Intercity Travel for Metropolitan Access in Northern New England

September 2020

A Research Report from the National Center for Sustainable Transportation

Muriel Adams, University of Vermont

National Center for Sustainable Transportation

THE UNIVERSITY OF VERMONT TRANSPORTATION RESEARCH CENTER

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No. NCST-USC-RR-20-27	2. Government Accession No. N/A	3. Recipient's Catalog No. N/A
4. Title and Subtitle Intercity Travel for Metropolitan Access in Northern New England		**5. Report Date** September 2020
		6. Performing Organization Code N/A
7. Author(s) Muriel Adams		**8. Performing Organization Report No.** N/A
9. Performing Organization Name and Address University of Vermont Transportation Research Center Farrell Hall 210 Colchester Avenue, Burlington, VT 05405		**10. Work Unit No.** N/A
		11. Contract or Grant No. USDOT Grant 69A3551747114
12. Sponsoring Agency Name and Address U.S. Department of Transportation Office of the Assistant Secretary for Research and Technology 1200 New Jersey Avenue, SE, Washington, DC 20590		**13. Type of Report and Period Covered** Research Report (May 2019 – May 2020)
		14. Sponsoring Agency Code USDOT OST-R

15. Supplementary Notes
DOI: https://doi.org/10.7922/G2BR8QG0

16. Abstract
This project builds on the prior work of Dr. Brian H. Y. Lee and Sean Neely, who focused on travel behavior between non-metropolitan areas and large metropolitan centers, because of its impacts on quality of life, multimodal planning, and rural economies. This project studies travel from home locations in northern New England (Vermont, New Hampshire, Maine, and Massachusetts, excluding the Boston-Cambridge-Quincy Metropolitan Statistical Area), going to Boston, New York City, Philadelphia, and Washington, DC. Data were collected in The Intercity Travel, Information, and Technology Survey Questionnaire conducted by Resource Systems Group (RSG Inc.) in May 2014. A total of 2560 valid survey responses were collected using a paid panel purchased by the consultant. Neely previously developed generic mode choice models for these intercity trips but lacked the mode-specific travel times and access measures for more advanced mode choice models. The goal of this current project was to perform more advanced spatial data re-tabulation to generate new mode-specific predictor variables, especially measures of air access using Federal Aviation Administration datasets. The internet access measures were also refined and alternative measures of a zip code location's ruralness were generated. Zip code home location was used for generation of on-road travel times and distances to destination as well as an Amtrak station access measure. In addition to the data development, some specific research questions were pursued with the data: (1) How many trips per year do rural residents take in the Northeast United States to major metropolitan areas? And (2) What socio-economic, location, and accessibility variables are associated with rural trip generation to metropolitan areas?

17. Key Words Access; Air travel; Intercity travel; Metropolitan areas; Mode choice; Rural areas; Socioeconomic factors; Surveys; Travel behavior; Trip generation	**18. Distribution Statement** No restrictions.		
19. Security Classif. (of this report) Unclassified	**20. Security Classif. (of this page)** Unclassified	**21. No. of Pages** 48	**22. Price** N/A

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

NCST

About the National Center for Sustainable Transportation

The National Center for Sustainable Transportation is a consortium of leading universities committed to advancing an environmentally sustainable transportation system through cutting-edge research, direct policy engagement, and education of our future leaders. Consortium members include: University of California, Davis; University of California, Riverside; University of Southern California; California State University, Long Beach; Georgia Institute of Technology; and University of Vermont. More information can be found at: ncst.ucdavis.edu.

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report is funded, partially or entirely, by a grant from the U.S. Department of Transportation's University Transportation Centers Program. However, the U.S. Government assumes no liability for the contents or use thereof.

Acknowledgments

NCST

Intercity Travel for Metropolitan Access in Northern New England

A National Center for Sustainable Transportation Research Report

September 2020

Muriel Adams, Department of Civil and Environmental Engineering, University of Vermont

NCST

TABLE OF CONTENTS

List of Tables

List of Figures

Intercity Travel for Metropolitan Access in Northern New England

Introduction

This project was completed as a National Center for Sustainable Transportation graduate student research project at the University of Vermont. The work builds on the prior work of Dr. Brian H. Y. Lee and the National Center for Sustainable Transportation graduate student Sean Neely who focused on travel behavior between non-metropolitan areas and large metropolitan centers, because of its impacts on quality of life, multimodal planning, and rural economies. This project studies travel <u>from</u> home locations in northern New England (Vermont, New Hampshire, Maine, and Massachusetts, excluding the Boston-Cambridge-Quincy Metropolitan Statistical Area), going <u>to</u> Boston, New York City, Philadelphia, and Washington, DC. Data were collected in T*he Intercity Travel, Information, and Technology Survey Questionnaire* conducted by Resource Systems Group (RSG Inc.) for the University of Vermont's Transportation Research Center (UVM TRC) and the New England Transportation Institute (NETI), with funding from the US Department of Transportation (USDOT) NCST in May 2014. A total of 2560 valid survey responses were collected using a paid panel purchased by the consultant. In his Master's thesis, Neely developed generic mode choice models for these intercity trips but lacked the mode-specific travel times and access measures for more advanced mode choice models. The goal of this current project was to perform more advanced spatial data re-tabulation to generate new mode-specific predictor variables especially measures of air access using Federal Aviation Administration datasets. The Internet access measures were also refined and alternative measures of a zip code location's ruralness were generated. Zip code home location was used for generation of on-road travel times and distances to destination as well as an Amtrak station access measure.

In addition to the data development, some specific research questions were pursued with the data:

1. How many trips per year do rural residents take in the Northeast United States to major metropolitan areas?
2. What socio-economic, location, and accessibility variables are associated with rural trip generation to metropolitan areas?

Data Description

The survey questionnaire for the Intercity Travel, Information, and Technology Survey can be found in Appendix A: Survey Instrument. The study area (Figure 1) comprised part of Massachusetts, New Hampshire, Vermont, and Maine. The Boston Metropolitan Statistical Area, or MSA was excluded as the focus was on rural travel. Participants were recruited to be a part of the paid panel from a private firm by RSG Inc. The three main parts of the survey used in this project were as follows:

- Self-reported number trips taken in the prior one-year to each of Boston, New York City, Philadelphia, and Washington DC;

