |
| 1. Sponsoring Agency Name and Address   Oregon Dept. of Transportation  Research Section and Federal Highway Admin.  555 13th Street NE, Suite 1 400 Seventh Street, SW  Salem, OR 97301 Washington, DC 20590-0003 | | | | 1. Type of Report and Period Covered   \_\_\_\_\_\_\_ Report | |
| 1. Sponsoring Agency Code | |
| 1. Supplementary Notes | | | | | |
| 1. Abstract | | | | | |
| 17. Key Words | | 18. Distribution Statement  Copies available from NTIS, and online at <http://www.oregon.gov/ODOT/TD/TP_RES/> | | | |
| 19. Security Classification (of this report)  Unclassified | 1. Security Classification (of this page)   Unclassified | | 21. No. of Pages  XXX | | 22. Price |

Technical Report Form dot f 1700.7 (8-72) Reproduction of completed page authorized Printed on recycled paper

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **SI\* (MODERN METRIC) CONVERSION FACTORS** | | | | | | | | | | | | | | | | |
| **APPROXIMATE CONVERSIONS TO SI UNITS** | | | | | | | | **APPROXIMATE CONVERSIONS FROM SI UNITS** | | | | | | | | |
| Symbol | | When You Know | Multiply By | | To Find | Symbol | | Symbol | | When You Know | Multiply By | | To Find | Symbol | | |
| **LENGTH** | | | | | | | | **LENGTH** | | | | | | | | |
| in | inches | | | 25.4 | millimeters | | mm | mm | millimeters | | | 0.039 | inches | | in | |
| ft | feet | | | 0.305 | meters | | m | m | meters | | | 3.28 | feet | | ft | |
| yd | yards | | | 0.914 | meters | | m | m | meters | | | 1.09 | yards | | yd | |
| mi | miles | | | 1.61 | kilometers | | km | km | kilometers | | | 0.621 | miles | | mi | |
| **AREA** | | | | | | | | **AREA** | | | | | | | | |
| in2 | square inches | | | 645.2 | millimeters squared | | mm2 | mm2 | millimeters squared | | | 0.0016 | square inches | | in2 | |
| ft2 | square feet | | | 0.093 | meters squared | | m2 | m2 | meters squared | | | 10.764 | square feet | | ft2 | |
| yd2 | square yards | | | 0.836 | meters squared | | m2 | m2 | meters squared | | | 1.196 | square yards | | yd2 | |
| ac | acres | | | 0.405 | hectares | | ha | ha | hectares | | | 2.47 | acres | | ac | |
| mi2 | square miles | | | 2.59 | kilometers squared | | km2 | km2 | kilometers squared | | | 0.386 | square miles | | mi2 | |
| **VOLUME** | | | | | | | | **VOLUME** | | | | | | | | |
| fl oz | fluid ounces | | | 29.57 | milliliters | | ml | ml | milliliters | | | 0.034 | fluid ounces | | fl oz | |
| gal | gallons | | | 3.785 | liters | | L | L | liters | | | 0.264 | gallons | | gal | |
| ft3 | cubic feet | | | 0.028 | meters cubed | | m3 | m3 | meters cubed | | | 35.315 | cubic feet | | ft3 | |
| yd3 | cubic yards | | | 0.765 | meters cubed | | m3 | m3 | meters cubed | | | 1.308 | cubic yards | | yd3 | |
| NOTE: Volumes greater than 1000 L shall be shown in m3. | | | | | | | |  |  | | |  |  | |  | |
| **MASS** | | | | | | | | **MASS** | | | | | | | | |
| oz | ounces | | | 28.35 | grams | | g | g | grams | | | 0.035 | ounces | | | oz |
| lb | pounds | | | 0.454 | kilograms | | kg | kg | kilograms | | | 2.205 | pounds | | | lb |
| T | short tons (2000 lb) | | | 0.907 | megagrams | | Mg | Mg | megagrams | | | 1.102 | short tons (2000 lb) | | | T |
| **TEMPERATURE (exact)** | | | | | | | | **TEMPERATURE (exact)** | | | | | | | | |
| °F | Fahrenheit | | | (F-32)/1.8 | Celsius | | °C | °C | Celsius | | | 1.8C+32 | Fahrenheit | | °F | |
| \*SI is the symbol for the International System of Measurement | | | | | | | | | | | | | | | | |

acknowledgements

Disclaimer

This document is disseminated under the sponsorship of the Oregon Department of Transportation and the United States Department of Transportation in the interest of information exchange. The State of Oregon and the United States Government assume no liability of its contents or use thereof.

The contents of this report reflect the view of the authors who are solely responsible for the facts and accuracy of the material presented. The contents do not necessarily reflect the official views of the Oregon Department of Transportation or the United States Department of Transportation.

The State of Oregon and the United States Government do not endorse products of manufacturers. Trademarks or manufacturers’ names appear herein only because they are considered essential to the object of this document.

This report does not constitute a standard, specification, or regulation.

**Table of Contents**

[1 Introduction 8](#_Toc485616129)

[2 Data Sources 8](#_Toc485616130)

[2.1 Descriptive Statistics 9](#_Toc485616131)

[2.1.1 2009 NHTS 9](#_Toc485616132)

[2.1.1.1 Travel Mode Reclassificiation 10](#_Toc485616133)

[2.1.1.2 Descriptive Statistics 11](#_Toc485616134)

[2.2 SmartLocation Database (SLD) 19](#_Toc485616135)

[2.3 Place Types 19](#_Toc485616136)

[2.4 Model Structures 21](#_Toc485616137)

[2.4.1 Current GreenSTEP DVMT models 21](#_Toc485616138)

[2.4.2 Proposed New Models 23](#_Toc485616139)

[2.4.2.1 AADVMT Model (Power-transformed linear regression model) 23](#_Toc485616140)

[2.4.2.2 AADVMT Model (hurdle model) 25](#_Toc485616141)

[2.4.2.3 Person Miles Traveled by Mode Models 30](#_Toc485616142)

[2.4.2.3.1 Transit Miles Traveled Model (hurdle model) 30](#_Toc485616143)

[2.4.2.3.2 Walk Miles Traveled Model (hurdle model) 34](#_Toc485616144)

[2.4.2.3.3 Bike Miles Traveled Model (hurdle model) 38](#_Toc485616145)

[2.4.2.4 Trip Frequency-Length (TFL) Models 41](#_Toc485616146)

[2.4.2.4.1 Trip Frequency Models 41](#_Toc485616147)

[2.4.2.4.1.1 Transit Trip Frequency Model 42](#_Toc485616148)

[2.4.2.4.1.2 Walking Trip Frequency Model 46](#_Toc485616149)

[2.4.2.4.1.3 Biking Trip Frequency Model 51](#_Toc485616150)

[2.4.2.4.2 Average Trip Length Models 54](#_Toc485616151)

[2.4.2.4.2.1 Transit Trip Length Model 55](#_Toc485616152)

[2.4.2.4.2.2 Walking Trip Length Model 57](#_Toc485616153)

[2.4.2.4.2.3 Biking Trip Length Model 59](#_Toc485616154)

[2.4.3 Considered Model Structures 61](#_Toc485616155)

[2.4.3.1 Person Miles Traveled by Mode (PMT) Model 61](#_Toc485616156)

[2.5 Next step 62](#_Toc485616157)

[2.6 Other considerations 62](#_Toc485616158)

[2.7 References 62](#_Toc485616159)

**LIST OF TABLES**

[Table 2.1 Auto Modes 10](#_Toc485616098)

[Table 2.2 Transit Modes 10](#_Toc485616099)

[Table 2.3 Bike 10](#_Toc485616100)

[Table 2.4 Walk 11](#_Toc485616101)

[Table 2.5 Other Modes (not being modelled) 11](#_Toc485616102)

[Table 2.6 Name of selected variables 11](#_Toc485616103)

[Table 2.7 Trip frequencies by mode (unweighted) 14](#_Toc485616104)

[Table 2.8 Shares of trips by trip purpose and mode 14](#_Toc485616105)

[Table 2.9 Summary of trip distance by mode 16](#_Toc485616106)

[Table 2.10 Descriptive of total household travel distance (miles) by mode used: 16](#_Toc485616107)

[Table 2.11 Descriptive of household travel time (minutes) by mode used: 17](#_Toc485616108)

[Table 2.12 Binomial Logit Models of Zero DVMT 21](#_Toc485616109)

[Table 2.13 ZeroDVMT model goodness-of-fit 22](#_Toc485616110)

[Table 2.14 Power-transformed Regression Models of DVMT (DVMT > 0) 22](#_Toc485616111)

[Table 2.15 DVMT model goodness-of-fit 23](#_Toc485616112)

[Table 2.16 Combined model goodness-of-fit 23](#_Toc485616113)

[Table 2.17 Power-transformed Regression Models of AADVMT 24](#_Toc485616114)

[Table 2.18 Validation of AADVMT Models 25](#_Toc485616115)

[Table 2.19 AADVMT model (Hurdle model) goodness-of-fit and validation 29](#_Toc485616116)

[Table 2.20 Transit Miles Model (hurdle model) performance 33](#_Toc485616117)

[Table 2.21 Walk miles model (hurdle model) performance 37](#_Toc485616118)

[Table 2.22 Bike miles model (hurdle model) performance 40](#_Toc485616119)

[Table 2.23 Transit trip frequency model performance 45](#_Toc485616120)

[Table 2.24 Walk trip frequency model performance 50](#_Toc485616121)

[Table 2.25 Biking trip frequency model performance 54](#_Toc485616122)

[Table 2.26 Power-transformed Regression Models of Average Transit Trip Length 55](#_Toc485616123)

[Table 2.27 Transit length model performance 56](#_Toc485616124)

[Table 2.28 Power-transformed Regression Models of Average Walk Trip Length 57](#_Toc485616125)

[Table 2.29 Walking trip length model performance 58](#_Toc485616126)

[Table 2.30 Power-transformed Regression Models of Average Bike Trip Length 59](#_Toc485616127)

[Table 2.31 Biking trip length model 60](#_Toc485616128)

**LIST OF FIGURES**

[Figure 2.1 Shares of trips by trip purpose and mode 15](#_Toc482779263)

[Figure 2.2 The distribution of raw trip distance (miles) 16](#_Toc482779264)

[Figure 2.3 Boxplot of household travel distance (mile) and time (minutes) by mode 17](#_Toc482779265)

[Figure 2.4 Distribution of VMT after transformation 18](#_Toc482779266)

[Figure 2.5 Temporal distribution of travel day 18](#_Toc482779267)

[Figure 2.6 Sensitivity Test of VehPerDriver for AADVMT Model (Power-transformed linear regression model) 25](#_Toc482779268)

[Figure 2.7 Sensitivity Test of LogIncome for AADVMT Model (Power-transformed linear regression model) 26](#_Toc482779269)

[Figure 2.8 Sensitivity test of D1B for AADVMT Model (Power-transformed linear regression model) 26](#_Toc482779270)

[Figure 2.9 Sensitivity test of TRPPOPDEN for AADVMT Model (Power-transformed linear regression model) 27](#_Toc482779271)

[Figure 2.10 Sensitivity test of Tranmilecap for AADVMT Model (Power-transformed linear regression model) 27](#_Toc482779272)

[Figure 2.11 Sensitivity test of VehPerDriver for AADVMT Model (hurdle model) 32](#_Toc482779273)

[Figure 2.12 Sensitivity test of LogIncome for AADVMT Model (hurdle model) 32](#_Toc482779274)

[Figure 2.13 Sensitivity test of D1B for AADVMT Model (hurdle model) 33](#_Toc482779275)

[Figure 2.14 Sensitivity test of TRPOPDEN for AADVMT Model (hurdle model) 33](#_Toc482779276)

[Figure 2.15 Sensitivity test of Tranmilescap for AADVMT Model (hurdle model) 34](#_Toc482779277)

[Figure 2.16 Transit miles distribution 35](#_Toc482779278)

[Figure 2.17 Sensitivity test of HHSIZE for transit miles traveled model (hurdle model) 39](#_Toc482779279)

[Figure 2.18 Sensitivity test of VehPerDriver for transit miles traveled model (hurdle model) 39](#_Toc482779280)

[Figure 2.19 Sensitivity test of D1D for transit miles traveled model (hurdle model) 40](#_Toc482779281)

[Figure 2.20 Sensitivity test of LogIncome for transit miles traveled model (hurdle model) 40](#_Toc482779282)

[Figure 2.21 Sensitivity test of Tranmilescap for transit miles traveled model (hurdle model) 41](#_Toc482779283)

[Figure 2.22 Walk miles distribution 41](#_Toc482779284)

[Figure 2.23 Sensitivity test of AADVMT for walk miles traveled model (hurdle model) 45](#_Toc482779285)

[Figure 2.24 Sensitivity test of VehPerDriver for walk miles traveled model (hurdle model) 46](#_Toc482779286)

[Figure 2.25 Sensitivity test of D1D for walk miles traveled model (hurdle model) 46](#_Toc482779287)

[Figure 2.26 Sensitivity test of LogIncome for walk miles traveled model (hurdle model) 47](#_Toc482779288)

[Figure 2.27 Sensitivity test of Tranmilescap for walk miles traveled model (hurdle model) 47](#_Toc482779289)

[Figure 2.28 Bike Miles distribution 48](#_Toc482779290)

[Figure 2.29 Sensitivity test of AADVMT for bike miles traveled model (hurdle model) 51](#_Toc482779291)

[Figure 2.30 Sensitivity test of VehPerDriver for bike miles traveled model (hurdle model) 51](#_Toc482779292)

[Figure 2.31 Sensitivity test of LogIncome for bike miles traveled model (hurdle model) 52](#_Toc482779293)

[Figure 2.32 Sensitivity test of Tranmilescap for bike miles traveled model (hurdle model) 52](#_Toc482779294)

[Figure 2.33 Flow Chart of Trip Frequency-Length Model 53](#_Toc482779295)

[Figure 2.34 Transit trip frequency distribution 54](#_Toc482779296)

[Figure 2.35 Sensitivity test of AADVMT for transit trip frequency model 57](#_Toc482779297)

[Figure 2.36 Sensitivity test of VehPerDriver for transit trip frequency model 58](#_Toc482779298)

[Figure 2.37 Sensitivity test of LogIncome for transit trip frequency model 58](#_Toc482779299)

[Figure 2.38 Sensitivity test of D1B for transit trip frequency model 59](#_Toc482779300)

[Figure 2.39 Sensitivity test of D1D for transit trip frequency model 59](#_Toc482779301)

[Figure 2.40 Sensitivity test of Tranmilescap for transit trip frequency model 60](#_Toc482779302)

[Figure 2.41 Walk trip frequency 60](#_Toc482779303)

[Figure 2.42 Sensitivity test of AADVMT for walking trip frequency model 65](#_Toc482779304)

[Figure 2.43 Sensitivity test of VehPerDriver for walking trip frequency model 65](#_Toc482779305)

[Figure 2.44 Sensitivity test of LogIncome for walking trip frequency model 66](#_Toc482779306)

[Figure 2.45 Sensitivity test of D1D for walking trip frequency model 66](#_Toc482779307)

[Figure 2.46 Sensitivity test of Tranmilescap for walking trip frequency model 67](#_Toc482779308)

[Figure 2.47 Bike trip frequency distribution 67](#_Toc482779309)

[Figure 2.48 Sensitivity test of AADVMT for biking trip frequency model 71](#_Toc482779310)

[Figure 2.49 Sensitivity test of VehPerDriver for biking trip frequency model 71](#_Toc482779311)

[Figure 2.50 Sensitivity test of LogIncome for biking trip frequency model 72](#_Toc482779312)

[Figure 2.51 Sensitivity test of D1A for biking trip frequency model 72](#_Toc482779313)

[Figure 2.52 Sensitivity test of D1C for biking trip frequency model 73](#_Toc482779314)

[Figure 2.53 Sensitivity test of Tranmilescap for biking trip frequency model 73](#_Toc482779315)

[Figure 2.54 Transit average trip distance distribution 74](#_Toc482779316)

[Figure 2.55 Sensitivity test of AADVMT for transit trip length model 76](#_Toc482779317)

[Figure 2.56 Sensitivity test of VehPerDriver for transit trip length model 77](#_Toc482779318)

[Figure 2.57 Sensitivity test of LogIncome for transit trip length model 77](#_Toc482779319)

[Figure 2.58 Sensitivity test of D1D for transit trip length model 78](#_Toc482779320)

[Figure 2.59 Sensitivity test of Tranmilescap for transit trip length model 78](#_Toc482779321)

[Figure 2.60 Walk average trip distance distribution 79](#_Toc482779322)

[Figure 2.61 Sensitivity test of VehPerDriver for walking trip length model 81](#_Toc482779323)

[Figure 2.62 Sensitivity test of LogIncome for walking trip length model 81](#_Toc482779324)

[Figure 2.63 Sensitivity test of D1D for walking trip length model 82](#_Toc482779325)

[Figure 2.64 Sensitivity test of Tranmilescap for walking trip length model 82](#_Toc482779326)

[Figure 2.65 Bike average trip distance 83](#_Toc482779327)

[Figure 2.66 Sensitivity test of AADVMT for biking trip length model 85](#_Toc482779328)

[Figure 2.67 Sensitivity test of LogIncome for biking trip length model 85](#_Toc482779329)

[Figure 2.68 Sensitivity test of D1D for biking trip length model 86](#_Toc482779330)

[Figure 2.69 Sensitivity test of UZAEMPDEN for biking trip length model 86](#_Toc482779331)

[Figure 2.70 Sensitivity test of Tranmilescap for biking trip length model 87](#_Toc482779332)

[Figure 2.71 Flow Chart of Person Miles Traveled by Mode (PMT) Model 87](#_Toc482779333)

# Introduction

Task 2 of the project is to "select one or more possible model designs for RSPM mode shift, estimate model parameters and evaluate the designs and estimated parameters with sensitivity tests and validation". More specifically, the plan is to select and estimate one or more possible designs of the mode choice model based on literature review and data exploration in Task 1 and to understand what mode shifts occur as vehicle travel is reduced, incorporating and testing interactions in RSPM. These approaches build on the existing RSPM module and utilize household and land use inputs and budget constraints already embedded in the RSPM tool. The PSU team will suggest functional form and independent variables for model estimation with associated data sources for estimation and validation. PSU researchers will also identify sensitivity tests to assess the upgraded model with literature elasticities, repeating some of the tests previously calculated by the RSPM to ensure these remain intact, as well as adding tests to evaluate the new functionality. The PSU team will discuss and coordinate with Brian Gregor in the model design and estimation process, as he implements the RSPM common framework, to make sure the design and data format match the latest RSPM modeling framework. ODOT staff shall review and adjust the proposed designs, estimation data and validation data/approach.

The deliverable of Task 2 is a working paper (this document) that describes model designs, data sources, estimation, results of sensitivity tests and validation; documented R scripts used to process and analyze data.

Members of the TAC/OSA contract will review and suggest adjustments to the PSU researchers for model design, estimation data and results, validation data, approach and results; guide the selection of the best model design.

# Data Sources

The primary data sources we identified and used for Task 2 are the **2009 National Household Travel Survey (NHTS)** and **2010 EPA Smart Location Database (SLD)**. Additional data sources include TTI's Urban Mobility Report dataset and National Transit Database. We retrieved the 2009 NHTS data with confidential block group level residence location (and Census Tract and ZIP code of workplace location), which is the ideal data set we eyed for modeling mode choice. With the confidential residential block group location, we joined the 2009 NHTS with the 2010 SLD to get a combined dataset of travel information and built environment/urban form variables of households' residential block group.

We also looked into alternative data sources, including the 2012 California Travel Survey (CATS), 2011 Oregon Household Travel and Activity Survey (OHAS), the 2014-2015 Puget Sound Regional Travel Study (PSRTS) and explored the potential to combine these 3 surveys to create a unified data set with diverse coverage. But with the 2009 NHTS data with confidential residential block group location, there is limited additional benefit of pursuing combining CATS, OHAS and PSRTS for a number of reasons. First, these three data sets were collected in different years, which may create odd effects that are included in the model estimation. Second, since each of the data sets was collected by a different State/agency, the information collected varies and how the variables are measured or coded are likely different, which adds extra work to create a unified data set at the least and may weaken the final models at the worst.

There are two areas that extra data would be still beneficial. One area we wish to have a better handle is the day-to-day variation in mode choice and total demand (for example, the amount of driving measured in vehicle miles traveled), so that we can predict long-term behavior from a daily model. However, NHTS, as well as the three travel surveys above only capture the travel information for one single day. In GreenSTEP, Brian Gregor assumed the stochasticity in household daily VMT model (a linear regression model with transformed VMT as the dependent variable) represents the day-to-day variation in VMT. Such approximation of weekly VMT from daily information may be imperfect. Verification of the relationship between daily and longer term VMT and an explicit model of weekly (or annual) VMT may be necessary. A few potential data sets would be helpful in looking into the relationship. In particular, the 2004 – 2006 Traffic Choices Study by the Puget Sound Regional Council. For a pilot project on congestion-based tolling sponsored by Federal Highway Administration, the study placed GPS data loggers into the vehicles of about 275 households in the Seattle metropolitan area. The project recorded roughly 18 months of trip data (from November 2004 to April 2006) and included more than 400 vehicles. Such long-term data would be helpful to look into the relationship between daily and long term VMT.

Another potential area we are looking into for improvement is the modeling of price elasticities of travel demand. Brian tested three different methods of capturing price elasticities: income effect, price coefficient and household budget model. There are a number of challenges to get realistic price elasticities, including 1) The lack of disaggregate panel data that can be used to study how household travel decisions change over time in response to changes in fuel prices; 2) The relatively low historical price of fuel; 3) The prospect for future fuel prices that may be several times greater than present prices; 4) A lack of research consensus on the magnitude of the effects; and, 5) The difficulty of sorting out short range and long range effects.

Because of these challenges, the first two methods do not have sufficient sensitivity and Brian adopted the household budget model. All the challenges Brian identified above remain for the current project. Using the household budget model as the baseline model, we hope to draw from literature around the world (for example, Graham and Glaister, 2002) on the magnitude of the price elasticities and explore alternative methods of incorporating the elasticities into the new model of travel demand. Tolling studies such as the Puget Sound Traffic Choices Study provide some useful information on the price elasticities of travel demand (even though not from fuel price change).

## Descriptive Statistics

### 2009 NHTS

In addition to surveyed households' socio-demographic characteristics, the 2009 NHTS (USDOT Federal Highway Administration, 2011) collected daily trips taken in a 24-hour period, and includes:

* purpose of the trip (work, shopping, etc.);
* means of transportation used (car, bus, subway, walk, etc.);
* how long the trip took, i.e., travel time;
* time of day when the trip took place;
* day of week when the trip took place; and
* if a private vehicle trip:
* number of people in the vehicle , i.e., vehicle occupancy;
* driver characteristics (age, sex, worker status, education level, etc.); and
* vehicle attributes (make, model, model year, amount of miles driven in a year).

The 2009 NHTS included 150,145 households, 308,901 household members and 1,079,763 trips.

#### Travel Mode Reclassificiation

According to [codebook for G34 TRPTRANS](http://nhts.ornl.gov/tables09/CodebookPage.aspx?id=1084), we re-classify the modes into 5 categories:

Table 2.1 Auto Modes

|  |  |
| --- | --- |
| **Code** | **Name** |
| 1 | Car |
| 2 | Van |
| 3 | SUV |
| 4 | pickup truck |
| 5 | other truck |
| 6 | recreational vehicle |
| 7 | motorcycle |

Table 2.2 Transit Modes

|  |  |
| --- | --- |
| **Code** | **Name** |
| 9 | transit bus |
| 10 | commuter bus |
| 11 | school bus |
| 12 | charter bus |
| 13 | city to city bus |
| 14 | Shuttle bus |
| 15 | Amtrak |
| 16 | Commuter train |
| 17 | Subway |
| 18 | Street car/trolley |

Table 2.3 Bike

|  |  |
| --- | --- |
| **Code** | **Name** |
| 22 | bicycle |

Table 2.4 Walk

|  |  |
| --- | --- |
| **Code** | **Name** |
| 23 | walk |

Table 2.5 Other Modes (not being modelled)

|  |  |
| --- | --- |
| **Code** | **Name** |
| 8 | Light electric veh (golf cart) |
| 19 | taxi cab |
| 20 | Ferry |
| 21 | airplanes |
| 24 | Special transit-people w/disabilitie |

#### Descriptive Statistics

Table 2.6 Name of selected variables

|  |  |
| --- | --- |
| **Variable Names** | **Description** |
| ANNMILES | [NHTS] Self-reported annualized mile estimate |
| BESTMILE | [NHTS] Best estimate of annual miles (ORNL) |
| TDAYDATE | [NHTS] Date of Travel Day (YYYYMM) |
| TRAVDAY | [NHTS] Travel day - day of week |
| DRVRCNT | [NHTS] Number of drivers in HH |
| HHSIZE | [NHTS] Count of HH members |
| HHVEHCNT | [NHTS] Count of HH vehicles |
| NUMADLT | [NHTS] Count of adult HHMs at least 18 years old |
| TRPMILES | [NHTS] Calculated Trip distance converted into miles |
| TRPTRANS | [NHTS] Transportation mode used on trip |
| TRIPPURP | [NHTS] General Trip Purpose (Home-Based Purpose types) |
| TRVL\_MIN | [NHTS] Derived trip time - minutes |
| TRVLCMIN | [NHTS] Calculated travel time |
| DVMT | [NHTS] Calculated Trip distance (miles) for Driver Trips |
| WRKCOUNT | [NHTS] Number of workers in HH |
| LIF\_CYC | [NHTS] Household Life Cycle (Single, Young Couple, Couple with children, Empty Nester) |
| Htppopdn |  |
| TOTPOP10\_1 | [PT] 2010 Total population within 1 mile buffer of BG centroid |
| EMPTOT\_2 | [PT] Total employment within 2 mile buffer of BG centroid |
| [PT] |
| E5\_RET10 | [SLD] 2010 Retail employment |
| E5\_SVC10 | [SLD] 2010 Service employment |
| D1D | [SLD] Gross activity density (employment + HUs) on unprotected land |
| D2A\_JPHH | [SLD] Jobs per household |
| D3amm | [SLD] Network density in terms of facility miles of multi-modal links per square mile |
| D3apo | [SLD] Network density in terms of facility miles of pedestrian-oriented links per square mile |
| D4a | [SLD] Distance from population weighted centroid to nearest transit stop (meters) |
| D4c | [SLD] Aggregate frequency of transit service within 0.25 miles of block group boundary per hour during evening peak period |
| D4d | [SLD] Aggregate frequency of transit service (D4c) per square mile |
| Fwylnmicap | [TTI] 2010 Urbanized Area freeway lane miles per capita |
| Tranmilescap | [NTD] 2009 Urbanized Area annual vehicle revenue miles per capita |
| ACCESS | [Place Type] Accessibility measure ACCESS = (2 \* EMPTOT\_2 \* TOTPOP10\_5) / 10000 \* (EMPTOT\_2 + TOTPOP10\_5), where EMPTOT\_2 is employment within 2-mile radius, and TOTPOP10\_5 is total 2010 population within 5-mile radius |

Table 2.7 Trip frequencies by mode (unweighted)

|  |  |  |
| --- | --- | --- |
| **mode** | **n** | **%** |
| Auto | 955345 | 88.477 |
| Walk | 93182 | 8.630 |
| Transit | 22483 | 2.082 |
| Bike | 8753 | 0.811 |

Table 2.8 Shares of trips by trip purpose and mode

|  |  |  |  |
| --- | --- | --- | --- |
| **TRIPPURP** | **mode** | **n** | **%** |
| HBO | Auto | 195189 | 84.670 |
| HBO | Bike | 1023 | 0.444 |
| HBO | Transit | 13157 | 5.707 |
| HBO | Walk | 21161 | 9.179 |
| HBSHOP | Auto | 243832 | 95.249 |
| HBSHOP | Bike | 1097 | 0.429 |
| HBSHOP | Transit | 1251 | 0.489 |
| HBSHOP | Walk | 9814 | 3.834 |
| HBSOCREC | Auto | 110582 | 71.482 |
| HBSOCREC | Bike | 4832 | 3.123 |
| HBSOCREC | Transit | 812 | 0.525 |
| HBSOCREC | Walk | 38473 | 24.870 |
| HBW | Auto | 102319 | 95.909 |
| HBW | Bike | 684 | 0.641 |
| HBW | Transit | 1671 | 1.566 |
| HBW | Walk | 2009 | 1.883 |
| NHB | Auto | 303423 | 91.432 |
| NHB | Bike | 1117 | 0.337 |
| NHB | Transit | 5592 | 1.685 |
| NHB | Walk | 21725 | 6.546 |

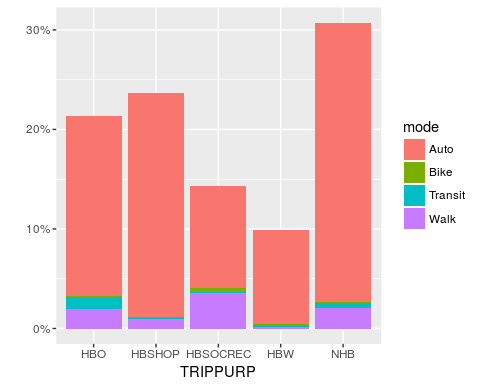


Figure 2.1 Shares of trips by trip purpose and mode

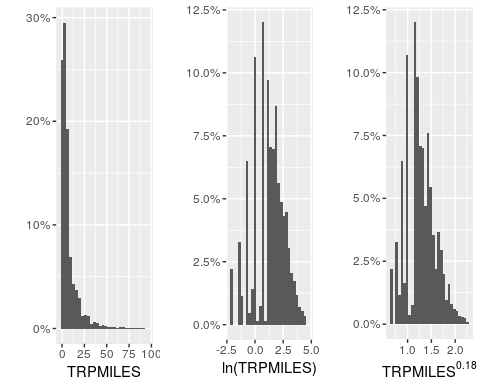
The distribution of raw trip distance (miles) is very skewed 

Figure 2.2 The distribution of raw trip distance (miles)

Table 2.9 Summary of trip distance by mode

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **mode** | **n** | **5%** | **25%** | **50%** | **75%** | **95%** | **99%** | **max** | **mean** | **sd** |
| Auto | 955345 | 0.556 | 2.000 | 4.0 | 10.000 | 29 | 57 | 91 | 8.027 | 10.831 |
| Bike | 8753 | 0.111 | 0.556 | 1.0 | 2.889 | 8 | 17 | 22 | 2.211 | 3.113 |
| Transit | 22483 | 0.556 | 2.000 | 4.0 | 9.000 | 26 | 55 | 95 | 7.727 | 10.301 |
| Walk | 93182 | 0.111 | 0.222 | 0.5 | 0.778 | 2 | 3 | 4 | 0.646 | 0.612 |

Exclude trips by vehicles w/ commercial license plates and with distance above 99 percentile

Given the skewedness of trip distance, a cutoff of 99-percentile of trip distance for each mode is used. Results below are after applying the cutoff.