- Details about the most recent trip taken to any of these four cities; and

- Socioeconomic and household details, especially latitude and longitude, of the household zip code location.
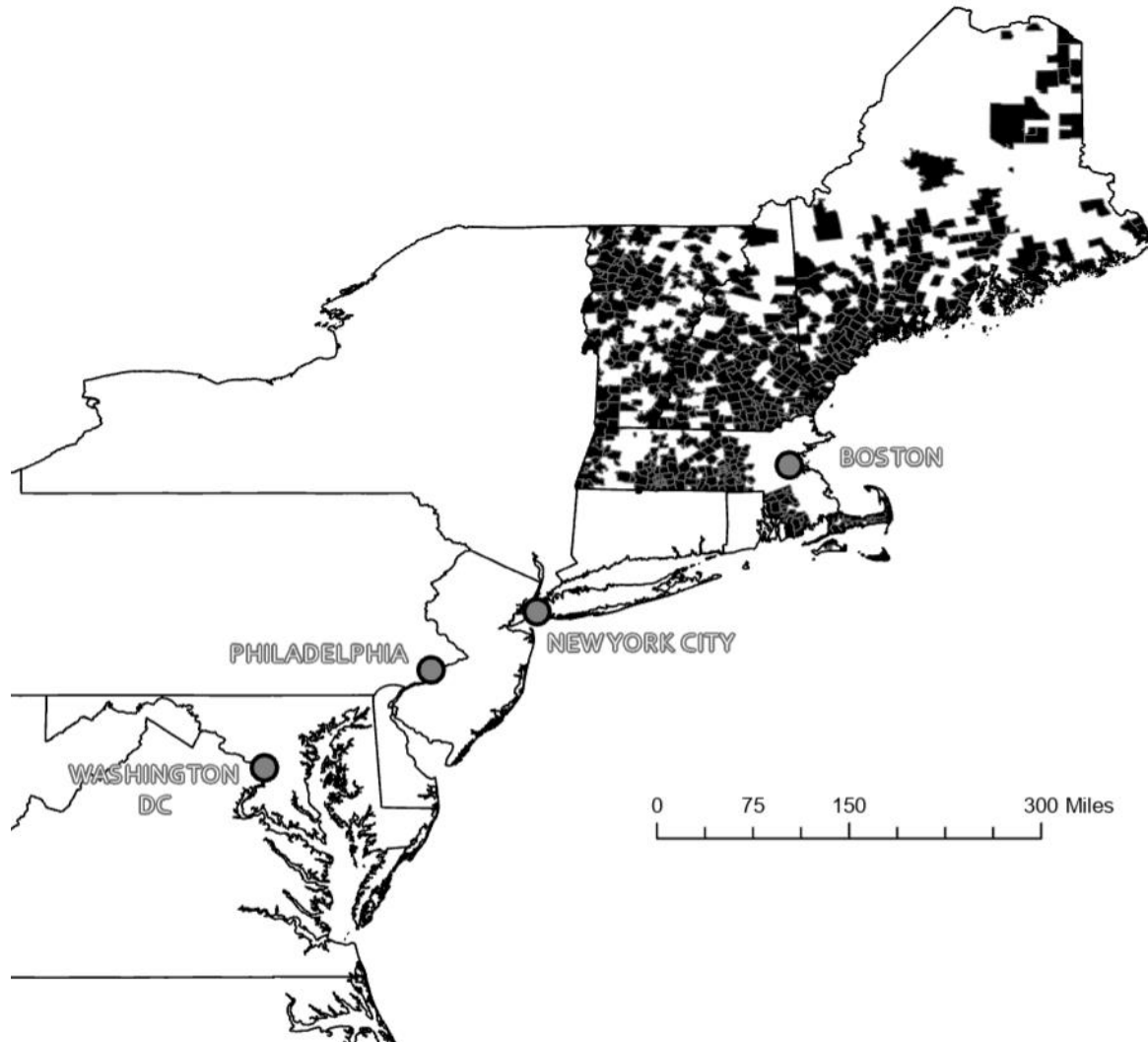
Figure 1. Zip codes Locations of Households in the Study and Destination Cities

As shown in Table 1, very few of the residents surveyed had taken a trip to one of the major metropolitan areas in the past year. As indicated in Table 2, those travelers did use a range of modes suggesting the database does have utility for mode choice study.

Table 1. Percent of Participants Traveling to Destination Cities

	Destinations			
	Boston	**New York City**	**Philadelphia**	**Washington DC**
Percent of Sample Taking Trips to Each Destination	18%	31%	11%	18%

Table 2. Primary Mode Choice of Participants to Destination Cities

Mode	**Boston**	**New York City**	**Philadelphia**	**Washington DC**
Car	78%	47%	49%	37%
Transit	18%	34%	11%	10%
Airplane	3%	16%	38%	52%

New Predictor Variable Generation

The variables in this survey can be broken into categories, based on type as well as how they were calculated. The procedures for calculation of the following new variables, indicated in bold, can be found in Appendix B: Guide to Updated NETI Survey Results Table. The list of variables is as follows:

- Socio-economic Variables
    - **HH Vehicles**
    - **Number of Adults in HH**
    - **Number of Adults in HH with driver's license**
    - **Number of Kids in HH**
    - **Number of Kids in HH with driver's license**
    - **HH Income Levels and Indicators**
    - Gender
    - Age
    - **Education Levels and Indicators**
    - **Internet Access Levels**
    - Smartphone Ownership

- Outcome Variables:
    - Number of trips to each destination city in the past year

- Location Variables: (origin zip code based)
    - **Alternate zip code**
    - **Rural Urban Commuting Code (RUCA) and levels**
    - **Population density number and levels**
    - **Proximity to Boston**
    - **Origin latitude and longitude**

- Accessibility Variables: (origin and destination based)
    - **Greater circle distance for origin to destination**
    - For personal vehicle trips:
        - **On-road distance**
        - **Uncongested on-road time**
    - For air trips:
        - **Travel time (access to airport, wait time, flight time, transfer time, and egress from airport)**
        - **Price of ticket**
        - **Access time to origin airport**
        - **Frequency of trip available on an average day in the year**
        - **Weighted average number of transfers for the trip**
    - **Greater circle distance from origin to closest Amtrak station**
    - **Primary mode choice**
    - **Mode reclassification based on manual interpretation of respondent's other mode entry**

- o Additional Variables: (applies to most recent trip, not included in modelling)
    - o **Purpose reclassification based on manual interpretation of respondent's other purpose entry**
    - o Number of adults and number of children on the trip
    - o **Driving and air trip variables for most recent trip (reference list of personal vehicle and air trip variables above)**