Table 2.10 Descriptive of total household travel distance (miles) by mode used:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| mode | n | 5% | 25% | 50% | 75% | 95% | 99% | max | mean | sd |
| Auto | 127999 | 4.000 | 17.000 | 40.00 | 80.00 | 183.00 | 308.12 | 1205.0 | 59.91 | 64.91 |
| Bike | 3412 | 0.222 | 1.111 | 3.00 | 7.00 | 20.00 | 37.10 | 76.0 | 5.67 | 7.61 |
| Transit | 9107 | 1.000 | 4.000 | 10.00 | 22.00 | 66.00 | 130.00 | 434.0 | 19.08 | 27.31 |
| Walk | 32780 | 0.222 | 0.556 | 1.11 | 2.22 | 5.44 | 8.67 | 40.2 | 1.84 | 1.88 |

Table 2.11 Descriptive of household travel time (minutes) by mode used:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| mode | n | 5% | 25% | 50% | 75% | 95% | 99% | max | mean | sd |
| Auto | 127999 | 19 | 50 | 97 | 167 | 325 | 500 | 2084 | 124.9 | 105.6 |
| Bike | 3412 | 5 | 19 | 30 | 60 | 140 | 240 | 515 | 48.6 | 49.6 |
| Transit | 9107 | 13 | 31 | 60 | 106 | 220 | 380 | 1155 | 82.7 | 79.5 |
| Walk | 32780 | 4 | 15 | 30 | 55 | 118 | 196 | 1110 | 40.6 | 42.4 |

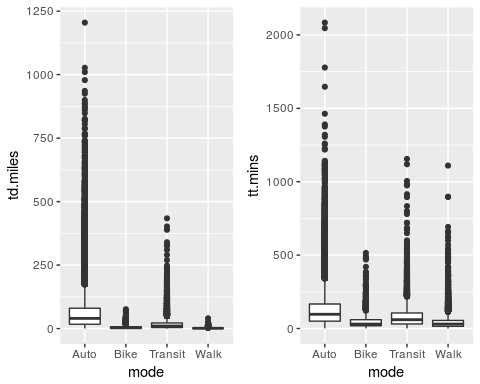


Figure 2.3 Boxplot of household travel distance (mile) and time (minutes) by mode

**Survey day VMT and Annual vehicle miles**

* DVMT - Calculated Trip distance (miles) for Auto Trips
* ANNMILES - Self-reported annualized mile estimate;
* BESTMILE - Best estimate of annual miles (by ORNL)

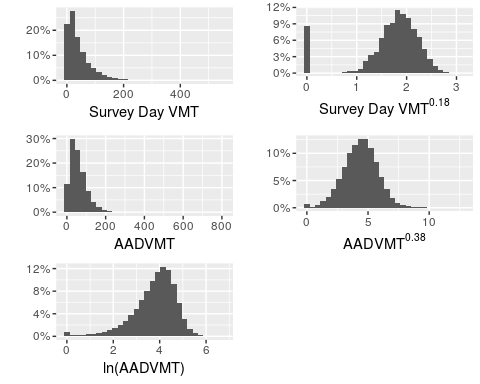


Figure 2.4 Distribution of VMT after transformation

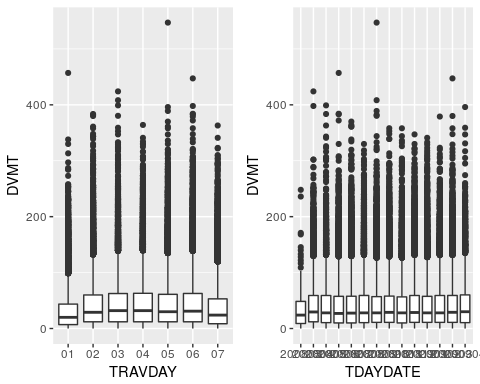


Figure 2.5 Temporal distribution of travel day

## SmartLocation Database (SLD)

The Smart Location Database (Ramsey and Bell, 2014) is a nationwide geographic data resource for measuring location efficiency. It includes more than 90 attributes summarizing characteristics such as housing density, diversity of land use, neighborhood design, destination accessibility, transit service, employment, and demographics. Most attributes are available for every census block group in the United States. The variables in SLD are largely organized according to the 5D built environment measures: Density, Diversity, Design, Transit, Destination, in addition to demographics and employment. A complete list of the variables can be found [here](https://www.epa.gov/sites/production/files/2014-03/documents/sld_userguide.pdf).

The confidential NHTS data contain Census Block Group information of households' residence Census block group (2010 geography), which is joined with SLD to retrieve land use features for these locations. Land use information in SLD provide a rich set of factors that are documented in existing research literature to have influence on households' travel behavior including mode choices and travel distance.

All households in the 2009 NHTS data have a matched block group in the SLD.

## Place Types

Place types are land uses categories that are useful for describing development patterns and their relationship to human behavior (e.g. travel behavior) and well being (e.g. health) (Gregor, 2016). In the RSPM mode shift project, we use place types as a means to simplify the work for RSPM users when they create scenarios.

This project adopts the work by Brian Gregor and others and establishes categories over the following 3 dimensions:

* (flag) Location Type: categorizes the the general urban context of the place (e.g. large urbanized area, small city, etc.).

1. Urbanized: A contiguous area of urban development which has a large population. Criteria: population within 5 miles >= 30,000 and population within 1 Mile >= 1,000;
2. Urban Near Urbanized: Urban development (e.g. cities, towns, communities) located in the fringe of an urbanized area but are not part of the contiguous urbanized area. Criteria: Population within 15 Miles >= 60,000 and Population within 2 Mile >= 2,000;
3. Rural Near Urbanized: Urban development not located in the fringe of an urbanized area. Criteria: Population within 15 Miles >= 60,000 and Population within 2 Mile < 2,000
4. Urban Not Near Urbanized: Urban development not located in the fringe of an urbanized area. Criteria: Population within 15 Miles <= 60,000 and Population within 2 Mile >= 2,000
5. Rural Not Near Urbanized: Rural development not located in the fringe of an urbanized area. Criteria: Population within 15 Miles <= 60,000 and Population within 2 Mile <= 2,000

* Area Type: categorizes the spatial relationship of urban places to the urban center (e.g. urban center, suburbs, etc.).

1. Regional Center: Places within urbanized areas that have high levels of population accessibility to jobs and developed at densities and having transportation networks that would allow a substantial portion of the population to get to jobs or other activities by non-auto transport modes. Criteria: if ACCESS is high, and DENSITY is medium or high, and DESIGN is high;
2. Close In Community: Places within urbanized areas and other urban areas that are located near regional centers or are places with relatively high levels of population accessibility to jobs within urban areas that are not urbanized. Criteria: if ACCESS is high, and DENSITY is medium or high, but DESIGN is not high, or if ACCESS is high and DENSITY is low, or if ACCESS is medium and DENSITY is medium or high;
3. Suburb/Town: Places in urbanized areas, smaller urban areas, and towns that have lower population accessibility to jobs. Criteria: if ACCESS is high but DENSITY is very low, or if ACCESS is very low or low and DENSITY is not very low;
4. Low Density/Rural: Low density places with low job accessibiity located primarily in rural areas, but may occassionally be found in large vacant tracts in urbanized areas. Criteria: in all other cases.

* Development Type: categorizes the general character of land uses occupying the place (e.g. residential, employment, mixed, etc.)

1. Low Density/Rural: These are places that have very low density development in urban or rural areas. In urban areas these can include large tracts of park land or greenfields. Criteria: if DENSITY is very low;
2. Employment: These are places where there are more jobs than households and do not qualify as mixed-use as described below. Criteria: if not Mixed and Diversity1 is greater than 1 (i.e. more jobs than households);
3. Residential: These are places where there are more households than jobs and do not qualify as mixed-use as described below. Criteria: if not Mixed and Diversity1 is less than 1 (i.e. more households than jobs);
4. Mixed: These are places where there are a mixture of jobs and households that meet a specified ratio of the two uses. Criteria: if DIVERSITY is high and DENSITY is medium or high and DESIGN is medium or high;
5. Mixed High: These are places that are mixed and have relatively high densities. Critera: if Mixed and DENSITY is high and DESIGN is high;
6. Transit-Oriented Development (TOD): These are places that are mixed, have relatively high densities, and have relatively high levels of public transit service. Criteria: if Mixed High and TRANSIT is high, or if Employment and TRANSIT is high and DESIGN is high.

By default, the accessibility measure ACCESS = (2 \* EMPTOT\_2 \* TOTPOP10\_5) / 10000 \* (EMPTOT\_2 + TOTPOP10\_5), where EMPTOT\_2 is employment within 2-mile radius, and TOTPOP10\_5 is total 2010 population within 5-mile radius. The break points for very low, low, medium, and high are 0.1, 0.5 and 2, respectively.

The Density level uses D1D variable in SLD - gross activity density (employment + HUs) on unprotected land (per acre) - with break points of 0.1, 1, and 5.

The Design measure is based on two variables from the SLD: D3amm variable (network density in terms of multimodal links per square mile) and D3apo variable (network density in terms of facility miles of pedestrian-oriented links per square mile). The default break points for D3amm are 1.3, 2.5, and 3.3, while those for D3apo are 12.5, 15.6, and 20. The final value of the Design measure is the maximum value of the two. For example, if the D3amm value is low and D3apo value is medium, the final value of the design measure would be medium.

Diversity Level is a measure of the mixing of jobs and households in the block group. It is based on measures in the SLD: D2A\_JPHH (ratio of jobs to households in the block group and the ratio of retail and service jobs to the number of households (E5\_RET10 + E5\_SVC10)/HH.

Transit Level is a measure of the level of transit service derived from the SLD D4c (aggregate frequency of transit service within 0.25 miles of block group boundary per hour during evening peak period). The threshold values for the 4 levels are 1, 20, and 150.

Based on discussion with the TAC, in particular, Brian and Tara, we primilarily use the place types as an intermediate step to faciliate scenario creation, but not as independent variables directly included in model specification.

## Model Structures

### Current GreenSTEP DVMT models

GreenSTEP focuses on Daily Vehicle Mile Travel (VMT) by drivers in its household travel model and does not explicitly models non-driving travel (for example, by transit or non-motorized modes), except for diversion of short-distance trips to bike. The current household travel model in GreenSTEP has two sequential (conditional) model: a binary model of whether a household will have non-zero VMT and a regression model of the actual VMT for households with non-zero VMT. Such a model structure provides a good balance between behavioral realism and simplicity and performance:

1. ZeroDVMT model

, and

Table 2.12 Binomial Logit Models of Zero DVMT

|  |  |  |
| --- | --- | --- |
|  | metro | Nonmetro |
| DrvAgePop | 0.065\*\*\*  (0.019) | -0.070\*\*\*  (0.021) |
| LogIncome | -0.453\*\*  (0.022) | -0.435\*\*\*  (0.020) |
| HTPPOPDN | -0.003  (0.008) | 0.002  (0.008) |
| Age65Plus | -0.101\*\*\*  (0.026) | -0.013  (0.022) |
| HHVEHCNT | -0.522\*\*\*  (0.029) | -0.304\*\*\*  (0.022) |
| ZeroVeh | 3.730\*\*\*  (0.094) | 3.680\*\*\*  (0.091) |
| Tranmilescap | 0.023\*\*\*  (0.001) |  |
| Constant | 2.650\*\*\*  (0.224) | 2.520\*\*\*  (0.202) |
| Observations | 53.461 | 70.324 |
| Log Likelihood | -11,856.000 | -14,397.000 |
| Akaike Inf. Crit. | 23,728.000 | 28,808.000 |

Note: p<0.1; **p<0.05;** p<0.01

Table 2.13 ZeroDVMT model goodness-of-fit

|  |  |  |
| --- | --- | --- |
| metro | auc | pseudo.r2 |
| metro | 0.806 | 0.306 |
| non\_metro | 0.743 | 0.198 |

1. DVMT model

Table 2.14 **Power-transformed Regression Models of DVMT (DVMT > 0)**

|  |  |  |
| --- | --- | --- |
|  | metro | nonmetro |
| CENSUS\_RNE | -0.004  (0.008) | 0.051\*\*\*  (0.005) |
| CENSUS\_RS | 0.011\*  (0.006) | 0.054\*\*\*  (0.004) |
| CENSUS\_RW | -0.008  (0.006) | 0.015\*\*\*  (0.005) |
| LogIncome | 0.081\*\*\*  (0.002) | 0.076\*\*\*  (0.002) |
| HTPPOPDN | -0.004\*\*\*  (0.001) | 0.004\*\*\*  (0.001) |
| HHVEHCNT | 0.058\*\*\*  (0.002) | 0.054\*\*\*  (0.001) |
| ZeroVeh | 0.036  (0.024) | 0.066\*\*\*  (0.025) |
| Tranmilescap | -0.001\*\*\*  (0.0003) |  |
| Fwylnmicap | 0.029\*\*\*  (0.006) |  |
| DrvAgePop | 0.045\*\*\*  (0.002) | 0.054\*\*\*  (0.002) |
| Age65Plus | -0.049\*\*\*  (0.002) | -0.052\*\*\*  (0.002) |
| HTPPOPDN:Tranmilescap | 0.0001\*  (0.0001) |  |
| Constant | 0.765\*\*\*  (0.023) | 0.781\*\*\*  (0.019) |
| Observation | 48,249 | 65,356 |
| R2 | 0.170 | 0.163 |
| Adjusted R2 | 0.169 | 0.163 |
| Residual Std. Error | 0.307 (df=48236) | 0.327 (df=65346) |
| F Statistic | 821.000\*\*\* (df=12; 48236) | 1,416.000\*\*\* (df=9; 65346) |

Note: p<0.1; **p<0.05;** p<0.01

Table 2.15 DVMT model goodness-of-fit

|  |  |  |  |
| --- | --- | --- | --- |
| metro | rmse | nrmse | r.squared |
| metro | 37.2 | 37.2 | 0.170 |
| non\_metro | 44.6 | 44.6 | 0.163 |

1. Combined model

Table 2.16 Combined model goodness-of-fit

|  |  |  |
| --- | --- | --- |
| metro | rmse | nrmse |
| metro | 55.0 | 1.47 |
| non\_metro | 65.3 | 1.46 |

Another related model in GreenSTEP is the household budget model that captures the price elasticity of travel. The budget approach to modeling is based on the perspective that households make their travel decisions within money and time budget constraints. According to Brian's research on historical consumer expenditure survey data, household spending on gasoline and other variable costs is done within a household transportation budget that is relatively stable, as households shift expenses between transportation budget categories when gasoline prices fluctuate. Households will necessarily reduce their travel in direct proportion to the cost increase only when fuel prices or other variable costs increase to the point where it is no longer possible to shift money from other parts of the transportation budget (Gregor, 2010). Brian assumes the transition between inelastic and elastic behavior will not be abrupt unless there is little time for the household to recognize the impact of the cost increases on the budget or respond to the cost increases. If the changes are more gradual, the transition will be less abrupt. Given the focus of GreenSTEP/RSPM on long term forecasting, we would only need to model long run elasticities.

### Proposed New Models

#### AADVMT Model (Power-transformed linear regression model)

Instead of modeling DVMT and then approximating annual VMT from it, an alternative is to directly model annual average daily VMT (AADVMT). Both 2001 and 2009 NHTS contain annual mile estimates provided by ORNL, from which we can derive AADVMT.

ln(AADVMT) = f(HH variables, 5D variables)

1. Estimated Parameters

Table 2.17 **Power-transformed Regression Models of AADVMT**

|  |  |  |
| --- | --- | --- |
|  | metro | nonmetro |
| DrvAgePop |  | 0.274\*\*\*  (0.009) |
| HHSIZE | 0.261\*\*\*  (0.005) | 0.113\*\*\*  (0.007) |
| WRKCOUNT | 0.407\*\*\*  (0.007) | 0.349\*\*\*  (0.006) |
| CENSUS\_RNE |  | -0.124\*\*\*  (0.015) |
| CENSUS\_RS |  | 0.074\*\*\*  (0.012) |
| CENSUS\_RW |  | -0.132\*\*\*  (0.016) |
| LogIncome | 0.357\*\*\*  (0.007) | 0.392\*\*\*  (0.005) |
| Age65Plus | -0.024\*\*\*  (0.008) | -0.037\*\*\*  (0.006) |
| ns(log1p(VehPerDriver), 3)1 | 2.120\*\*\*  (0.052) | 2.320\*\*\*  (0.041) |
| ns(log1p(VehPerDriver), 3)2 | 4.330\*\*\*  (0.215) | 4.140\*\*\*  (0.167) |
| ns(log1p(VehPerDriver), 3)3 | 2.120\*\*\*  (0.236) | 2.570\*\*\*  (0.159) |
| log1p(TRPOPDEN) | -0.047\*\*\*  (0.011) | -0.059\*\*\*  (0.011) |
| log1p(EMPTOT\_5) | -0.074\*\*\*  (0.006) | -0.049\*\*\*  (0.003) |
| Tranmilescap | -0.003\*\*\*  (0.0005) |  |
| D1B | -0.002\*\*\*  (0.0004) | 0.008\*\*\*  (0.003) |
| D3bpo4 | -0.0004\*  (0.0002) |  |
| D2A\_EPHHM |  | 0.070\*\*\*  (0.024) |
| ACCESS | 0.002  (0.002) | -0.027\*\*\*  (0.008) |
| Tranmilescap:D4c | -0.00001\*\*\*  (0.00000) |  |
| D1B:D2A\_EPHHM |  | -0.026\*\*\*  (0.005) |
| Constant | -1.730\*\*\*  (0.122) | -2.280\*\*\*  (0.089) |
| Observations | 41,497 | 73,899 |
| R2 | 0.390 | 0.417 |
| Adjusted R2 | 0.390 | 0.417 |
| Residual Std. Error | 1.040 (df=41482) | 1.040 (df=73881) |
| F Statistic | 1,893.000\*\*\* (df=14; 41482) | 3,107.000\*\*\* (df=17; 73881) |