One of the most important first steps in generating spatial variables for this rural focused study was to define the type of rural location a person lives in. This is challenging because it is not simply a measure of the immediate area around one's home but also a function of the area surrounding the home. The size of buffer are to define the surrounding area is subject to debate. Population density is one way to distinguish between urban and rural living conditions (Figure 2). This captures the observed relationship between the proximity of one's neighbors and the number of people within a square mile. This is not an ideal measure of access, as access to some proximate urban areas can offer many non-residential uses, but are often grouped near higher density residential areas. In general, zones that have higher density in our study areas might be considered micropolitan or semi-urban since major metro areas were excluded. The survey includes a large group of low density, or rural, areas, which fits the context of this study (Figure 3).
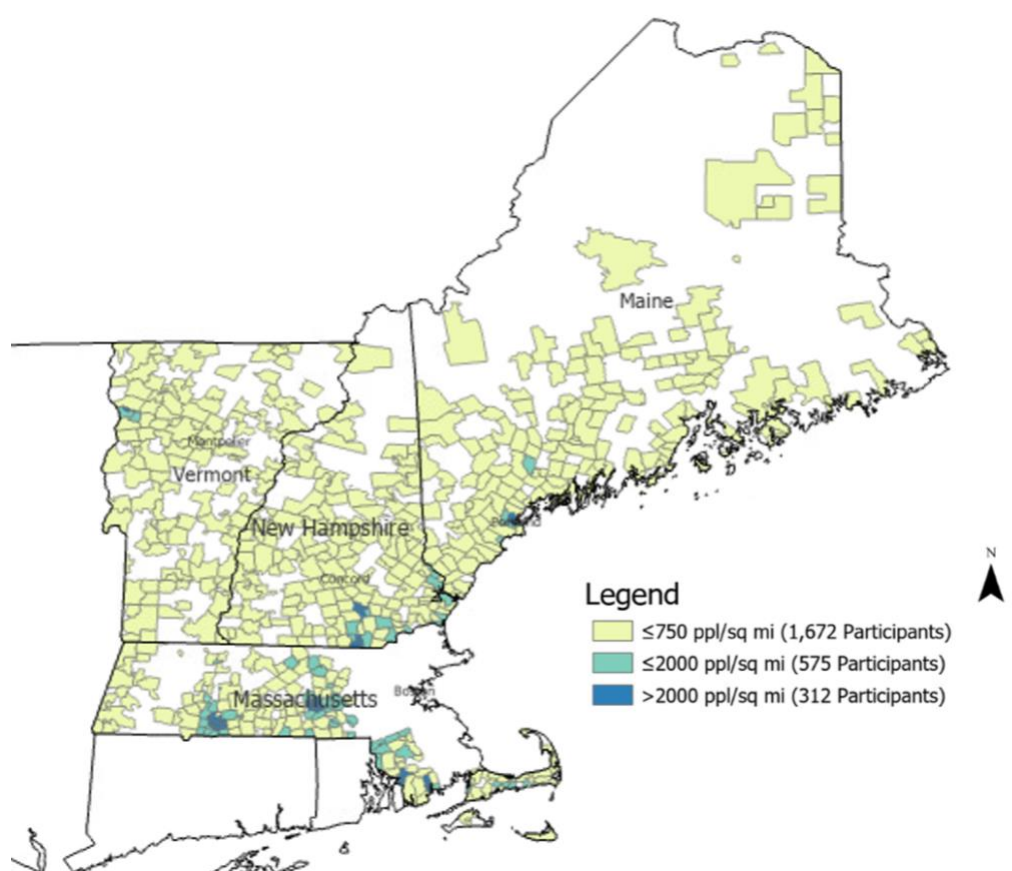
Figure 2. Zip Code Population Density in Study Area
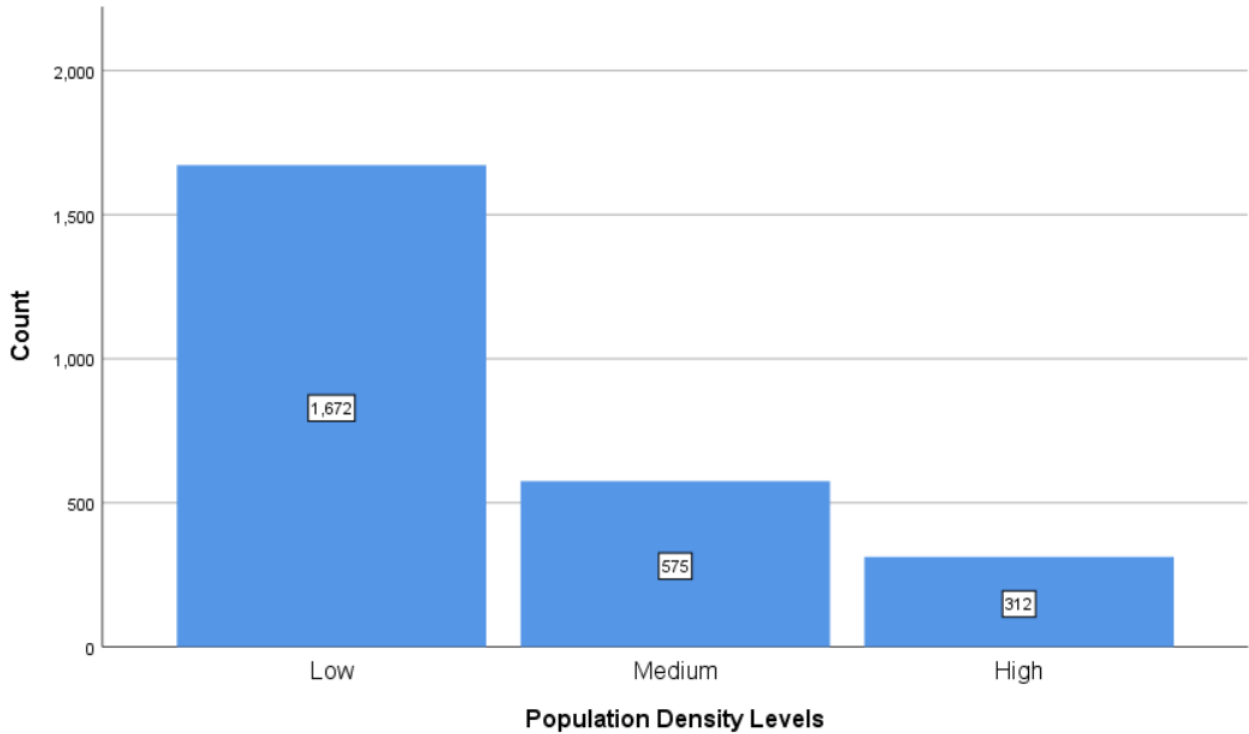
Figure 3. Population Density Distribution of Sample

Another way to measure rural versus urban lifestyle is through spatial commuting patterns. Rural zip code areas with more commuting volume to proximate urban areas suggests access to services in those areas. The volume of commuting by type can be used to define rural or urban. The traffic in each area is described by its Rural Urban Commuting Area code, or RUCA code. (Cromartie, 2019) The RUCA codes for this survey area can be found below in Figure 4. Based on these RUCA codes, the study area has a larger portion of urban or urban-focused areas than the population density suggests (Table 3).
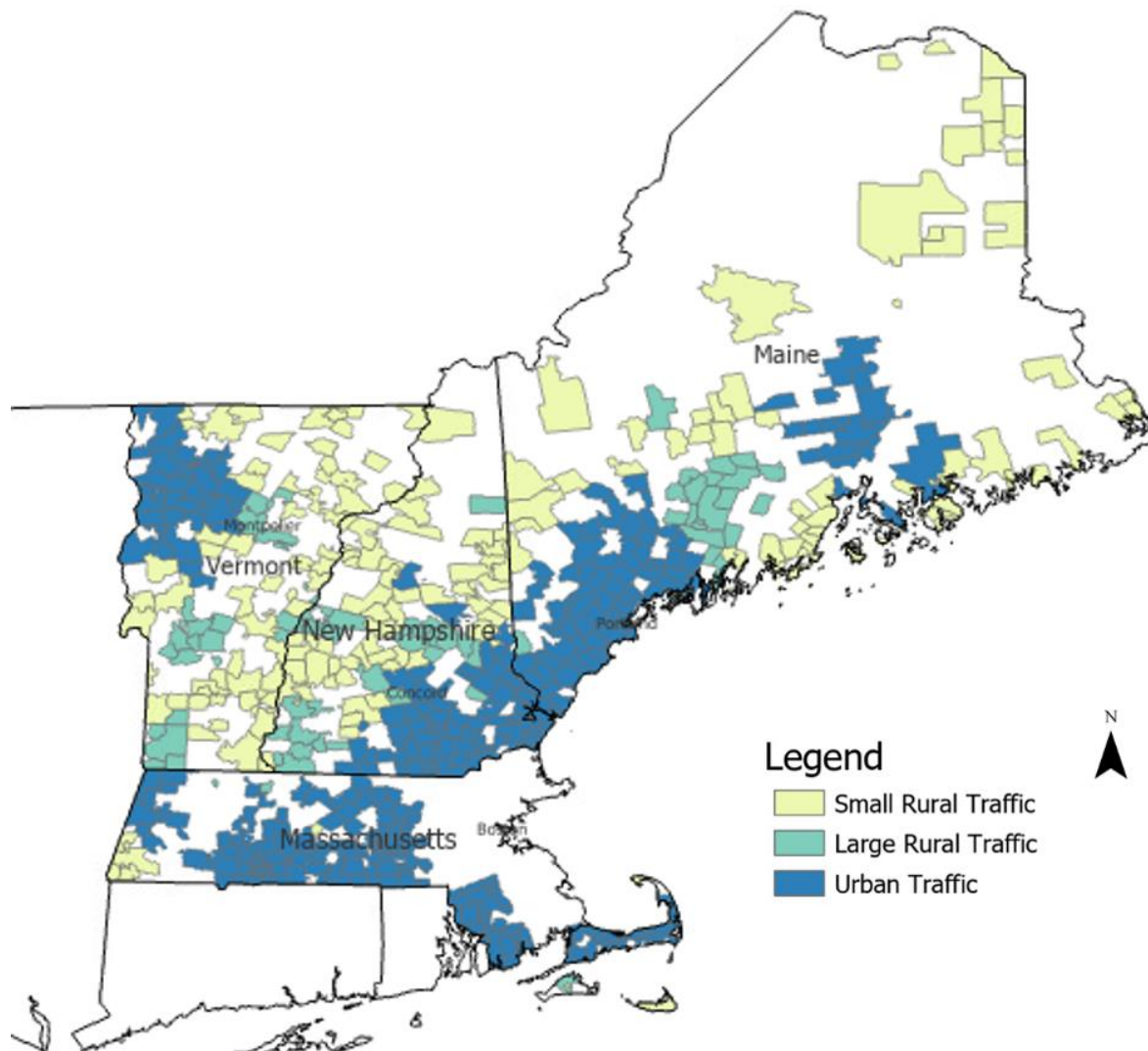
Figure 4. Rural Urban Commuting Area (RUCA) Codes of Study Area

Table 3. Number of Participants Traveling to Destination Cities by RUCA Class

Number of Participants in Category				
RUCA Class	MA	NH	ME	VT
Small Rural	12	91	96	125
Large Rural	9	79	55	71
Urban	916	557	370	179

In rural areas, Internet access can play an important role in access and mobility. Increasingly Internet access is queried in travel surveys. In this case, a number of specific combinations were asked (column 3 of Table 4) and these combinations were aggregated for modeling. This distribution of Internet access suggests most of the sample has at least some access to online information sources. The limited category is small but may be very important in terms of considering mobility choices.