Note: p<0.1; **p<0.05;** p<0.01

1. Validation

Table 2.18 Validation of AADVMT Models

|  |  |  |  |
| --- | --- | --- | --- |
| metro | rmse | nrmse | r.squared |
| metro | 30.7 | 0.573 | 0.390 |
| non\_metro | 34.2 | 0.566 | 0.417 |

#### AADVMT Model (hurdle model)

1. Estimated Parameters

## $metro  
##   
## Call:  
## pscl::hurdle(formula = AADVMT.int ~ HHSIZE + WRKCOUNT + CENSUS\_R +   
## LogIncome + Age65Plus + ns(log1p(VehPerDriver), 3) + log1p(EMPTOT\_5) +   
## D1D + D1B:D2A\_EPHHM + Tranmilescap + Tranmilescap:D4c + D3bpo4 +   
## ACCESS | HHSIZE + WRKCOUNT + LogIncome + Age65Plus + ns(log1p(VehPerDriver),   
## 3) + log1p(TRPOPDEN) + D1D + D1B:D2A\_EPHHM, data = ., na.action = na.omit,   
## dist = "negbin")  
##   
## Pearson residuals:  
## Min 1Q Median 3Q Max   
## -1.663 -0.686 -0.162 0.487 10.175   
##   
## Count model coefficients (truncated negbin with log link):  
## Estimate Std. Error z value  
## (Intercept) 1.01205323 0.07871768 12.86  
## HHSIZE 0.16429554 0.00284087 57.83  
## WRKCOUNT 0.20046525 0.00430530 46.56  
## CENSUS\_RNE -0.06553070 0.01627132 -4.03  
## CENSUS\_RS 0.02571395 0.01375002 1.87  
## CENSUS\_RW -0.03530183 0.01382868 -2.55  
## LogIncome 0.17903454 0.00419696 42.66  
## Age65Plus -0.02988608 0.00452710 -6.60  
## ns(log1p(VehPerDriver), 3)1 1.17448224 0.03150512 37.28  
## ns(log1p(VehPerDriver), 3)2 1.89891522 0.14345867 13.24  
## ns(log1p(VehPerDriver), 3)3 1.01860008 0.14084883 7.23  
## log1p(EMPTOT\_5) -0.04421578 0.00329058 -13.44  
## D1D -0.00047001 0.00020798 -2.26  
## Tranmilescap -0.00078248 0.00034049 -2.30  
## D3bpo4 -0.00033312 0.00012605 -2.64  
## ACCESS 0.00215679 0.00121729 1.77  
## D1B:D2A\_EPHHM -0.00156197 0.00056375 -2.77  
## Tranmilescap:D4c -0.00000713 0.00000236 -3.02  
## Log(theta) 1.05398938 0.00753693 139.84  
## Pr(>|z|)   
## (Intercept) < 0.0000000000000002 \*\*\*  
## HHSIZE < 0.0000000000000002 \*\*\*  
## WRKCOUNT < 0.0000000000000002 \*\*\*  
## CENSUS\_RNE 0.00005640306333 \*\*\*  
## CENSUS\_RS 0.0615 .   
## CENSUS\_RW 0.0107 \*   
## LogIncome < 0.0000000000000002 \*\*\*  
## Age65Plus 0.00000000004068 \*\*\*  
## ns(log1p(VehPerDriver), 3)1 < 0.0000000000000002 \*\*\*  
## ns(log1p(VehPerDriver), 3)2 < 0.0000000000000002 \*\*\*  
## ns(log1p(VehPerDriver), 3)3 0.00000000000048 \*\*\*  
## log1p(EMPTOT\_5) < 0.0000000000000002 \*\*\*  
## D1D 0.0238 \*   
## Tranmilescap 0.0216 \*   
## D3bpo4 0.0082 \*\*   
## ACCESS 0.0764 .   
## D1B:D2A\_EPHHM 0.0056 \*\*   
## Tranmilescap:D4c 0.0025 \*\*   
## Log(theta) < 0.0000000000000002 \*\*\*  
## Zero hurdle model coefficients (binomial with logit link):  
## Estimate Std. Error z value  
## (Intercept) -8.05823 0.65520 -12.30  
## HHSIZE -0.12906 0.03851 -3.35  
## WRKCOUNT 1.19329 0.09631 12.39  
## LogIncome 0.71847 0.05045 14.24  
## Age65Plus 0.46273 0.07781 5.95  
## ns(log1p(VehPerDriver), 3)1 1.35483 0.51821 2.61  
## ns(log1p(VehPerDriver), 3)2 13.97910 1.93649 7.22  
## ns(log1p(VehPerDriver), 3)3 9.41377 3.39600 2.77  
## log1p(TRPOPDEN) -0.48856 0.07674 -6.37  
## D1D 0.00880 0.00482 1.82  
## D1B:D2A\_EPHHM -0.02015 0.00826 -2.44  
## Pr(>|z|)   
## (Intercept) < 0.0000000000000002 \*\*\*  
## HHSIZE 0.0008 \*\*\*  
## WRKCOUNT < 0.0000000000000002 \*\*\*  
## LogIncome < 0.0000000000000002 \*\*\*  
## Age65Plus 0.00000000273370 \*\*\*  
## ns(log1p(VehPerDriver), 3)1 0.0089 \*\*   
## ns(log1p(VehPerDriver), 3)2 0.00000000000052 \*\*\*  
## ns(log1p(VehPerDriver), 3)3 0.0056 \*\*   
## log1p(TRPOPDEN) 0.00000000019327 \*\*\*  
## D1D 0.0681 .   
## D1B:D2A\_EPHHM 0.0147 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1   
##   
## Theta: count = 2.869  
## Number of iterations in BFGS optimization: 32   
## Log-likelihood: -1.95e+05 on 30 Df  
##   
## $non\_metro  
##   
## Call:  
## pscl::hurdle(formula = AADVMT.int ~ HHSIZE + WRKCOUNT + CENSUS\_R +   
## LogIncome + Age65Plus + ns(log1p(VehPerDriver), 3) + log1p(TRPOPDEN) +   
## log1p(EMPTOT\_5) + D1D + D1B:D2A\_EPHHM + ACCESS | WRKCOUNT +   
## CENSUS\_R + LogIncome + Age65Plus + ns(log1p(VehPerDriver), 3) +   
## log1p(EMPTOT\_5) + D1B:D2A\_EPHHM + ACCESS, data = ., na.action = na.omit,   
## dist = "negbin")  
##   
## Pearson residuals:  
## Min 1Q Median 3Q Max   
## -1.705 -0.685 -0.159 0.479 11.759   
##   
## Count model coefficients (truncated negbin with log link):  
## Estimate Std. Error z value  
## (Intercept) 0.939339 0.055473 16.93  
## HHSIZE 0.173281 0.002258 76.75  
## WRKCOUNT 0.183136 0.003206 57.12  
## CENSUS\_RNE -0.057008 0.008371 -6.81  
## CENSUS\_RS 0.053989 0.006653 8.12  
## CENSUS\_RW -0.045698 0.008971 -5.09  
## LogIncome 0.200394 0.002981 67.23  
## Age65Plus -0.015876 0.003241 -4.90  
## ns(log1p(VehPerDriver), 3)1 1.144715 0.025061 45.68  
## ns(log1p(VehPerDriver), 3)2 1.404949 0.107054 13.12  
## ns(log1p(VehPerDriver), 3)3 1.024127 0.089391 11.46  
## log1p(TRPOPDEN) -0.034741 0.005444 -6.38  
## log1p(EMPTOT\_5) -0.024967 0.001603 -15.57  
## D1D -0.000789 0.000958 -0.82  
## ACCESS -0.013824 0.005239 -2.64  
## D1B:D2A\_EPHHM -0.004449 0.002213 -2.01  
## Log(theta) 1.091443 0.005492 198.75  
## Pr(>|z|)   
## (Intercept) < 0.0000000000000002 \*\*\*  
## HHSIZE < 0.0000000000000002 \*\*\*  
## WRKCOUNT < 0.0000000000000002 \*\*\*  
## CENSUS\_RNE 0.00000000000973672 \*\*\*  
## CENSUS\_RS 0.00000000000000048 \*\*\*  
## CENSUS\_RW 0.00000035091456787 \*\*\*  
## LogIncome < 0.0000000000000002 \*\*\*  
## Age65Plus 0.00000096504772293 \*\*\*  
## ns(log1p(VehPerDriver), 3)1 < 0.0000000000000002 \*\*\*  
## ns(log1p(VehPerDriver), 3)2 < 0.0000000000000002 \*\*\*  
## ns(log1p(VehPerDriver), 3)3 < 0.0000000000000002 \*\*\*  
## log1p(TRPOPDEN) 0.00000000017563824 \*\*\*  
## log1p(EMPTOT\_5) < 0.0000000000000002 \*\*\*  
## D1D 0.4103   
## ACCESS 0.0083 \*\*   
## D1B:D2A\_EPHHM 0.0444 \*   
## Log(theta) < 0.0000000000000002 \*\*\*  
## Zero hurdle model coefficients (binomial with logit link):  
## Estimate Std. Error z value  
## (Intercept) -7.8887 0.5919 -13.33  
## WRKCOUNT 1.1217 0.0903 12.43  
## CENSUS\_RNE -0.4442 0.1889 -2.35  
## CENSUS\_RS -0.3819 0.1579 -2.42  
## CENSUS\_RW -0.5114 0.1919 -2.66  
## LogIncome 0.7016 0.0461 15.22  
## Age65Plus 0.3877 0.0640 6.06  
## ns(log1p(VehPerDriver), 3)1 2.1057 0.4062 5.18  
## ns(log1p(VehPerDriver), 3)2 13.3883 1.3937 9.61  
## ns(log1p(VehPerDriver), 3)3 7.3197 2.3514 3.11  
## log1p(EMPTOT\_5) -0.0827 0.0260 -3.18  
## ACCESS -0.0664 0.0611 -1.09  
## D1B:D2A\_EPHHM -0.0459 0.0249 -1.84  
## Pr(>|z|)   
## (Intercept) < 0.0000000000000002 \*\*\*  
## WRKCOUNT < 0.0000000000000002 \*\*\*  
## CENSUS\_RNE 0.0187 \*   
## CENSUS\_RS 0.0156 \*   
## CENSUS\_RW 0.0077 \*\*   
## LogIncome < 0.0000000000000002 \*\*\*  
## Age65Plus 0.0000000014 \*\*\*  
## ns(log1p(VehPerDriver), 3)1 0.0000002179 \*\*\*  
## ns(log1p(VehPerDriver), 3)2 < 0.0000000000000002 \*\*\*  
## ns(log1p(VehPerDriver), 3)3 0.0019 \*\*   
## log1p(EMPTOT\_5) 0.0015 \*\*   
## ACCESS 0.2770   
## D1B:D2A\_EPHHM 0.0653 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1   
##   
## Theta: count = 2.979  
## Number of iterations in BFGS optimization: 25   
## Log-likelihood: -3.55e+05 on 30 Df

1. Model goodness-of-fit and Validation

Table 2.19 AADVMT model (Hurdle model) goodness-of-fit and validation

|  |  |  |  |
| --- | --- | --- | --- |
| metro | rmse | nrmse | pseudo.r2 |
| metro | 30.9 | 0.577 | 0.0445 |
| non\_metro | 35.0 | 0.579 | 0.0458 |

The AADVMT hurdle model adds more complexity yet brings little benefits in terms of prediction accuracies (rmse) or sensitivities, thus the AADVMT power-transformed model is preferred.