Table 4. Percent of Participants with Levels of Internet Access

Levels	Percent	Combinations	Percent
Limited	1.1%	public	0.2%
		work or school	0.6%
		work and school	0.0%
		public and (work or school)	0.2%
		public, work, and school	0.0%
Sometimes	47.7%	mobile	0.2%
		home	27.7%
		mobile and public	0.0%
		mobile and (Work or School)	0.7%
		home and public	4.1%
		home and (Work or School)	12.2%
		mobile, work, and school	0.0%
		home, work, and school	0.5%
		mobile, public and (work or school)	0.3%
		home, public and (work or school)	1.5%
		mobile, public, work, and school	0.0%
		home, public, work, and school	0.4%
Always	51.3%	home and mobile	16.6%
		home, mobile and (work or school)	26.0%
		home, mobile, public, and (work or school)	6.2%
		home, mobile, work, and school	1.3%
		home, mobile, public, work, and school	1.2%

Drivers and vehicles per household can be very important in rural areas where travel is very auto dependent. Most of the sample had between 1 and 1.5 vehicles per driver, with a small

number of households above 2 (Table 5). When personal vehicles are not available transit can be important. For local urban trips, transit service might consist of a bus or subway but for rural households and their access to metropolitan areas, rural and motor coach is important. While we were unable to calculate access to motor coach service, we were able to use the home zip code location to estimate distance to the nearest AMTRAK station (Table 6). This access varied significantly across the sample. The distribution of Amtrak distance demonstrates the large differences that can exist in ease of access to public transit options, with over 200 miles to get to the closest station for some of the remote locations.

Table 5. Percent of Participants with Vehicles per Drivers in Household

Vehicles per Driver in household	Percent of Sample
< 0.5	2%
< 1	14%
< 1.5	67%
< 2	11%
≥ 2	7%
Missing	1%

Table 6. Description of Vehicles per Driver in Household and Amtrak Distance

	Amtrak Distance (mi)
Mean	21.5
St Dev	28.2
Minimum	0.2
Maximum	242.9

NCST

Travel to Major Metropolitan Areas

Figure 5 and Table 7 illustrate the number of trips taken in the prior year. Although these self-reported numbers are low, there is a significant range with some rural residents traveling for a large number of trips. In Table 8, the bold numbers demonstrate a statistically significant relationship between the RUCA classes and the trips to each destination city. The large number of trips to Boston is attributable to the proximity of some surveyed areas; Boston was closer. Figure 6 shows the average number of trips by zip code to each destination city. The RUCA urban areas are generally circled in red. There is no discernable pattern in number of trips to the destination cities when comparing rural to more urban areas in the study areas.
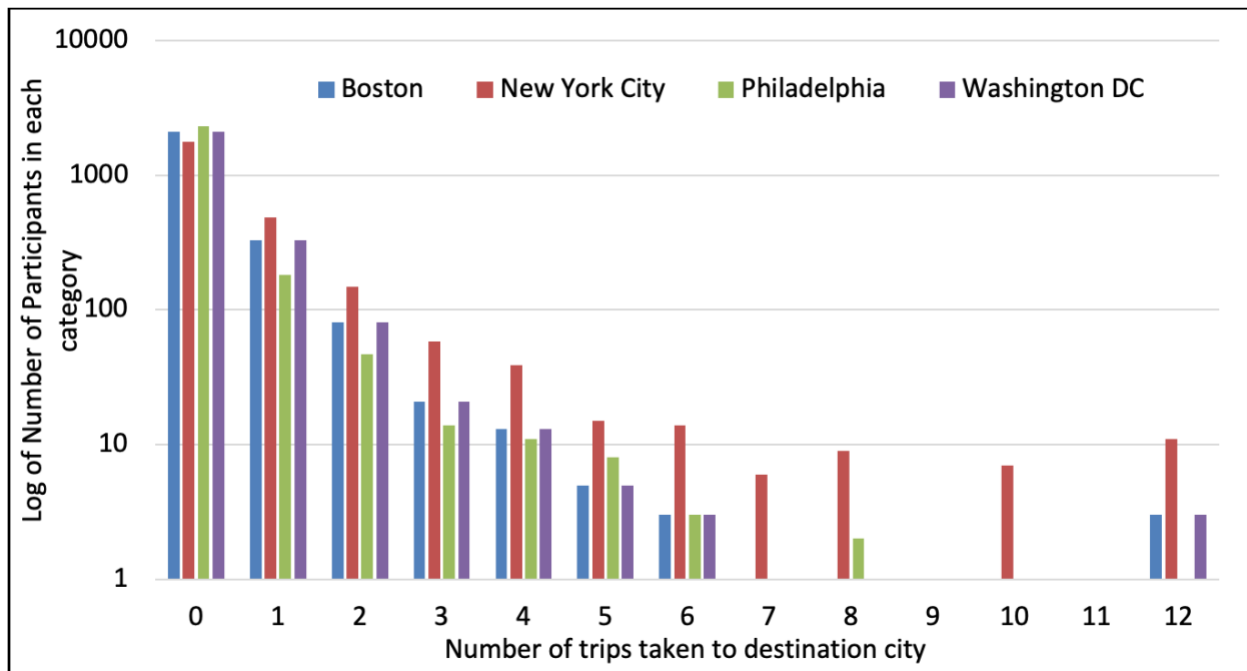
Figure 5. Distribution of Number of Trips to each Destination City