#### Person Miles Traveled by Mode Models

##### Transit Miles Traveled Model (hurdle model)

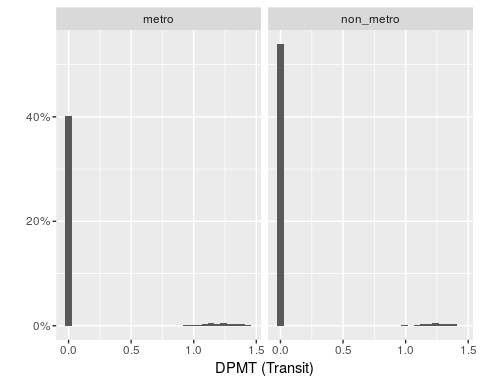


Figure 2.16 Transit miles distribution

1. Estimated Parameters

* counterintuitive coefficients D1D, D2A\_EPHHM

## $metro  
##   
## Call:  
## pscl::hurdle(formula = int\_round(td.miles.Transit) ~ log1p(VehPerDriver) +   
## HHSIZE + WRKCOUNT + LIF\_CYC + Age0to14 + LogIncome + D1D + D2A\_EPHHM +   
## Fwylnmicap + Tranmilescap + D4c | AADVMT + VehPerDriver + HHSIZE +   
## WRKCOUNT + LIF\_CYC + Age0to14 + D1D + D3bmm4 + Fwylnmicap +   
## Tranmilescap:D4c + LogIncome, data = .)  
##   
## Pearson residuals:  
## Min 1Q Median 3Q Max   
## -3.221 -0.268 -0.153 -0.117 27.418   
##   
## Count model coefficients (truncated poisson with log link):  
## Estimate Std. Error z value  
## (Intercept) 1.9212620 0.0995875 19.29  
## log1p(VehPerDriver) -0.1291904 0.0355742 -3.63  
## HHSIZE 0.0653538 0.0063841 10.24  
## WRKCOUNT 0.0268560 0.0084115 3.19  
## LIF\_CYCCouple w/o children 0.2762891 0.0214246 12.90  
## LIF\_CYCEmpty Nester 0.2650850 0.0240630 11.02  
## LIF\_CYCSingle 0.3283556 0.0369382 8.89  
## Age0to14 -0.0193269 0.0086041 -2.25  
## LogIncome 0.0239053 0.0084591 2.83  
## D1D -0.0014416 0.0001659 -8.69  
## D2A\_EPHHM -0.0562551 0.0272781 -2.06  
## Fwylnmicap -0.1631917 0.0418730 -3.90  
## Tranmilescap 0.0027340 0.0004983 5.49  
## D4c 0.0002471 0.0000742 3.33  
## Pr(>|z|)   
## (Intercept) < 0.0000000000000002 \*\*\*  
## log1p(VehPerDriver) 0.00028 \*\*\*  
## HHSIZE < 0.0000000000000002 \*\*\*  
## WRKCOUNT 0.00141 \*\*   
## LIF\_CYCCouple w/o children < 0.0000000000000002 \*\*\*  
## LIF\_CYCEmpty Nester < 0.0000000000000002 \*\*\*  
## LIF\_CYCSingle < 0.0000000000000002 \*\*\*  
## Age0to14 0.02469 \*   
## LogIncome 0.00471 \*\*   
## D1D < 0.0000000000000002 \*\*\*  
## D2A\_EPHHM 0.03918 \*   
## Fwylnmicap 0.000097270 \*\*\*  
## Tranmilescap 0.000000041 \*\*\*  
## D4c 0.00087 \*\*\*  
## Zero hurdle model coefficients (binomial with logit link):  
## Estimate Std. Error z value  
## (Intercept) -0.9625233 0.3684409 -2.61  
## AADVMT -0.0037454 0.0006157 -6.08  
## VehPerDriver -0.7904598 0.0731318 -10.81  
## HHSIZE 0.0620903 0.0249369 2.49  
## WRKCOUNT 0.3095209 0.0307155 10.08  
## LIF\_CYCCouple w/o children -1.1173552 0.0773660 -14.44  
## LIF\_CYCEmpty Nester -1.6122166 0.0893579 -18.04  
## LIF\_CYCSingle -1.3448620 0.1312501 -10.25  
## Age0to14 0.4511670 0.0310501 14.53  
## D1D 0.0060544 0.0008202 7.38  
## D3bmm4 0.0030048 0.0018077 1.66  
## Fwylnmicap -0.2489311 0.1327192 -1.88  
## LogIncome -0.0744235 0.0313442 -2.37  
## Tranmilescap:D4c 0.0001073 0.0000167 6.43  
## Pr(>|z|)   
## (Intercept) 0.009 \*\*   
## AADVMT 0.00000000117985 \*\*\*  
## VehPerDriver < 0.0000000000000002 \*\*\*  
## HHSIZE 0.013 \*   
## WRKCOUNT < 0.0000000000000002 \*\*\*  
## LIF\_CYCCouple w/o children < 0.0000000000000002 \*\*\*  
## LIF\_CYCEmpty Nester < 0.0000000000000002 \*\*\*  
## LIF\_CYCSingle < 0.0000000000000002 \*\*\*  
## Age0to14 < 0.0000000000000002 \*\*\*  
## D1D 0.00000000000016 \*\*\*  
## D3bmm4 0.096 .   
## Fwylnmicap 0.061 .   
## LogIncome 0.018 \*   
## Tranmilescap:D4c 0.00000000013145 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1   
##   
## Number of iterations in BFGS optimization: 30   
## Log-likelihood: -2.06e+04 on 28 Df  
##   
## $non\_metro  
##   
## Call:  
## pscl::hurdle(formula = int\_round(td.miles.Transit) ~ log1p(VehPerDriver) +   
## HHSIZE + WRKCOUNT + LIF\_CYC + LogIncome + UZAPOPDEN + D2A\_EPHHM |   
## AADVMT + VehPerDriver + HHSIZE + WRKCOUNT + LIF\_CYC + Age0to14 +   
## Age65Plus + LogIncome + D3apo + WRKCOUNT + CENSUS\_R, data = .)  
##   
## Pearson residuals:  
## Min 1Q Median 3Q Max   
## -1.1727 -0.1312 -0.0775 -0.0651 61.3687   
##   
## Count model coefficients (truncated poisson with log link):  
## Estimate Std. Error z value  
## (Intercept) 2.35475 0.14358 16.40  
## log1p(VehPerDriver) 0.07368 0.04643 1.59  
## HHSIZE 0.08088 0.00833 9.71  
## WRKCOUNT 0.00169 0.01235 0.14  
## LIF\_CYCCouple w/o children -0.10792 0.05199 -2.08  
## LIF\_CYCEmpty Nester 0.30676 0.04500 6.82  
## LIF\_CYCSingle -0.13659 0.10128 -1.35  
## LogIncome -0.02427 0.01286 -1.89  
## UZAPOPDEN -0.04786 0.00784 -6.11  
## D2A\_EPHHM -0.24876 0.03980 -6.25  
## Pr(>|z|)   
## (Intercept) < 0.0000000000000002 \*\*\*  
## log1p(VehPerDriver) 0.113   
## HHSIZE < 0.0000000000000002 \*\*\*  
## WRKCOUNT 0.891   
## LIF\_CYCCouple w/o children 0.038 \*   
## LIF\_CYCEmpty Nester 0.0000000000093 \*\*\*  
## LIF\_CYCSingle 0.177   
## LogIncome 0.059 .   
## UZAPOPDEN 0.0000000010246 \*\*\*  
## D2A\_EPHHM 0.0000000004118 \*\*\*  
## Zero hurdle model coefficients (binomial with logit link):  
## Estimate Std. Error z value  
## (Intercept) -0.499103 0.492886 -1.01  
## AADVMT -0.000277 0.000651 -0.43  
## VehPerDriver -0.132132 0.076127 -1.74  
## HHSIZE -0.034550 0.035261 -0.98  
## WRKCOUNT 0.123666 0.042565 2.91  
## LIF\_CYCCouple w/o children -2.399990 0.145637 -16.48  
## LIF\_CYCEmpty Nester -3.071090 0.178348 -17.22  
## LIF\_CYCSingle -2.887385 0.262570 -11.00  
## Age0to14 0.535248 0.040731 13.14  
## Age65Plus 0.130595 0.083443 1.57  
## LogIncome -0.151872 0.043401 -3.50  
## D3apo -0.019118 0.005504 -3.47  
## CENSUS\_RNE 0.666180 0.109336 6.09  
## CENSUS\_RS 0.075224 0.089570 0.84  
## CENSUS\_RW -0.190835 0.114349 -1.67  
## Pr(>|z|)   
## (Intercept) 0.31124   
## AADVMT 0.67033   
## VehPerDriver 0.08262 .   
## HHSIZE 0.32717   
## WRKCOUNT 0.00367 \*\*   
## LIF\_CYCCouple w/o children < 0.0000000000000002 \*\*\*  
## LIF\_CYCEmpty Nester < 0.0000000000000002 \*\*\*  
## LIF\_CYCSingle < 0.0000000000000002 \*\*\*  
## Age0to14 < 0.0000000000000002 \*\*\*  
## Age65Plus 0.11757   
## LogIncome 0.00047 \*\*\*  
## D3apo 0.00051 \*\*\*  
## CENSUS\_RNE 0.0000000011 \*\*\*  
## CENSUS\_RS 0.40100   
## CENSUS\_RW 0.09514 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1   
##   
## Number of iterations in BFGS optimization: 35   
## Log-likelihood: -9.83e+03 on 25 Df

1. Model goodness-of-fit and Validation

Table 2.20 Transit Miles Model (hurdle model) performance

|  |  |  |  |
| --- | --- | --- | --- |
| metro | rmse | nrmse | pseudo.r2 |
| metro | 3.37 | 4.36 | 0.440 |
| non\_metro | 2.33 | 4.75 | 0.678 |

##### Walk Miles Traveled Model (hurdle model)

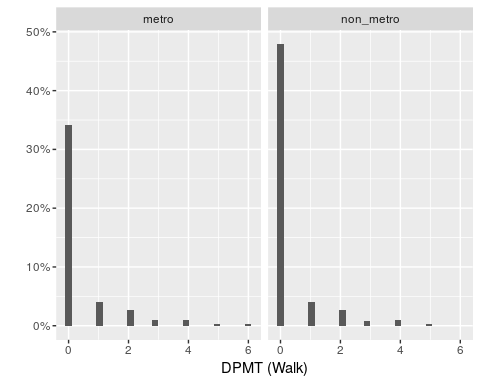


Figure 2.22 Walk miles distribution

1. Estimated Parameters

* counterintuitive coefficients AADVMT

## $metro  
##   
## Call:  
## pscl::hurdle(formula = int\_round(td.miles.Walk) ~ VehPerDriver +   
## HHSIZE + LIF\_CYC + Age0to14 + D3bpo4 + Fwylnmicap + LogIncome |   
## AADVMT + VehPerDriver + HHSIZE + LIF\_CYC + Age0to14 + D1D +   
## D2A\_EPHHM + D3bmm4 + D3bpo4 + ACCESS + WRKCOUNT + Fwylnmicap +   
## Tranmilescap + LogIncome + D3apo + D4c, data = .)  
##   
## Pearson residuals:  
## Min 1Q Median 3Q Max   
## -1.759 -0.463 -0.391 -0.310 10.660   
##   
## Count model coefficients (truncated poisson with log link):  
## Estimate Std. Error z value  
## (Intercept) -0.907874 0.178813 -5.08  
## VehPerDriver -0.049935 0.023479 -2.13  
## HHSIZE 0.060988 0.011543 5.28  
## LIF\_CYCCouple w/o children 0.119167 0.032594 3.66  
## LIF\_CYCEmpty Nester 0.101605 0.033228 3.06  
## LIF\_CYCSingle -0.013599 0.056549 -0.24  
## Age0to14 0.029688 0.016598 1.79  
## D3bpo4 0.000793 0.000324 2.45  
## Fwylnmicap -0.192943 0.060215 -3.20  
## LogIncome 0.120746 0.014843 8.14  
## Pr(>|z|)   
## (Intercept) 0.00000038298006387 \*\*\*  
## VehPerDriver 0.03344 \*   
## HHSIZE 0.00000012661306718 \*\*\*  
## LIF\_CYCCouple w/o children 0.00026 \*\*\*  
## LIF\_CYCEmpty Nester 0.00223 \*\*   
## LIF\_CYCSingle 0.80996   
## Age0to14 0.07368 .   
## D3bpo4 0.01444 \*   
## Fwylnmicap 0.00135 \*\*   
## LogIncome 0.00000000000000041 \*\*\*  
## Zero hurdle model coefficients (binomial with logit link):  
## Estimate Std. Error z value  
## (Intercept) -4.013593 0.238696 -16.81  
## AADVMT -0.001190 0.000339 -3.51  
## VehPerDriver -0.146407 0.032266 -4.54  
## HHSIZE 0.116488 0.017506 6.65  
## LIF\_CYCCouple w/o children -0.143977 0.044730 -3.22  
## LIF\_CYCEmpty Nester -0.137506 0.046731 -2.94  
## LIF\_CYCSingle -0.298566 0.068666 -4.35  
## Age0to14 0.068833 0.024435 2.82  
## D1D 0.001792 0.000910 1.97  
## D2A\_EPHHM 0.120234 0.059099 2.03  
## D3bmm4 0.008061 0.001149 7.02  
## D3bpo4 0.001364 0.000649 2.10  
## ACCESS 0.027498 0.004874 5.64  
## WRKCOUNT 0.114347 0.019439 5.88  
## Fwylnmicap -0.332528 0.083261 -3.99  
## Tranmilescap 0.007673 0.001168 6.57  
## LogIncome 0.205425 0.019768 10.39  
## D3apo 0.011566 0.002910 3.97  
## D4c 0.000489 0.000196 2.49  
## Pr(>|z|)   
## (Intercept) < 0.0000000000000002 \*\*\*  
## AADVMT 0.00045 \*\*\*  
## VehPerDriver 0.0000056918330 \*\*\*  
## HHSIZE 0.0000000000285 \*\*\*  
## LIF\_CYCCouple w/o children 0.00129 \*\*   
## LIF\_CYCEmpty Nester 0.00326 \*\*   
## LIF\_CYCSingle 0.0000137312743 \*\*\*  
## Age0to14 0.00485 \*\*   
## D1D 0.04903 \*   
## D2A\_EPHHM 0.04191 \*   
## D3bmm4 0.0000000000023 \*\*\*  
## D3bpo4 0.03552 \*   
## ACCESS 0.0000000168321 \*\*\*  
## WRKCOUNT 0.0000000040428 \*\*\*  
## Fwylnmicap 0.0000650219698 \*\*\*  
## Tranmilescap 0.0000000000499 \*\*\*  
## LogIncome < 0.0000000000000002 \*\*\*  
## D3apo 0.0000706825964 \*\*\*  
## D4c 0.01275 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1   
##   
## Number of iterations in BFGS optimization: 22   
## Log-likelihood: -3.11e+04 on 29 Df  
##   
## $non\_metro  
##   
## Call:  
## pscl::hurdle(formula = int\_round(td.miles.Walk) ~ HHSIZE + LIF\_CYC +   
## D3apo + LogIncome | AADVMT + VehPerDriver + HHSIZE + LIF\_CYC +   
## Age0to14 + D1D + D2A\_EPHHM + WRKCOUNT + LogIncome + D3apo, data = .)  
##   
## Pearson residuals:  
## Min 1Q Median 3Q Max   
## -0.944 -0.386 -0.338 -0.283 13.422   
##   
## Count model coefficients (truncated poisson with log link):  
## Estimate Std. Error z value  
## (Intercept) -1.35795 0.16263 -8.35  
## HHSIZE 0.04621 0.01102 4.19  
## LIF\_CYCCouple w/o children 0.06395 0.03266 1.96  
## LIF\_CYCEmpty Nester 0.07535 0.03276 2.30  
## LIF\_CYCSingle 0.11151 0.05509 2.02  
## D3apo 0.00407 0.00149 2.73  
## LogIncome 0.14341 0.01398 10.26  
## Pr(>|z|)   
## (Intercept) < 0.0000000000000002 \*\*\*  
## HHSIZE 0.000027 \*\*\*  
## LIF\_CYCCouple w/o children 0.0502 .   
## LIF\_CYCEmpty Nester 0.0215 \*   
## LIF\_CYCSingle 0.0430 \*   
## D3apo 0.0063 \*\*   
## LogIncome < 0.0000000000000002 \*\*\*  
## Zero hurdle model coefficients (binomial with logit link):  
## Estimate Std. Error z value  
## (Intercept) -5.004156 0.189416 -26.42  
## AADVMT -0.000568 0.000254 -2.24  
## VehPerDriver -0.201467 0.023947 -8.41  
## HHSIZE 0.087666 0.016330 5.37  
## LIF\_CYCCouple w/o children -0.087148 0.039682 -2.20  
## LIF\_CYCEmpty Nester -0.062195 0.041998 -1.48  
## LIF\_CYCSingle -0.138422 0.062963 -2.20  
## Age0to14 0.112863 0.021986 5.13  
## D1D 0.020040 0.003220 6.22  
## D2A\_EPHHM 0.089203 0.050131 1.78  
## WRKCOUNT 0.073121 0.016876 4.33  
## LogIncome 0.286021 0.016677 17.15  
## D3apo 0.016426 0.002109 7.79  
## Pr(>|z|)   
## (Intercept) < 0.0000000000000002 \*\*\*  
## AADVMT 0.025 \*   
## VehPerDriver < 0.0000000000000002 \*\*\*  
## HHSIZE 0.0000000793832374 \*\*\*  
## LIF\_CYCCouple w/o children 0.028 \*   
## LIF\_CYCEmpty Nester 0.139   
## LIF\_CYCSingle 0.028 \*   
## Age0to14 0.0000002845411910 \*\*\*  
## D1D 0.0000000004846051 \*\*\*  
## D2A\_EPHHM 0.075 .   
## WRKCOUNT 0.0000147146567698 \*\*\*  
## LogIncome < 0.0000000000000002 \*\*\*  
## D3apo 0.0000000000000068 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1   
##   
## Number of iterations in BFGS optimization: 24   
## Log-likelihood: -4.14e+04 on 20 Df

1. Model goodness-of-fit and Validation

Table 2.21 Walk miles model (hurdle model) performance

|  |  |  |
| --- | --- | --- |
| metro | rmse | pseudo.r2 |
| metro | 1.013 | 0.340 |
| non\_metro | 0.806 | 0.122 |