Table 7. Mean Number of Trips to Destination Cities

	Boston	New York City	Philadelphia	Washington DC
Mean	2.9	0.6	0.2	0.3
St Dev	3.5	1.55	0.7	0.8
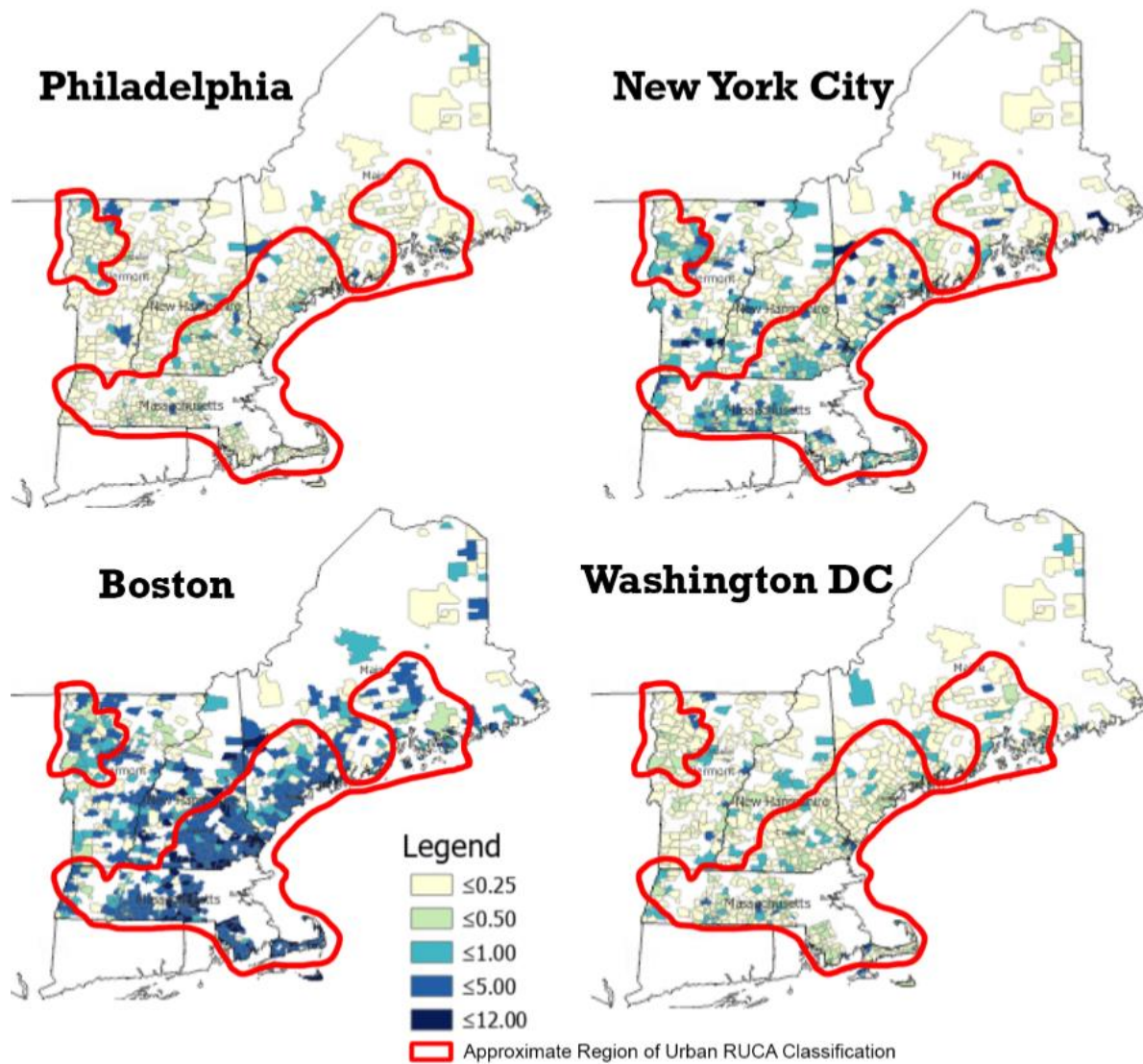
Figure 6. Distribution of Number of Trips by Residents in each Zip Code for each Destination City

Table 8. Mean of Trips to Destination Cities in Sample by RUCA Class

Cities\RUCA Code Class	Small Rural	Large Rural	Urban
Boston	**1.46**	**1.94**	**3.24**
New York City	0.59	0.43	0.65
Philadelphia	0.23	0.11	0.18
Washington DC	0.24	0.28	0.29

Poisson Regression Models of Trip Generation by Destination City

Poisson regression modeling was used due to the non-negative count outcome variable which was skewed towards zero. Dummy variables were built as alternatives to socioeconomic classifications due to sample size. Different combinations of each type of indicator for socioeconomic variables were applied. A base model was selected based on log likelihood and AIC values. The Nagelkerke Quasi-R^2 was calculated for reference for only the final base models presented in Table 9.

Table 9. Parameters and Variables of Base Model

| Variable	Category	Boston		New York City		Philadelphia		Washington DC	
		β	Odds Ratio	β	Odds Ratio	β	Odds Ratio	β	Odds Ratio
Intercept		**1.386**	**3.998**	-0.094	0.91	**-2.029**	**0.132**	**-1.946**	**0.143**
Gender (ref = female)	Male	**0.182**	**1.199**	**0.17**	**1.185**	**0.622**	**1.863**	**0.236**	**1.266**
Age (ref = 18-25 yrs)	85 and up	**-0.812**	**0.444**	-1.904	0.149	**-28.223**	**5.53E-13**	**-28.57**	**3.91E-13**
	75-84	**-0.516**	**0.597**	**-1.006**	**0.366**	-0.329	0.719	**-0.997**	**0.369**
	65-74	**-0.519**	**0.595**	**-0.667**	**0.513**	-0.531	0.588	-0.077	0.926
	55-64	**-0.319**	**0.727**	**-0.455**	**0.634**	-0.088	0.915	0.189	1.209
	45-54	**-0.304**	**0.738**	**-0.474**	**0.622**	-0.099	0.905	-0.048	0.953
	35-44	**-0.392**	**0.676**	**-0.661**	**0.516**	0.162	1.176	-0.239	0.788
	25-34	-0.056	0.946	**-0.337**	**0.714**	-0.028	0.972	-0.003	0.997
Education (ref = high school or less)	≥ some college	**0.286**	**1.332**	**0.369**	**1.447**	-0.348	0.706	**0.558**	**1.748**
	≥ Bachelor's degree	**0.248**	**1.281**	**0.422**	**1.524**	**0.818**	**2.265**	**0.543**	**1.721**
Income (ref = < $50k)	$200k and up	**0.661**	**1.936**	**0.889**	**2.433**	**0.642**	**1.9**	**0.929**	**2.532**
	$100k ≤ x < $200k	**0.364**	**1.44**	**0.287**	**1.332**	0.019	1.02	**0.396**	**1.485**
	$50k ≤ x < $100k	**0.194**	**1.214**	0.004	1.004	-0.038	0.962	0.144	1.154
Driving Distance from Zip Code Coordinates to Destination City		**-0.008**	**0.992**	**-0.003**	**0.997**	0	1	**-0.001**	**0.999**
Nagelkerke R^2		0.232		0.070		0.045		0.074	

Most of the socioeconomic variables that are statistically significant for the trip generation modelling are consistent with prior long-distance trip generation modeling. Age is more influential as the destination city moves farther away but is not statically significant. Men are more likely to have more trips to metro areas than women. As income increases, travel increases. As would be expected, the driving distance has about the same impact regardless of destination city. Increasing distance makes travel less likely.

Research questions for this project were related to rurality, Internet access and air/rail access of the origin home location. In order to select test variables, those test variables that were correlated to variables in the base model were removed. Each test variable was added to base model individually in turn. The log likelihood, AIC, and significance values for each new predictor model were considered to select the best final combination.

The following test variables were considered:

- Own smartphone (Y/N)
- Internet access (3 categories)
- Population density
- Zip code commute pattern (RUCA) (3 categories)
- Air Travel time (airport access, total flying time, number of transfers)
- Weighted average number of transfers for trip
- Straight line distance to nearest Amtrak station
- HH Vehicles per driver
- Children in Household (Y/N)

Based on the results below in Table 10, there are a few patterns found. Out of the variables tested, Internet access and smartphone ownership were the best variables to improve the model fit from the base model. Socioeconomic variables were more significant for the closer cities. Travel-related variables such as time and distance were not significant for this model. Smartphone ownership increases the modeled number of trips taken. Internet access decreases the number of trips taken for Boston and New York, but increases for Philadelphia and Washington DC. Internet was expected to increase the number of trips for all cities, so the decrease for the closer cities is an interesting finding requiring more research.

Based on the Nagelkerke R^2, one can observe that the models are weak, accounting for very little variation in the total amount of travel.