##### Bike Miles Traveled Model (hurdle model)

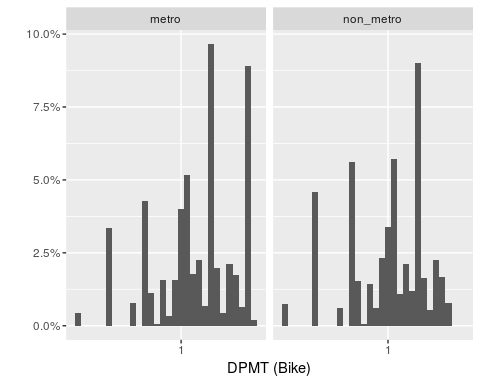


Figure 2.28 Bike Miles distribution

1. Estimated Parameters

* counterintuitive coefficients AADVMT

## $metro  
##   
## Call:  
## pscl::hurdle(formula = int\_round(td.miles.Bike) ~ LIF\_CYC + Age0to14 +   
## D3apo | log1p(VehPerDriver) + HHSIZE + WRKCOUNT + LIF\_CYC +   
## Age0to14 + D3bpo4 + Fwylnmicap + Tranmilescap:D4c + Tranmilescap,   
## data = .)  
##   
## Pearson residuals:  
## Min 1Q Median 3Q Max   
## -0.6224 -0.1198 -0.0901 -0.0763 22.6290   
##   
## Count model coefficients (truncated poisson with log link):  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 0.65068 0.10133 6.42 0.00000000013 \*\*\*  
## LIF\_CYCCouple w/o children 0.29517 0.09756 3.03 0.0025 \*\*   
## LIF\_CYCEmpty Nester 0.14039 0.09390 1.50 0.1349   
## LIF\_CYCSingle 0.14000 0.14464 0.97 0.3331   
## Age0to14 -0.11177 0.04795 -2.33 0.0197 \*   
## D3apo -0.00443 0.00509 -0.87 0.3840   
## Zero hurdle model coefficients (binomial with logit link):  
## Estimate Std. Error z value  
## (Intercept) -2.8757916 0.2657413 -10.82  
## log1p(VehPerDriver) -0.7692363 0.1750131 -4.40  
## HHSIZE 0.0723571 0.0412962 1.75  
## WRKCOUNT 0.1737007 0.0507142 3.43  
## LIF\_CYCCouple w/o children -0.5943595 0.1340230 -4.43  
## LIF\_CYCEmpty Nester -0.6467664 0.1358374 -4.76  
## LIF\_CYCSingle -0.6461547 0.2050500 -3.15  
## Age0to14 0.4389457 0.0533470 8.23  
## D3bpo4 0.0046401 0.0012066 3.85  
## Fwylnmicap -1.1209505 0.2404129 -4.66  
## Tranmilescap -0.0209388 0.0040677 -5.15  
## Tranmilescap:D4c -0.0000300 0.0000142 -2.10  
## Pr(>|z|)   
## (Intercept) < 0.0000000000000002 \*\*\*  
## log1p(VehPerDriver) 0.00001106 \*\*\*  
## HHSIZE 0.07975 .   
## WRKCOUNT 0.00061 \*\*\*  
## LIF\_CYCCouple w/o children 0.00000922 \*\*\*  
## LIF\_CYCEmpty Nester 0.00000192 \*\*\*  
## LIF\_CYCSingle 0.00163 \*\*   
## Age0to14 < 0.0000000000000002 \*\*\*  
## D3bpo4 0.00012 \*\*\*  
## Fwylnmicap 0.00000312 \*\*\*  
## Tranmilescap 0.00000026 \*\*\*  
## Tranmilescap:D4c 0.03535 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1   
##   
## Number of iterations in BFGS optimization: 24   
## Log-likelihood: -4.5e+03 on 18 Df  
##   
## $non\_metro  
##   
## Call:  
## pscl::hurdle(formula = int\_round(td.miles.Bike) ~ Age0to14 + D3bpo4 |   
## AADVMT + log1p(VehPerDriver) + LIF\_CYC + Age0to14 + LogIncome,   
## data = .)  
##   
## Pearson residuals:  
## Min 1Q Median 3Q Max   
## -0.5517 -0.0963 -0.0617 -0.0566 31.4488   
##   
## Count model coefficients (truncated poisson with log link):  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 0.22866 0.06086 3.76 0.00017 \*\*\*  
## Age0to14 -0.10417 0.04887 -2.13 0.03303 \*   
## D3bpo4 0.00261 0.00190 1.38 0.16839   
## Zero hurdle model coefficients (binomial with logit link):  
## Estimate Std. Error z value  
## (Intercept) -5.784563 0.664318 -8.71  
## AADVMT -0.001752 0.000908 -1.93  
## log1p(VehPerDriver) -0.649427 0.229176 -2.83  
## LIF\_CYCCouple w/o children -1.133926 0.142812 -7.94  
## LIF\_CYCEmpty Nester -1.134590 0.125830 -9.02  
## LIF\_CYCSingle -0.934868 0.211834 -4.41  
## Age0to14 0.474818 0.045337 10.47  
## LogIncome 0.194778 0.060384 3.23  
## Pr(>|z|)   
## (Intercept) < 0.0000000000000002 \*\*\*  
## AADVMT 0.0537 .   
## log1p(VehPerDriver) 0.0046 \*\*   
## LIF\_CYCCouple w/o children 0.000000000000002 \*\*\*  
## LIF\_CYCEmpty Nester < 0.0000000000000002 \*\*\*  
## LIF\_CYCSingle 0.000010184556702 \*\*\*  
## Age0to14 < 0.0000000000000002 \*\*\*  
## LogIncome 0.0013 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1   
##   
## Number of iterations in BFGS optimization: 16   
## Log-likelihood: -4.01e+03 on 11 Df

1. Model goodness-of-fit and Validation

Table 2.22 Bike miles model (hurdle model) performance

|  |  |  |
| --- | --- | --- |
| metro | rmse | pseudo.r2 |
| metro | 0.313 | 0.237 |
| non\_metro | 0.179 | 0.161 |

#### Trip Frequency-Length (TFL) Models

An alternative model structure we propose is a combination of household level models of trip frequency and average trip length by mode (Figure 2.33).

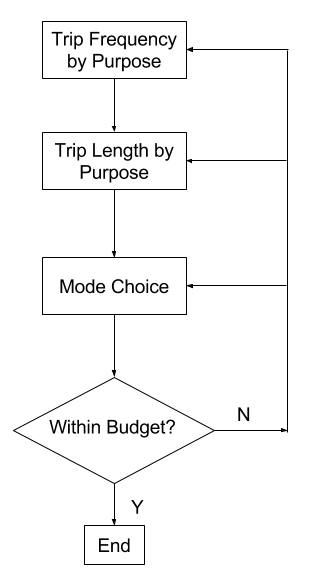


Figure 2.33 Flow Chart of Trip Frequency-Length Model

##### Trip Frequency Models

The trip frequency models of Transit, Bike, and Walk are hurdle models of dependent variable (# Trips): . The hurdle models allows the "inflated" zeros in Transit, Bike, and Walk trip counts to be accounted. It differs from a hurdle model in that a hurdle model allows zeros to arise from both the zero inflation process and the count process, while a hurdle model only allows zeros to arise from the zero hurdle process but not the count process. Like other models, the trip frequency models are segmented by metro and non-metro areas.

###### Transit Trip Frequency Model



Figure 2.34 Transit trip frequency distribution

1. estimated Parameters

## $metro  
##   
## Call:  
## pscl::hurdle(formula = ntrips.Transit ~ AADVMT + HHSIZE + LIF\_CYC +   
## Age0to14 + D1D + Tranmilescap + D4c | AADVMT + VehPerDriver +   
## HHSIZE + WRKCOUNT + LIF\_CYC + Age0to14 + D1D + Fwylnmicap +   
## Tranmilescap:D4c, data = ., na.action = na.omit)  
##   
## Pearson residuals:  
## Min 1Q Median 3Q Max   
## -2.457 -0.254 -0.155 -0.115 32.570   
##   
## Count model coefficients (truncated poisson with log link):  
## Estimate Std. Error z value  
## (Intercept) 0.048617 0.057046 0.85  
## AADVMT -0.001171 0.000338 -3.47  
## HHSIZE 0.112374 0.012259 9.17  
## LIF\_CYCCouple w/o children 0.148899 0.050201 2.97  
## LIF\_CYCEmpty Nester 0.150393 0.055856 2.69  
## LIF\_CYCSingle 0.006653 0.097939 0.07  
## Age0to14 0.167051 0.016449 10.16  
## D1D -0.000692 0.000313 -2.21  
## Tranmilescap 0.005202 0.000920 5.65  
## D4c 0.000694 0.000158 4.39  
## Pr(>|z|)   
## (Intercept) 0.39408   
## AADVMT 0.00053 \*\*\*  
## HHSIZE < 0.0000000000000002 \*\*\*  
## LIF\_CYCCouple w/o children 0.00302 \*\*   
## LIF\_CYCEmpty Nester 0.00709 \*\*   
## LIF\_CYCSingle 0.94584   
## Age0to14 < 0.0000000000000002 \*\*\*  
## D1D 0.02689 \*   
## Tranmilescap 0.000000016 \*\*\*  
## D4c 0.000011527 \*\*\*  
## Zero hurdle model coefficients (binomial with logit link):  
## Estimate Std. Error z value  
## (Intercept) -1.5660306 0.1334279 -11.74  
## AADVMT -0.0039284 0.0005527 -7.11  
## VehPerDriver -0.8753771 0.0668357 -13.10  
## HHSIZE 0.1124917 0.0225501 4.99  
## WRKCOUNT 0.2854531 0.0273473 10.44  
## LIF\_CYCCouple w/o children -0.9103851 0.0678056 -13.43  
## LIF\_CYCEmpty Nester -1.5156797 0.0790749 -19.17  
## LIF\_CYCSingle -1.1199809 0.1148641 -9.75  
## Age0to14 0.4018514 0.0285833 14.06  
## D1D 0.0063118 0.0007727 8.17  
## Fwylnmicap -0.4099763 0.1208247 -3.39  
## Tranmilescap:D4c 0.0001036 0.0000148 6.99  
## Pr(>|z|)   
## (Intercept) < 0.0000000000000002 \*\*\*  
## AADVMT 0.00000000000117882 \*\*\*  
## VehPerDriver < 0.0000000000000002 \*\*\*  
## HHSIZE 0.00000060841501968 \*\*\*  
## WRKCOUNT < 0.0000000000000002 \*\*\*  
## LIF\_CYCCouple w/o children < 0.0000000000000002 \*\*\*  
## LIF\_CYCEmpty Nester < 0.0000000000000002 \*\*\*  
## LIF\_CYCSingle < 0.0000000000000002 \*\*\*  
## Age0to14 < 0.0000000000000002 \*\*\*  
## D1D 0.00000000000000031 \*\*\*  
## Fwylnmicap 0.00069 \*\*\*  
## Tranmilescap:D4c 0.00000000000267718 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1   
##   
## Number of iterations in BFGS optimization: 32   
## Log-likelihood: -1.39e+04 on 22 Df  
##   
## $non\_metro  
##   
## Call:  
## pscl::hurdle(formula = ntrips.Transit ~ log1p(AADVMT) + log1p(VehPerDriver) +   
## HHSIZE + LIF\_CYC + Age0to14 + LogIncome + D1D | log1p(AADVMT) +   
## log1p(VehPerDriver) + WRKCOUNT + LIF\_CYC + Age0to14 + D1B +   
## D3bmm4 + LogIncome, data = ., na.action = na.omit)  
##   
## Pearson residuals:  
## Min 1Q Median 3Q Max   
## -2.0226 -0.1014 -0.0818 -0.0718 56.3148   
##   
## Count model coefficients (truncated poisson with log link):  
## Estimate Std. Error z value  
## (Intercept) 0.4628 0.1781 2.60  
## log1p(AADVMT) -0.0397 0.0188 -2.12  
## log1p(VehPerDriver) 0.1421 0.0636 2.23  
## HHSIZE 0.1421 0.0110 12.87  
## LIF\_CYCCouple w/o children 0.3869 0.0709 5.46  
## LIF\_CYCEmpty Nester 0.4382 0.0663 6.61  
## LIF\_CYCSingle 0.5946 0.1258 4.73  
## Age0to14 0.2310 0.0150 15.43  
## LogIncome -0.0467 0.0165 -2.83  
## D1D -0.0158 0.0047 -3.37  
## Pr(>|z|)   
## (Intercept) 0.00934 \*\*   
## log1p(AADVMT) 0.03442 \*   
## log1p(VehPerDriver) 0.02555 \*   
## HHSIZE < 0.0000000000000002 \*\*\*  
## LIF\_CYCCouple w/o children 0.000000048842 \*\*\*  
## LIF\_CYCEmpty Nester 0.000000000038 \*\*\*  
## LIF\_CYCSingle 0.000002285863 \*\*\*  
## Age0to14 < 0.0000000000000002 \*\*\*  
## LogIncome 0.00468 \*\*   
## D1D 0.00076 \*\*\*  
## Zero hurdle model coefficients (binomial with logit link):  
## Estimate Std. Error z value  
## (Intercept) -0.71019 0.27560 -2.58  
## log1p(AADVMT) -0.08173 0.02803 -2.92  
## log1p(VehPerDriver) -0.23302 0.09966 -2.34  
## WRKCOUNT 0.11309 0.02412 4.69  
## LIF\_CYCCouple w/o children -2.50520 0.08279 -30.26  
## LIF\_CYCEmpty Nester -2.92327 0.08555 -34.17  
## LIF\_CYCSingle -2.75239 0.15258 -18.04  
## Age0to14 0.53435 0.01964 27.21  
## D1B -0.03508 0.00693 -5.07  
## D3bmm4 -0.01023 0.00441 -2.32  
## LogIncome -0.07992 0.02634 -3.03  
## Pr(>|z|)   
## (Intercept) 0.0100 \*\*   
## log1p(AADVMT) 0.0035 \*\*   
## log1p(VehPerDriver) 0.0194 \*   
## WRKCOUNT 0.00000276 \*\*\*  
## LIF\_CYCCouple w/o children < 0.0000000000000002 \*\*\*  
## LIF\_CYCEmpty Nester < 0.0000000000000002 \*\*\*  
## LIF\_CYCSingle < 0.0000000000000002 \*\*\*  
## Age0to14 < 0.0000000000000002 \*\*\*  
## D1B 0.00000041 \*\*\*  
## D3bmm4 0.0204 \*   
## LogIncome 0.0024 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1   
##   
## Number of iterations in BFGS optimization: 23   
## Log-likelihood: -1.69e+04 on 21 Df

1. validation

Table 2.23 Transit trip frequency model performance

|  |  |  |
| --- | --- | --- |
| metro | rmse | pseudo.r2 |
| metro | 0.704 | 0.125 |
| non\_metro | 0.651 | 0.203 |

###### Walking Trip Frequency Model

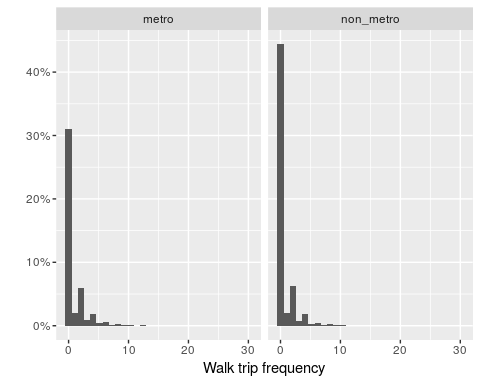


Figure 2.41 Walk trip frequency

1. estimated Parameters

## $metro  
##   
## Call:  
## pscl::hurdle(formula = ntrips.Walk ~ AADVMT + VehPerDriver + HHSIZE +   
## LIF\_CYC + Age0to14 + D1D + D2A\_EPHHM + D3bmm4 + D3bpo4 + ACCESS +   
## Fwylnmicap + Tranmilescap + LogIncome + D3apo + D4c | AADVMT +   
## VehPerDriver + HHSIZE + LIF\_CYC + Age0to14 + D1D + D2A\_EPHHM +   
## D3bmm4 + D3bpo4 + ACCESS + WRKCOUNT + Fwylnmicap + Tranmilescap +   
## LogIncome + D3apo + D4c, data = ., na.action = na.omit)  
##   
## Pearson residuals:  
## Min 1Q Median 3Q Max   
## -2.428 -0.546 -0.450 0.198 11.639   
##   
## Count model coefficients (truncated poisson with log link):  
## Estimate Std. Error z value  
## (Intercept) 0.3396326 0.1108162 3.06  
## AADVMT -0.0006296 0.0001593 -3.95  
## VehPerDriver -0.0627771 0.0159711 -3.93  
## HHSIZE 0.0661640 0.0071273 9.28  
## LIF\_CYCCouple w/o children 0.0818843 0.0204502 4.00  
## LIF\_CYCEmpty Nester 0.0149507 0.0209264 0.71  
## LIF\_CYCSingle -0.1564482 0.0359595 -4.35  
## Age0to14 0.1787000 0.0095290 18.75  
## D1D 0.0002525 0.0001596 1.58  
## D2A\_EPHHM 0.0763013 0.0287273 2.66  
## D3bmm4 0.0021534 0.0004074 5.29  
## D3bpo4 0.0007432 0.0002638 2.82  
## ACCESS 0.0043409 0.0010836 4.01  
## Fwylnmicap -0.2160661 0.0418904 -5.16  
## Tranmilescap 0.0020197 0.0005384 3.75  
## LogIncome 0.0381371 0.0089785 4.25  
## D3apo 0.0032860 0.0013631 2.41  
## D4c 0.0002665 0.0000787 3.39  
## Pr(>|z|)   
## (Intercept) 0.00218 \*\*   
## AADVMT 0.00007773 \*\*\*  
## VehPerDriver 0.00008471 \*\*\*  
## HHSIZE < 0.0000000000000002 \*\*\*  
## LIF\_CYCCouple w/o children 0.00006226 \*\*\*  
## LIF\_CYCEmpty Nester 0.47495   
## LIF\_CYCSingle 0.00001357 \*\*\*  
## Age0to14 < 0.0000000000000002 \*\*\*  
## D1D 0.11373   
## D2A\_EPHHM 0.00791 \*\*   
## D3bmm4 0.00000013 \*\*\*  
## D3bpo4 0.00483 \*\*   
## ACCESS 0.00006176 \*\*\*  
## Fwylnmicap 0.00000025 \*\*\*  
## Tranmilescap 0.00018 \*\*\*  
## LogIncome 0.00002161 \*\*\*  
## D3apo 0.01592 \*   
## D4c 0.00071 \*\*\*  
## Zero hurdle model coefficients (binomial with logit link):  
## Estimate Std. Error z value  
## (Intercept) -3.139966 0.212511 -14.78  
## AADVMT -0.001089 0.000306 -3.55  
## VehPerDriver -0.151875 0.028867 -5.26  
## HHSIZE 0.115032 0.016143 7.13  
## LIF\_CYCCouple w/o children -0.158409 0.040800 -3.88  
## LIF\_CYCEmpty Nester -0.201468 0.042687 -4.72  
## LIF\_CYCSingle -0.273041 0.061517 -4.44  
## Age0to14 0.100385 0.022590 4.44  
## D1D 0.002613 0.000949 2.75  
## D2A\_EPHHM 0.111833 0.053647 2.08  
## D3bmm4 0.007243 0.001089 6.65  
## D3bpo4 0.002233 0.000599 3.73  
## ACCESS 0.028885 0.005245 5.51  
## WRKCOUNT 0.127150 0.017774 7.15  
## Fwylnmicap -0.296850 0.074853 -3.97  
## Tranmilescap 0.008907 0.001074 8.29  
## LogIncome 0.159085 0.017562 9.06  
## D3apo 0.010425 0.002659 3.92  
## D4c 0.000461 0.000188 2.46  
## Pr(>|z|)   
## (Intercept) < 0.0000000000000002 \*\*\*  
## AADVMT 0.00038 \*\*\*  
## VehPerDriver 0.00000014305766 \*\*\*  
## HHSIZE 0.00000000000104 \*\*\*  
## LIF\_CYCCouple w/o children 0.00010 \*\*\*  
## LIF\_CYCEmpty Nester 0.00000236229526 \*\*\*  
## LIF\_CYCSingle 0.00000906148248 \*\*\*  
## Age0to14 0.00000883779981 \*\*\*  
## D1D 0.00589 \*\*   
## D2A\_EPHHM 0.03711 \*   
## D3bmm4 0.00000000002937 \*\*\*  
## D3bpo4 0.00019 \*\*\*  
## ACCESS 0.00000003642928 \*\*\*  
## WRKCOUNT 0.00000000000085 \*\*\*  
## Fwylnmicap 0.00007315368140 \*\*\*  
## Tranmilescap < 0.0000000000000002 \*\*\*  
## LogIncome < 0.0000000000000002 \*\*\*  
## D3apo 0.00008836073023 \*\*\*  
## D4c 0.01399 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1   
##   
## Number of iterations in BFGS optimization: 33   
## Log-likelihood: -4.32e+04 on 37 Df  
##   
## $non\_metro  
##   
## Call:  
## pscl::hurdle(formula = ntrips.Walk ~ AADVMT + VehPerDriver + HHSIZE +   
## LIF\_CYC + Age0to14 + D1D + D2A\_EPHHM + D3bpo4 + ACCESS + WRKCOUNT +   
## LogIncome | AADVMT + VehPerDriver + HHSIZE + LIF\_CYC + Age0to14 +   
## D1D + D2A\_EPHHM + D3bpo4 + WRKCOUNT + LogIncome + D3apo, data = .,   
## na.action = na.omit)  
##   
## Pearson residuals:  
## Min 1Q Median 3Q Max   
## -1.325 -0.464 -0.401 -0.320 20.853   
##   
## Count model coefficients (truncated poisson with log link):  
## Estimate Std. Error z value  
## (Intercept) 0.299015 0.095775 3.12  
## AADVMT -0.000333 0.000131 -2.54  
## VehPerDriver -0.060192 0.012722 -4.73  
## HHSIZE 0.043963 0.007597 5.79  
## LIF\_CYCCouple w/o children 0.049262 0.019986 2.46  
## LIF\_CYCEmpty Nester -0.007278 0.021042 -0.35  
## LIF\_CYCSingle -0.051017 0.034032 -1.50  
## Age0to14 0.132969 0.009983 13.32  
## D1D 0.011622 0.001456 7.98  
## D2A\_EPHHM 0.067366 0.026231 2.57  
## D3bpo4 0.001211 0.000291 4.16  
## ACCESS -0.026325 0.009260 -2.84  
## WRKCOUNT -0.015735 0.008499 -1.85  
## LogIncome 0.045475 0.008372 5.43  
## Pr(>|z|)   
## (Intercept) 0.0018 \*\*   
## AADVMT 0.0111 \*   
## VehPerDriver 0.0000022320118256 \*\*\*  
## HHSIZE 0.0000000071777780 \*\*\*  
## LIF\_CYCCouple w/o children 0.0137 \*   
## LIF\_CYCEmpty Nester 0.7294   
## LIF\_CYCSingle 0.1338   
## Age0to14 < 0.0000000000000002 \*\*\*  
## D1D 0.0000000000000014 \*\*\*  
## D2A\_EPHHM 0.0102 \*   
## D3bpo4 0.0000311868571506 \*\*\*  
## ACCESS 0.0045 \*\*   
## WRKCOUNT 0.0641 .   
## LogIncome 0.0000000557378292 \*\*\*  
## Zero hurdle model coefficients (binomial with logit link):  
## Estimate Std. Error z value  
## (Intercept) -4.035095 0.162199 -24.88  
## AADVMT -0.000651 0.000222 -2.93  
## VehPerDriver -0.193868 0.020660 -9.38  
## HHSIZE 0.092010 0.014375 6.40  
## LIF\_CYCCouple w/o children -0.143740 0.034839 -4.13  
## LIF\_CYCEmpty Nester -0.128328 0.036772 -3.49  
## LIF\_CYCSingle -0.199495 0.054693 -3.65  
## Age0to14 0.104998 0.019551 5.37  
## D1D 0.022151 0.003020 7.34  
## D2A\_EPHHM 0.115400 0.044049 2.62  
## D3bpo4 -0.002354 0.000756 -3.11  
## WRKCOUNT 0.078126 0.014819 5.27  
## LogIncome 0.235719 0.014362 16.41  
## D3apo 0.020123 0.002598 7.75  
## Pr(>|z|)   
## (Intercept) < 0.0000000000000002 \*\*\*  
## AADVMT 0.00340 \*\*   
## VehPerDriver < 0.0000000000000002 \*\*\*  
## HHSIZE 0.0000000001544648 \*\*\*  
## LIF\_CYCCouple w/o children 0.0000369449170576 \*\*\*  
## LIF\_CYCEmpty Nester 0.00048 \*\*\*  
## LIF\_CYCSingle 0.00026 \*\*\*  
## Age0to14 0.0000000785595015 \*\*\*  
## D1D 0.0000000000002207 \*\*\*  
## D2A\_EPHHM 0.00880 \*\*   
## D3bpo4 0.00184 \*\*   
## WRKCOUNT 0.0000001349459596 \*\*\*  
## LogIncome < 0.0000000000000002 \*\*\*  
## D3apo 0.0000000000000095 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1   
##   
## Number of iterations in BFGS optimization: 24   
## Log-likelihood: -6.02e+04 on 28 Df

1. validation

Table 2.24 Walk trip frequency model performance

|  |  |  |
| --- | --- | --- |
| metro | rmse | pseudo.r2 |
| metro | 1.69 | 0.0406 |
| non\_metro | 1.42 | 0.0217 |

###### Biking Trip Frequency Model



Figure 2.47 Bike trip frequency distribution

1. estimated Parameters

## $metro  
##   
## Call:  
## pscl::hurdle(formula = ntrips.Bike ~ AADVMT + Age0to14 + Age65Plus +   
## D1C + D3bpo4 + WRKCOUNT + LogIncome | log1p(AADVMT) + HHSIZE +   
## LIF\_CYC + Age0to14 + Age65Plus + D2A\_EPHHM + D3bpo4 + WRKCOUNT +   
## Fwylnmicap + Tranmilescap + LogIncome, data = ., na.action = na.omit)  
##   
## Pearson residuals:  
## Min 1Q Median 3Q Max   
## -0.7027 -0.1645 -0.1205 -0.0981 39.8895   
##   
## Count model coefficients (truncated poisson with log link):  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -0.119569 0.320612 -0.37 0.7092   
## AADVMT -0.001340 0.000482 -2.78 0.0054 \*\*   
## Age0to14 0.151201 0.019000 7.96 0.0000000000000018 \*\*\*  
## Age65Plus 0.093488 0.035817 2.61 0.0090 \*\*   
## D1C 0.003128 0.001747 1.79 0.0734 .   
## D3bpo4 0.002008 0.000723 2.78 0.0054 \*\*   
## WRKCOUNT 0.046955 0.024350 1.93 0.0538 .   
## LogIncome 0.069766 0.029169 2.39 0.0168 \*   
## Zero hurdle model coefficients (binomial with logit link):  
## Estimate Std. Error z value  
## (Intercept) -4.967972 0.488708 -10.17  
## log1p(AADVMT) -0.074752 0.038759 -1.93  
## HHSIZE 0.085512 0.031548 2.71  
## LIF\_CYCCouple w/o children -0.539160 0.092424 -5.83  
## LIF\_CYCEmpty Nester -0.446249 0.116576 -3.83  
## LIF\_CYCSingle -0.674839 0.154108 -4.38  
## Age0to14 0.389527 0.038843 10.03  
## Age65Plus -0.238551 0.063021 -3.79  
## D2A\_EPHHM 0.312249 0.117192 2.66  
## D3bpo4 0.003756 0.000975 3.85  
## WRKCOUNT 0.183997 0.038413 4.79  
## Fwylnmicap -0.921581 0.128641 -7.16  
## Tranmilescap -0.017346 0.002804 -6.19  
## LogIncome 0.192054 0.043330 4.43  
## Pr(>|z|)   
## (Intercept) < 0.0000000000000002 \*\*\*  
## log1p(AADVMT) 0.05377 .   
## HHSIZE 0.00672 \*\*   
## LIF\_CYCCouple w/o children 0.00000000542706 \*\*\*  
## LIF\_CYCEmpty Nester 0.00013 \*\*\*  
## LIF\_CYCSingle 0.00001192270338 \*\*\*  
## Age0to14 < 0.0000000000000002 \*\*\*  
## Age65Plus 0.00015 \*\*\*  
## D2A\_EPHHM 0.00771 \*\*   
## D3bpo4 0.00012 \*\*\*  
## WRKCOUNT 0.00000166819099 \*\*\*  
## Fwylnmicap 0.00000000000078 \*\*\*  
## Tranmilescap 0.00000000061729 \*\*\*  
## LogIncome 0.00000931996713 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1   
##   
## Number of iterations in BFGS optimization: 30   
## Log-likelihood: -8.65e+03 on 22 Df  
##   
## $non\_metro  
##   
## Call:  
## pscl::hurdle(formula = ntrips.Bike ~ AADVMT + VehPerDriver + HHSIZE +   
## LIF\_CYC + Age0to14 + Age65Plus + D1D + ACCESS + WRKCOUNT + LogIncome +   
## D3apo | AADVMT + VehPerDriver + LIF\_CYC + Age0to14 + Age65Plus +   
## D1A + D2A\_EPHHM + ACCESS + WRKCOUNT + LogIncome + D3apo, data = .,   
## na.action = na.omit)  
##   
## Pearson residuals:  
## Min 1Q Median 3Q Max   
## -0.6605 -0.1437 -0.1056 -0.0873 39.0882   
##   
## Count model coefficients (truncated poisson with log link):  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 1.536486 0.288182 5.33 0.000000097 \*\*\*  
## AADVMT -0.001198 0.000465 -2.57 0.01003 \*   
## VehPerDriver -0.061889 0.041843 -1.48 0.13912   
## HHSIZE 0.048332 0.022306 2.17 0.03025 \*   
## LIF\_CYCCouple w/o children 0.109894 0.066463 1.65 0.09823 .   
## LIF\_CYCEmpty Nester 0.205788 0.078116 2.63 0.00843 \*\*   
## LIF\_CYCSingle -0.141329 0.123715 -1.14 0.25330   
## Age0to14 0.112147 0.029309 3.83 0.00013 \*\*\*  
## Age65Plus -0.145641 0.043436 -3.35 0.00080 \*\*\*  
## D1D 0.002849 0.006946 0.41 0.68167   
## ACCESS 0.084040 0.028946 2.90 0.00369 \*\*   
## WRKCOUNT 0.039836 0.026798 1.49 0.13714   
## LogIncome -0.089071 0.025349 -3.51 0.00044 \*\*\*  
## D3apo 0.006663 0.003702 1.80 0.07189 .   
## Zero hurdle model coefficients (binomial with logit link):  
## Estimate Std. Error z value  
## (Intercept) -6.360728 0.451172 -14.10  
## AADVMT -0.002043 0.000619 -3.30  
## VehPerDriver -0.203577 0.061281 -3.32  
## LIF\_CYCCouple w/o children -0.770327 0.082411 -9.35  
## LIF\_CYCEmpty Nester -0.645316 0.103512 -6.23  
## LIF\_CYCSingle -0.830495 0.137582 -6.04  
## Age0to14 0.400223 0.032650 12.26  
## Age65Plus -0.233184 0.058701 -3.97  
## D1A 0.039241 0.014639 2.68  
## D2A\_EPHHM 0.233317 0.120469 1.94  
## ACCESS 0.030270 0.040782 0.74  
## WRKCOUNT 0.130113 0.037873 3.44  
## LogIncome 0.257638 0.041040 6.28  
## D3apo 0.027031 0.005232 5.17  
## Pr(>|z|)   
## (Intercept) < 0.0000000000000002 \*\*\*  
## AADVMT 0.00096 \*\*\*  
## VehPerDriver 0.00089 \*\*\*  
## LIF\_CYCCouple w/o children < 0.0000000000000002 \*\*\*  
## LIF\_CYCEmpty Nester 0.00000000045 \*\*\*  
## LIF\_CYCSingle 0.00000000158 \*\*\*  
## Age0to14 < 0.0000000000000002 \*\*\*  
## Age65Plus 0.00007114663 \*\*\*  
## D1A 0.00735 \*\*   
## D2A\_EPHHM 0.05278 .   
## ACCESS 0.45795   
## WRKCOUNT 0.00059 \*\*\*  
## LogIncome 0.00000000034 \*\*\*  
## D3apo 0.00000023806 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1   
##   
## Number of iterations in BFGS optimization: 27   
## Log-likelihood: -9.42e+03 on 28 Df

1. validation

Table 2.25 Biking trip frequency model performance

|  |  |  |
| --- | --- | --- |
| metro | rmse | pseudo.r2 |
| metro | 0.512 | 0.0562 |
| non\_metro | 0.456 | 0.0578 |

##### Average Trip Length Models

The average trip length models are linear regression models with dependent variable (TRPMILES) power-transformed: . These models are similar in model structure to the non-zero DVMT model in GreenSTEP, but for average trip length for Transit, Bike and Walk trips.

The TFL model option is simplified from the original Trip Frequency-Length-Mode (TFLM) Model, which models individual trips for each household in the sample. One of reason for this simplification was performance: even though it has advantages in that it allows trip information to be utilized in these models, for example, trip purpose and trip length, which are important factors in mode choice decision. In estimation of TFLM model with NHTS data, it needs to use the trip dataset, which has more than 1 million observations; while in simulation, it requires to create a dataset with one observation for every trip. Even though it can work, the requirement for memory and the penalty of speed are high. We eventually settle with the simplified TFL model that caputres the essential of travel demand for non-driving modes.