Table 10. Parameters and Variables of Final Models

Variable	Category	Boston		New York City		Philadelphia		Washington DC	
		B	Odds Ratio	β	Odds Ratio	β	Odds Ratio	β	Odds Ratio
Intercept		1.17	3.22	**-0.05**	**0.95**	-2.57	0.08	-2.37	0.09
Gender (ref = female)	Male	**0.61**	**1.84**	**0.18**	**1.20**	**0.61**	**1.85**	**0.22**	**1.25**
Age (ref = 18-25 yrs)	85 and up	-0.34	0.72	-1.39	0.25	-27.80	0.00	-28.30	0.00
	75-84	**-1.63**	**0.20**	**-0.77**	**0.46**	-0.08	0.92	-0.83	0.44
	65-74	**-0.48**	**0.62**	**-0.48**	**0.62**	-0.33	0.72	0.06	1.06
	55-64	**-0.51**	**0.60**	**-0.31**	**0.73**	0.08	1.08	0.25	1.29
	45-54	**-0.43**	**0.65**	**-0.38**	**0.69**	0.02	1.02	0.03	1.03
	35-44	-0.18	0.84	**-0.66**	**0.52**	0.22	1.24	-0.22	0.81
	25-34	-0.17	0.84	**-0.34**	**0.71**	-0.05	0.95	-0.01	1.00
Education (ref = high school or less)	≥ some college	**0.46**	**1.59**	**0.39**	**1.48**	**-0.42**	**0.66**	**0.54**	**1.71**
	≥ Bachelor's degree	**0.42**	**1.52**	**0.38**	**1.46**	**0.80**	**2.21**	**0.50**	**1.66**
Income (ref = < $50k)	$200k and up	**0.80**	**2.22**	**0.75**	**2.12**	**0.45**	**1.57**	**0.82**	**2.27**
	$100k ≤ x < $200k	**0.21**	**1.24**	**0.19**	**1.20**	-0.15	0.86	**0.26**	**1.30**
	$50k ≤ x < $100k	0.11	1.11	-0.04	0.96	-0.11	0.90	0.10	1.11
Owns Smartphone (ref= not)	Yes	0.12	1.13	**0.31**	**1.37**	**0.46**	**1.58**	0.05	1.05
Internet Access Levels (ref= limited)	Always	**-0.73**	**0.48**	-0.21	0.81	0.32	1.37	0.42	1.52
	Sometimes	**-0.94**	**0.39**	-0.44	0.64	0.14	1.15	0.11	1.11
Driving Distance from Zip Code Coordinates to Destination City		**-0.01**	**1.00**	**0.00**	**1.00**	0.00	1.00	0.00	1.00
Nagelkerke R^2		0.45		0.09		0.05		0.093	

Conclusions

Survey results indicates that technology and Internet access appear to be statistically significant in long distance travel decisions from the rural areas in northeastern United States to major metropolitan areas on the eastern seaboard. In addition, the socioeconomic factors that influence travel for urban areas also hold for people from rural areas. Finally, accessibility variables beyond full trip driving distance, including air and rail travel variables, were not shown to be statistically significant factors to trip generation for this survey and may be a function of the skewed study area.

This project's limitation include factors related to survey execution. Questions about the previous full year of travel leave open bias related to recall. The travel variable calculations are limited based on the error related to use of the zip codes for home location and also the use of city center for destination location.

For future research, use of trip purpose and mode parameters is recommended. Another possible expansion is to focus on the destination cities separately, instead of modelling them with the same parameters. Exploration of different types of rural areas is recommended (both density of the origin locations as well as overall distance to metropolitan areas). The distances in this case were not long enough to make the air travel variables useful. A comparison can be explored between regions of the United States, especially with the addition of the rural areas in the south and Midwest regions. This would allow analysis of longer distances between origin and destination cities, potentially changing the influence of air travel variables on the model.

References

Cromartie, J. (2019, July 3). Rural-Urban Commuting Area Codes. From
https://www.ers.usda.gov/data-products/rural-urban-commuting-area-codes.aspx

Neely, S. (2016). Factors Influencing Mode Choice for Intercity Travel from Northern New
England to Major Northeastern Cities. UC Davis: National Center for Sustainable Transportation.
Retrieved from https://escholarship.org/uc/item/6zw5v1jz

Appendix A: Survey Instrument

Intercity Travel, Information, and Technology Survey Questionnaire

Part 1: Recent intercity travel trips and general travel preferences

Section 1-A: The following questions are about your recent trips.

1. How many times have you visited one of the following cities in the past twelve months? (Exclude trips where the city was not the primary destination and you only passed through it on the way to another destination) It may be helpful to refer to your calendar or daybook to recall your trips from the last twelve months
[Column = Frequency (drop-down box from 0 to 11, then 12 or more); Row = City]
   - Boston
   - New York City
   - Philadelphia
   - Washington, DC

[If only one city has frequency > 0], then identify this city as <recent city>, go to 3].

[If more than one city has frequency > 0], then go to 3].

[If 0 cities visited, then skip to Section 1-C]

[if visited at least one city above in past twelve months]

2. What mode(s) of transportation have you used for your trip(s) to each city in the past twelve months? Please select all that apply. [Column = Mode; Row = City]
   - Personal auto/car
   - Rental car (including car share) or a borrowed car
   - Intercity bus (e.g., Greyhound, Peter Pan, Megabus)
   - Intercity rail (e.g., Amtrak)
   - Airplane
   - Other

[Programmer: only show rows for cities that were visited in past 12 months]

[if visited at least one city above in past twelve months]

3. [*If intercity bus or intercity rail selected for ANY city*] How do you usually get information about routes and schedules for bus or rail trips? *Please select all that apply.*
   - Use pamphlets or other printed material
   - Ask a friend or family member
   - Visit the station
   - Call the bus or rail company
   - Search the internet
   - Use smart phone or tablet apps
   - Other, please specify:

[Programmer: randomize order of answer options]

[if # of cities visited > 1]
4. Which city did you visit most recently? [Choices are from those cities visited with frequency > 0]
   Answer = <recent city>

Section 1-B: The following questions are about your MOST RECENT trip to <recent city>.
5. [Skip if frequency to # cities visited = 1]
   What mode(s) of transportation did you use for your MOST RECENT trip to <recent city>?
   Please select all that apply.
   - Personal auto/car
   - Rental car (including car share) or a borrowed car
   - Intercity bus (e.g., Greyhound, Peter Pan, Megabus)
   - Intercity rail (e.g., Amtrak)
   - Airplane
   - Other, please specify:

6. What was the purpose of your most recent trip to <recent city>? *Please select all that apply.*
   - Leisure/vacation
   - Visit friends
   - Business
   - Family event
   - Other, please specify:

7. How many people travelled with you on your most recent trip to <recent city>? (Exclude those who did not make at least part of the journey with you)
   [Drop-Down for each age group = Number (drop-down box from 0 to 10, then 11 or more); Row = Age Group]
   - # Adults (18 and over):
   - # Children (under 18):

[if bus, rail, or plane trip]
8. How did you plan this trip and book your tickets? *Please select all that apply*.
   - Went to the airline, bus, or train website
   - Went to a travel website (e.g., Expedia.com, Kayak.com)
   - Called the airline, bus company, or train line
   - Through a travel agency
   - A friend or family member booked it for me
   - Other, please specify: _____

9. [NIGHTS] How many nights did you stay for your most recent trip to < recent city >?
   [Drop-down box from 0 to 6, then 7 or more]

Section 1-C: The following are general travel and communication questions about you and your household.

10. How many registered vehicles (in working order) are available to your household?
Please include all cars, pickup trucks, minivans, and motorcycles/scooters to which your household has regular access, whether owned, leased, or a company vehicle.
[Drop-down box from 0 to 9, then 10 or more]

11. Do you have a driver's license?
    - Yes
    - No

12. How many people live in your household? How many of you are licensed drivers? [Two columns: People (including yourself), Drivers (including yourself) – (drop-down boxes from 0 to 9, then 10 or more); Row = Age Group]
    - # Adults (18 and over):
    - # Children (under 18):

13. How do you access the internet? *Please select all that apply.*
    - Internet service at home
    - Internet service at school
    - Internet service at work
    - Public internet service (e.g., at the library, community center)
    - Mobile device with a cellular data plan (e.g., smart phone, internet-enabled tablet)
    - Other, please specify:

Part 2: Travel preferences

In this section, consider the following statements and select how much you agree or disagree on a scale from 1 (completely agree) to 7 (completely disagree).
[programmer: scale formatted as shown below – columns evenly spaced (though feel free to make this prettier and use this format throughout the survey for this scale):

Completely Agree			Neutral			Completely Disagree
1	2	3	4	5	6	7

14. I feel I am less dependent on cars than my parents are/were.

15. I need to drive my car to get where I need to go.

16. I love the freedom and independence I get from owning one or more cars.

17. It would be hard for me to reduce my driving mileage.

18. For me to be able to leave the driving to someone else (e.g., a bus driver) would be desirable.

19. It would be desirable for my household to be able to have fewer cars.

20. Being able to freely perform tasks, including using a laptop, tablet, or smartphone is an important reason for me to choose bus or train travel.

21. Having reliable WiFi internet access while I travel on a bus or train is important to me.

22. When taking a bus or train, being able to plan my trip and buy tickets online is important to me.

23. It would be important to me to receive email or text message updates about my bus or train trip.

24. I find tablet or smartphone apps for travel and trip planning to be helpful.

25. When the government tries to improve things, it never works.

26. If everyone works together, we could improve the environment and future for the earth.

27. People like me take the bus or the train.

28. I would be willing to pay more when I travel if it would help the environment.

29. I tend to use the fastest form of transportation, regardless of cost.

30. For me, the whole idea of being on a bus or train with other people I do not know seems uncomfortable.

31. I enjoy being out and about and observing people.

32. I don't mind traveling with people I do not know.

33. Having my privacy is important to me when I travel.

34. When I choose a home, I value having adequate space for parking two or more cars.

35. When I choose a neighborhood to live in, I like to be able to walk to a commercial or village center.

36. Living in a multiple family building (e.g., apartment, condo) wouldn't give me enough privacy.

37. I like living in a neighborhood where there is a lot going on.

38. I am confident that if I want to, I can do things that I have never done before.

39. I worry about crime or other disturbing behavior on buses and trains, or while walking in and around the stops/stations.

40. It is important to me to control the radio and the air conditioning in the car.

41. I feel really stressed when driving for a long time in congestion in and around big cities.

42. I prefer to use the most comfortable transportation mode regardless of cost or time.

43. Having a low-stress trip is more important than reaching my destination quickly.

44. I get very annoyed being stuck behind a slow driver.

45. I am usually in a hurry when I make a trip.

46. With my schedule, minimizing time spent traveling is very important to me.

47. I would use the bus or train more often if it were cheaper to ride.

48. Rather than owning a car, I would prefer to borrow, share, or rent a car just for when I need it.

Part 3: An imaginary situation

Imagine that someone has asked you to travel from your home to Manhattan in New York City (NYC) for an important appointment next month and you have decided to go. You will stay one night at a hotel and travel alone. Your host will pay for your hotel costs but not for getting you there; you would be responsible for all costs of gas, parking, or any fares. Assume that, for one reason or another, you have already decided that you will not take any part of the trip by plane.

You now need to choose between taking the entire trip by car (whether yours or not) and taking at least part of the trip by intercity bus or train.

Please imagine the situation described as you answer the questions in the sections that follow.

49. Knowing what you know right now, what mode(s) of transportation do you think are AVAILABLE to you for this trip to NYC? *Please select all that apply.*
    - Personal auto/car
    - Rental car (including car share) or a borrowed car
    - Intercity bus (e.g., Greyhound, Peter Pan, Megabus)
    - Intercity rail (e.g., Amtrak)
    - Other, please specify:

50. How likely are you to choose to take a bus or train for a trip like this to NYC next month?
    - Definitely
    - Very likely
    - Likely
    - Neutral
    - Unlikely
    - Very unlikely
    - Definitely not

51. If you learned there would be no WiFi, and no electrical outlet on the bus or train for this trip, would that make to you less likely to choose a bus or train for this trip?
    - Much less likely
    - Somewhat less likely
    - No change
    - Not applicable to me

TEST GROUP ONLY:
Now we would like you to review a website related to your imaginary trip to NYC. This website will show you some travel options from your home area to Times Square in NYC by combinations of bus and rail.

When you click on the link below, a second window with this website will open.
You can center the website within the pop-up screen using the up/down arrows on the right. You can ask to see more rail and bus services by using the down arrow in the center of your pop-up window.
When you are done reviewing the website, please close the second window and click "next" to continue.

Please click <here> to review this website.

[Test group only]

Please consider the website you looked at earlier and select the how much you agree or disagree on a scale from 1 (completely agree) to 7 (completely disagree).

52. There are more options than what I expected to travel to NYC by bus and train.

53. After seeing the bus and train options for traveling to NYC, I just don't think there's a good way for me to get there by either bus or train.

[Both groups]

54. Having information like this on my smartphone or computer might make it easier for me to understand the kinds of bus and train services available to me.

55. Having so many potential travel options by bus and train is confusing.

Please continue to imagine the NYC trip situation described, consider the following statements, and select how much you agree or disagree on a scale from 1 (completely agree) to 7 (completely disagree).

56. When I drive long distances (like from my home area to NYC), I can get tired and stressed.

57. I worry about the difficulty in finding a parking space at a reasonable cost when I get to NYC.

58. I am concerned that the schedule of the bus or train only lets me travel a few times per day, and I need to be flexible.

59. I could deal with the limited schedules offered by a bus or train for this trip from my home to NYC.

60. I like the idea that I might see and meet new people on a bus or train to NYC.

61. I don't like the idea of riding with a lot of people that I don't know on a bus or train.

62. If I took a bus or train to NYC, I might have to be with people whose behavior I find unpleasant.

63. I could be with other people who share my values when I take a bus or train on a trip like this.

64. I think that taking a BUS to NYC would take a lot longer than driving.

65. I think that taking a TRAIN to NYC would take a lot longer than driving.

66. Without thinking about it much, I would guess that the cost of taking the trip by BUS would be less than the cost of the car trip (including gas, tolls, and parking).

67. Without thinking about it much, I would guess that the cost of taking the trip by TRAIN would be less than the cost of the car trip (including gas, tolls, and parking.)

68. It would be really important to me to minimize costs when I plan this trip to NYC next month.

69. I really want to minimize the time I spend on the trip to NYC, even if that means more stress or higher costs.

70. Being able to use my laptop, tablet, or smartphone when traveling makes me more interested in taking a bus or train to NYC.

71. I am the kind of person who would take my own car to NYC.

72. Most people whose opinions I value would approve of my taking this trip by bus or train.

73. My family would think that I should take this kind of trip by car or plane.

74. My colleagues would likely think that it is strange not to go by a car or plane to NYC.

75. When my friends go to NYC, they always take a bus or train.

76. When my family members go to NYC, they always take a bus or train.

77. It might be unsafe to make this trip by bus or train.

78. The experience at the NYC bus or train station would be so unpleasant that I would try to avoid it.

79. It would be easy for me to get the schedules for a bus or train between here and NYC, and I would understand them.

80. I like the idea of taking a bus or train instead of driving for this trip to NYC.

81. I think that the most RATIONAL choice would be to take a bus or train instead of a car.

82. I think that the most PLEASURABLE choice would be to take a bus or train instead of a car.

83. I think that the most STRESSFUL choice would be to take a bus or train instead of a car.

84. All other things being equal, if a bus was cheaper, but less reliable than a train, I would choose to take a bus.

85. I am confident that if I wanted to, I could take a bus or train for such a trip to NYC next month.

86. I would make an effort to choose a bus or train for such a trip to NYC next month.

87. For me to take a bus or train for such a trip to NYC the next month would be impossible.

88. In this imaginary situation, I would plan to take a bus or train for this trip to NYC next month.

[Test group only]

89. I would trust the person who invited me to NYC to recommend how I should travel.

90. I don't know all the things I NEED to do to make this trip work by bus or train.

91. Given what you know about bus and train services to NYC, how likely are you to choose a bus or train for a trip to NYC next month (like the one described in the imaginary situation)?
    - Definitely
    - Very likely
    - Likely
    - Neutral
    - Unlikely
    - Very unlikely
    - Definitely not

[Test Group Only – If likelihood to take train/bus changed from Question 52]
92. We noticed that you are now <more/less> likely to take the train or bus to NYC. Please tell us why you have changed your mind.
   - [open-end]

[Both groups]
93. Thank you for sharing your opinions about the imaginary trip to NYC. In your real life, how seriously would you consider taking a bus or train to NYC?
   - Definitely would consider
   - Very likely would consider
   - Likely would consider
   - Neutral
   - Unlikely to consider
   - Very unlikely to consider
   - Definitely not consider

Part 4: Other information about you and your household.