###### Transit Trip Length Model

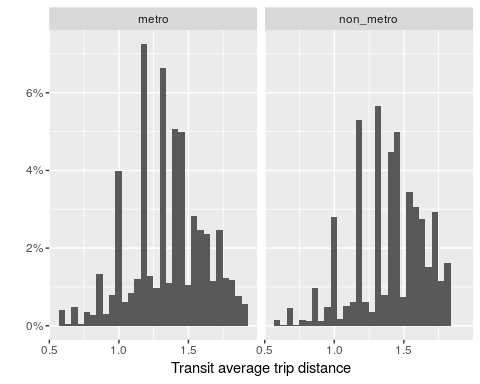


Figure 2.54 Transit average trip distance distribution

1. estimated Parameters

Table 2.26 **Power-transformed Regression Models of Average Transit Trip Length**

|  |  |  |
| --- | --- | --- |
|  | metro | nonmetro |
| log1p(AADVMT) | -0.026\*\*\*  (0.003) | -0.005\*\*  (0.002) |
| log1p(VehPerDriver) | -0.188\*\*\*  (0.014) | -0.013  (0.008) |
| HHSIZE | 0.027\*\*\*  (0.003) | 0.010\*\*\*  (0.003) |
| LIF\_CYCCouple w/o children | -0.076\*\*\*  (0.009) | -0.163\*\*\*  (0.006) |
| LIF\_CYCEmpty Nester | -0.098\*\*\*  (0.009) | -0.158\*\*\*  (0.006) |
| LIF\_CYCSingle | -0.073\*\*\*  (0.013) | -0.152\*\*\*  (0.009) |
| Age0to14 | 0.093\*\*\*  (0.005) | 0.158\*\*\*  (0.004) |
| D1D | 0.001\*\*\*  (0.0002) | 0.001\*  (0.001) |
| D2A\_EPHHM | -0.028\*\*  (0.011) | -0.018\*\*  (0.008) |
| D3bmm4 | 0.0003  (0.0002) |  |
| D3bpo4 | 0.0005\*\*\*  (0.0001) |  |
| ACCESS | 0.002\*\*  (0.001) | -0.0001  (0.004) |
| WRKCOUNT | 0.039\*\*\*  (0.004) | 0.012\*\*\*  (0.003) |
| Fwylnmicap | 0.039\*\*\*  (0.015) |  |
| Tranmilescap | 0.005\*\*\*  (0.0002) |  |
| LogIncome | -0.013\*\*\*  (0.004) | -0.008\*\*\*  (0.002) |
| D3apo | -0.003\*\*\*  (0.001) | -0.003\*\*\*  (0.0003) |
| D4c | -0.0004\*\*\*  (0.0001) |  |
| Tranmilescap:D4c | 0.00002\*\*\*  (0.00000) |  |
| Constant | 0.361\*\*\*  (0.043) | 0.285\*\*\*  (0.026) |
| Observations | 38,676 | 67,755 |
| R2 | 0.101 | 0.132 |
| Adjusted R2 | 0.101 | 0.132 |
| Residual Std. Error | 0.465 (df = 38656) | 0.416 (df = 67741) |
| F Statistic | 229.000\*\*\* (df = 19; 38656) | 792.000\*\*\* (df = 13; 67741) |

Note: p<0.1;***p<0.05;***p<0.01

1. Validation

Table 2.27 Transit length model performance

|  |  |  |
| --- | --- | --- |
| metro | rmse | r2 |
| metro | 3.78 | 0.101 |
| non\_metro | 3.43 | 0.132 |

###### Walking Trip Length Model

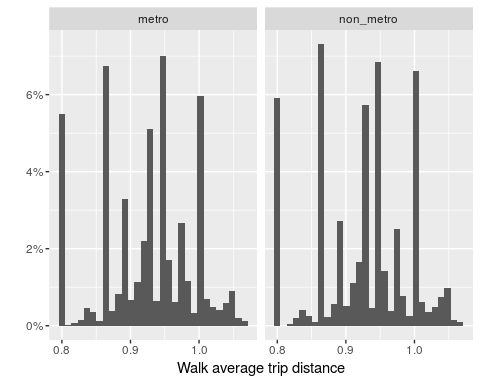


Figure 2.60 Walk average trip distance distribution

1. estimated Parameters

Table 2.28 **Power-transformed Regression Models of Average Walk Trip Length**

|  |  |  |
| --- | --- | --- |
|  | metro | nonmetro |
| log1p(AADVMT) | -0.026\*\*\*  (0.003) | -0.005\*\*  (0.002) |
| log1p(VehPerDriver) | -0.188\*\*\*  (0.014) | -0.013  (0.008) |
| HHSIZE | 0.027\*\*\*  (0.003) | 0.010\*\*\*  (0.003) |
| LIF\_CYCCouple w/o children | -0.076\*\*\*  (0.009) | -0.163\*\*\*  (0.006) |
| LIF\_CYCEmpty Nester | -0.098\*\*\*  (0.009) | -0.158\*\*\*  (0.006) |
| LIF\_CYCSingle | -0.073\*\*\*  (0.013) | -0.152\*\*\*  (0.009) |
| Age0to14 | 0.093\*\*\*  (0.005) | 0.158\*\*\*  (0.004) |
| D1D | 0.001\*\*\*  (0.0002) | 0.001\*  (0.001) |
| D2A\_EPHHM | -0.028\*\*  (0.011) | -0.018\*\*  (0.008) |
| D3bmm4 | 0.0003  (0.0002) |  |
| D3bpo4 | 0.0005\*\*\*  (0.0001) |  |
| ACCESS | 0.002\*\*  (0.001) | -0.0001  (0.004) |
| WRKCOUNT | 0.039\*\*\*  (0.004) | 0.012\*\*\*  (0.003) |
| Fwylnmicap | 0.039\*\*\*  (0.015) |  |
| Tranmilescap | 0.005\*\*\*  (0.0002) |  |
| LogIncome | -0.013\*\*\*  (0.004) | -0.008\*\*\*  (0.002) |
| D3apo | -0.003\*\*\*  (0.001) | -0.003\*\*\*  (0.0003) |
| D4c | -0.0004\*\*\*  (0.0001) |  |
| Tranmilescap:D4c | 0.00002\*\*\*  (0.00000) |  |
| Constant | 0.361\*\*\*  (0.043) | 0.285\*\*\*  (0.026) |
| Observations | 38,676 | 67,755 |
| R2 | 0.101 | 0.132 |
| Adjusted R2 | 0.101 | 0.132 |
| Residual Std. Error | 0.465 (df = 38656) | 0.416 (df = 67741) |
| F Statistic | 229.000\*\*\* (df = 19; 38656) | 792.000\*\*\* (df = 13; 67741) |

Note: p<0.1;***p<0.05;***p<0.01

1. validation

Table 2.29 Walking trip length model performance

|  |  |  |
| --- | --- | --- |
| metro | rmse | r2 |
| metro | 1.047 | 0.0514 |
| non\_metro | 0.434 | 0.0274 |

###### Biking Trip Length Model

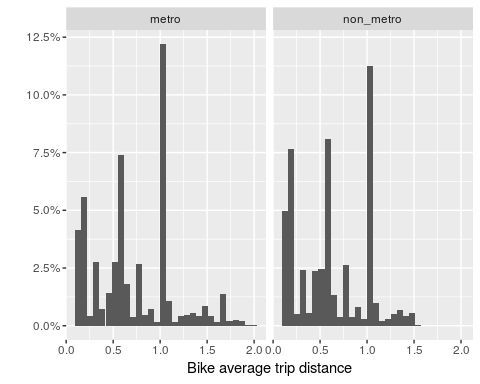


Figure 2.65 Bike average trip distance

1. estimated Parameters

Table 2.30 **Power-transformed Regression Models of Average Bike Trip Length**

|  |  |  |
| --- | --- | --- |
|  | metro | nonmetro |
| log1p(AADVMT) | -0.002  (0.001) |  |
| AADVMT |  | -0.0001\*\*\*  (0.00001) |
| VehPerDriver |  | -0.003\*\*  (0.001) |
| HHSIZE |  | 0.002\*  (0.001) |
| LIF\_CYCCouple w/o children | -0.017\*\*\*  (0.003) | -0.015\*\*\*  (0.002) |
| LIF\_CYCEmpty Nester | -0.014\*\*\*  (0.003) | -0.013\*\*\*  (0.003) |
| LIF\_CYCSingle | -0.018\*\*\*  (0.004) | -0.012\*\*\*  (0.004) |
| Age0to14 | 0.028\*\*\*  (0.002) | 0.020\*\*\*  (0.001) |
| UZAEMPDEN | 0.002\*  (0.001) |  |
| D2A\_EPHHM | 0.011\*\*\*  (0.004) |  |
| D3bpo4 | 0.0001\*\*\*  (0.00003) |  |
| D1D |  | 0.0005\*\*  (0.0002) |
| WRKCOUNT | 0.009\*\*\*  (0.001) | 0.004\*\*\*  (0.001) |
| Fwylnmicap | -0.021\*\*\*  (0.004) |  |
| Tranmilescap | -0.001\*\*\*  (0.0001) |  |
| LogIncome | 0.005\*\*\*  (0.001) | 0.006\*\*\*  (0.001) |
| D3apo |  | 0.001\*\*\*  (0.0001) |
| Constant | -0.010  (0.014) | -0.042\*\*\*  (0.010) |
| Observations | 50,547 | 67,755 |
| R2 | 0.020 | 0.017 |
| Adjusted R2 | 0.020 | 0.016 |
| Residual Std. Error | 0.189 (df = 50534) | 0.166 (df = 67743) |
| F Statistic | 87.000\*\*\* (df = 12; 50534) | 104.000\*\*\* (df = 11; 67743) |

Note: p<0.1;***p<0.05;***p<0.01

1. validation

Table 2.31 Biking trip length model

|  |  |  |
| --- | --- | --- |
| metro | rmse | r2 |
| metro | 0.765 | 0.0203 |
| non\_metro | 0.591 | 0.0166 |

### Considered Model Structures

#### Person Miles Traveled by Mode (PMT) Model

The PMT model is made up two sequential models: a total person miles traveled (PMT) and a mode allocation model (Figure 2.71).

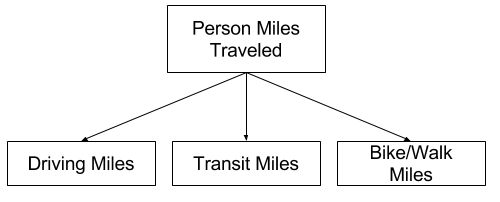


Figure 2.71 Flow Chart of Person Miles Traveled by Mode (PMT) Model

The total person miles traveled is a household level model of total person miles traveled by all household members. It is a linear regression model with pmt (log transformed) as the dependent variable: .

The model can be segmented by life stage of a household (e.g. single, young couple, full nesters, empty nesters) and place types, etc for better model fit and predicting power.

The mode allocation model captures the percentage of PMT by modes for households and allocates total PMT to each mode in prediction. In estimation, we first choose a base mode, compute the ratio of PMT percentage for all other modes relative to that for the base mode, and then use log of the ratio (i.e., log-odds ratio) as the dependent variable of the mode allocation model. We will estimate models if there are modes in total. In prediction, we first predict the log-odds ratios from each of the models, exponentiate the predicted log-odds ratios to get odds ratios, and apply the additional condition that the odds for all modes sum up to 1 to get the predicted PMT percentage for each mode. The model structure is consistent with a multinomial logit model that is commonly used in mode choice modeling.

, and .

The advantage of the PMT model is that the model structure is similar to the existing household travel model in GreenSTEP, and consistent with mode choice models in travel demand modeling. The disadvantages includes:

1. PMT is modeled at an aggregated household level and some of the traveler/trip information that are useful for mode choice modeling is lost. For example, a household will likely have different probability of choosing walking for 2 trips of half mile each than for 1 trip of 1 mile.
2. The NHTS data is dominated by driving when mode shares are measured by distance. The small share of transit and bike/walk mode may bring large variance of the odds ratio variable.
3. Finally, special handling is required when any of the shares are 0 among the modes being modeled (Auto, Transit, Bike/Walk), which is relatively common for daily travel.

After reporting to the TAC in October 2016, we converged to suspend the work on PMT models and focus on TFL models (below).

## Next step

1. ~~Re-retrieve residential location for 2009 NHTS with 2010 census block geography;~~
2. ~~Finalize model specification for each model structure, incorporating SLD variables and place type information~~;
3. ~~Performance testing of model structure/specification~~;
4. ~~Explore potential simplified mode choice model if necessary?;~~ (not using mode choice models)
5. ~~(Task 3) Implement selected model structure/specification as a package for the unified RSPM framework;~~
6. Explore alternatives for reasonable price elasticities.

## Other considerations

Statistical significance, theoretical foundation, and predicting power: because of the large sample size (n>15,000) of 2009 NHTS, it is easy to get a large number of significant coefficients, but they do not necessarily make for a good predictive model. On the other hand, models solely focusing on predictive power (for example, those based on machine learning algorithms) may lack of theoretical basis thus may break down when predict outcomes for conditions far from the base year range. One thing that is particularly hard to do for predictive model is for them to capture behavior that has not been observed in data, for example, potential non-linearity of price elasticities when price rise.

## References

1. Circella, Giovanni, Susan Handy, and Marlon G. Boarnet, 2014. Impacts of Gas Price on Passenger Vehicle Use and Greenhouse Gas Emissions, California Air Resources Board, Technical Background Document.
2. Dong, Jing, Diane Davidson, Frank Southworth and Tim Reuscher, 2012. Analysis of Automobile Travel Demand Elasticities With Respect To Travel Cost, Federal Highway Administration

Graham, D.J., Glaister, S., 2002. The demand for automobile fuel: a survey of elasticities. J. Transp. Econ. Policy JTEP 36, 1–25.

Gregor, B., 2010. Modeling the Effects of Vehicle Travel Costs on Household Vehicle Travel. Oregon DOT.

Ramsey, K., Bell, A., 2014. Smart Location Database Version 2.0 User Guide (Technique Report). U.S. EPA.

USDOT Federal Highway Administration, 2011. 2009 National Household Travel Survey.

sessionInfo()

## R version 3.3.3 (2017-03-06)  
## Platform: x86\_64-pc-linux-gnu (64-bit)  
## Running under: Ubuntu 16.04.2 LTS  
##   
## locale:  
## [1] LC\_CTYPE=en\_US.UTF-8 LC\_NUMERIC=C   
## [3] LC\_TIME=en\_US.UTF-8 LC\_COLLATE=en\_US.UTF-8   
## [5] LC\_MONETARY=en\_US.UTF-8 LC\_MESSAGES=en\_US.UTF-8   
## [7] LC\_PAPER=en\_US.UTF-8 LC\_NAME=C   
## [9] LC\_ADDRESS=C LC\_TELEPHONE=C   
## [11] LC\_MEASUREMENT=en\_US.UTF-8 LC\_IDENTIFICATION=C   
##   
## attached base packages:  
## [1] splines stats graphics grDevices utils datasets methods   
## [8] base   
##   
## other attached packages:  
## [1] stringr\_1.2.0 modelr\_0.1.0 pander\_0.6.0 stargazer\_5.2   
## [5] magrittr\_1.5 scales\_0.4.1 dplyr\_0.5.0 purrr\_0.2.2   
## [9] readr\_1.0.0 tidyr\_0.6.1 tibble\_1.3.0 ggplot2\_2.2.1   
## [13] tidyverse\_1.1.1 pastecs\_1.3-18 boot\_1.3-18 moments\_0.14   
## [17] knitr\_1.15.1 gridExtra\_2.2.1 pacman\_0.4.1   
##   
## loaded via a namespace (and not attached):  
## [1] gtools\_3.5.0 reshape2\_1.4.2 haven\_1.0.0   
## [4] lattice\_0.20-35 colorspace\_1.3-2 htmltools\_0.3.5   
## [7] yaml\_2.1.14 foreign\_0.8-67 DBI\_0.6-1   
## [10] readxl\_0.1.1 plyr\_1.8.4 munsell\_0.4.3   
## [13] gtable\_0.2.0 rvest\_0.3.2 caTools\_1.17.1   
## [16] psych\_1.6.9 evaluate\_0.10 labeling\_0.3   
## [19] forcats\_0.2.0 pscl\_1.4.9 parallel\_3.3.3   
## [22] highr\_0.6 broom\_0.4.1 Rcpp\_0.12.10   
## [25] KernSmooth\_2.23-15 ROCR\_1.0-7 backports\_1.0.5   
## [28] gdata\_2.17.0 jsonlite\_1.4 gplots\_3.0.1   
## [31] mnormt\_1.5-5 hms\_0.2 digest\_0.6.12   
## [34] stringi\_1.1.5 grid\_3.3.3 rprojroot\_1.2   
## [37] bitops\_1.0-6 tools\_3.3.3 lazyeval\_0.2.0   
## [40] MASS\_7.3-45 xml2\_1.1.1 lubridate\_1.6.0   
## [43] assertthat\_0.2.0 rmarkdown\_1.4 httr\_1.2.1   
## [46] R6\_2.2.0 nlme\_3.1-131