94. Which of the following do you own? *Please select all that apply.*
   - Desktop computer
   - Laptop
   - Smartphone
   - Tablet (e.g., iPad, Windows 8 Tablet)
   - Standalone GPS Navigation Device (e.g., Garmin, TomTom)
   - None of the above

95. What is your age?
   - 18-24
   - 25-34
   - 35-44
   - 45-54
   - 55-64
   - 65-74
   - 75-84
   - 85 or older

96. What is your gender?
   - Female
   - Male

97. What is your highest completed level of education?
   - Less than high school diploma
   - High school diploma or equivalent
   - Some college, no degree
   - Associate degree
   - Bachelor's degree

- Graduate or professional degree

98. What is your annual household income? If you are unsure of the answer, please give your best estimate.
   - Under $25,000
   - $25,000 - $49,999
   - $50,000 - $74,999
   - $75,000 - $$99,999
   - $100,000 - $149,999
   - $150,000 - $199,999
   - $200,000 - $249,999
   - $250,000 or more

Appendix B: Guide to Updated NETI Survey Results Table

Legend:

YT = Year Trips

MRT = Most Recent Trip

MT = Minimum Time (used in reference to Airplane Trip variable calculations)

= index for destination city,

- Many variables are measured separately for each destination city (1 = Boston, 2 = New York, 3 = Philadelphia, 4 = Washington DC), so the variable listed as mode_#_V3 in this document will correspond to variables mode_1_V3, mode_2_V3, mode_3_V3, and mode_4_V3 in the data table.

Spatial Variables

- Variable: homezip_V2
    - Description: Zip Code reassignment (correction) for PO box locations and missing polygons. homezip_1_1 is the original zip code as recorded in the survey
    - Process:
        1. IN ARCGIS, import participant data from Excel, convert to geodatabase
        2. IN ARCGIS, use "summary statistics" of participant data table by zip code to get frequency of responses in each zip code
        3. IN ARCGIS, join summary stats of participant table to zip code shapefile (from popden_1) by zip code/homezip_1 variable
        4. IN ARCGIS, select zip codes that have participants by inverting selection of zip codes where frequency of participants is null, copy to new shapefile
        5. IN ARCGIS, confirm all zip codes are represented by joining zip code file to participant summary stats (reverse order of step 4), select by attribute for any null data in the zip shapefile fields. If there are zip codes missing, copy to new shapefile. If there are no zip codes missing, skip steps 7-9.
        6. IN ARCGIS, confirm the missing zips are not in the full zip code shapefile downloaded
        7. IN ARCGIS, If the missing zips are not in the full zip code file (step 7) then check the location of these missing zips using google maps, build new field and record the zip code in GIS occupying the location of the missing zip on google maps
        8. In SPSS: Build a new field in the original participants table of corrected zip codes, calculate all of the field to be equal to the original zip code, then recalculate the corrected zip of the rows with alternate zip codes to be equal to the new recorded zip code found in step 8.
- Variable: RUCA_2010
    - Description: Area type based on commuting patterns in the zip code
        1. Calculated based on the census tracks information collected in 2010

![NCST logo]

- o Process:
    1. Download RUCA file for zip codes from
       https://agid.acl.gov/Resources/TechDocs/RUCA_3.10_zip_code_file.xlsx
    2. In ARCGIS, load the RUCA file and convert to geodatabase
    3. In ARCGIS, join the RUCA file to the Zip code table based on the homezip_V2 variable
    4. In the event there are multiple entries in the RUCA file for that zip code, set the program to use the minimum value available

- Variable: Popden_1
    - o Description: population density of zip code area in people per square mile
    - o Process:
        1. Download zip code polygon file from
           https://www.arcgis.com/home/item.html?id=8d2012a2016e484dafaac0451f9aea24
        2. In ARCGIS, load the zip code file and convert to geodatabase
        3. In ArcGIS, use these polygons for all other special joins and special calculations in ArcGIS

- Variable: BOS_1
    - o Description: dummy variable for being within 75 miles of the edge of Boston
    - o Process:
        1. In ArcGIS, combine Boston city border with towns considered in Boston MSA Area
        2. Use the Buffer tool in ArcGIS to extend a boundary 75 miles out from the Combined Boston MSA Area
        3. Build a new variable and fill all rows with 0 (indicating farther than 75 mi from Boston)
        4. Select all zip code areas with at least partial area within the 75mile boundary, reset those rows to be 1 for the new variable (indicating within 75 mi from Boston)

- Variable: origin_lat and origin_long
    - o Description: Zip Code center point coordinates for participant's home location
    - o Process:
        1. IN ARCGIS, use "summary statistics" of participant data table by the new zip code in variable homezip_V2 to get frequency of responses in each zip code
        2. In ARCGIS, Join summary stats of participant table to zip code shapefile by zip code
        3. In ARCGIS, select zip codes that have participants by inverting selection of zip codes where frequency of participants is null, copy to new shapefile
        4. In ArcGIS, use "Feature to Point" of resulting zip codes to get centroids
        5. In ArcGIS, use "add xy coordinates" of centroids to put lat & long categories into point table
        6. In ARCGIS, Join the point table to the participant table by zip code

Coding Variables

- Variable: Commute_V1
  - Description: Dummy variable - Far from Urban versus Close to Urban Indicator based on RUCA Codes
  - Process:
    1. Set new variable to equal 0 for all cases (indicating rural)
    2. Select all RUCA codes that indicate an urban zone and set the new variable for those rows to equal 1 (indicating urban)
       1. Classified based on category C at https://depts.washington.edu/uwruca/ruca-uses.php

- Variable: Commute_V2
  - Description: Category for Area Type - Far from Urban, Middle Distance, and Close to Urban Indicator based on RUCA Codes
  - Process:
    1. Set new variable to equal 1 for RUCA Classes for Far from urban areas
    2. Set new variable to equal 2 for RUCA Classes for Middle Distance areas
    3. Set new variable to equal 3 for RUCA Classes for Close to urban areas
    4. Set new variable to equal 0 for missing RUCA Classes
       1. Classified based on category B at https://depts.washington.edu/uwruca/ruca-uses.php

- Variable: popden_1_V1
  - Description: Low, Medium, and High population density Indicator based on popden_1
  - Process:
    1. Set variable to equal 3 for all rows
    2. Set variable to equal 2 for population densities less than 2000 ppl/sq mi
    3. Set variable to equal 1 for population densities less than 750 ppl/sq mi
    4. Set variable to equal 0 for missing population densities

- Variable: cityfreq_2_#
  - Description: Number of trips to city # in the past year, reclassified from cityfreq_1_#
    1. Only for year trip counts table
  - Process:
    1. Set variable equal to cityfreq_1_# minus 1, so that the number displayed is equal to the actual number of trips taken (the highest value of the new variable represents that number of trips <u>or more</u> in the past year)

- Variable: mode_#_V3
  - Description: Transportation Mode: #: Combined version 3
    1. Primary transportation mode to each city for trips in the last year
    2. Only for year trip counts table
  - Process:
    1. Assign indicator values for primary transportation mode for the trip

1. 1 = Personal or Rental Vehicle
2. 2 = Bus or Rail
3. 3 = Airplane
4. 4 = Other

- Variable: modes_6_x_V1
  - Description: Mode(s) used on most recent trip - Other, reclassified
    1. Only for Most Recent Trip table
  - Process:
    1. Manually input values for main categories that should have been selected based on modes_1_6_x
       1. 0 = other
       2. 1 = Personal Car
       3. 2 = Rental Car
       4. 3 = Bus
       5. 4 = Train
       6. 5 = Airplane

- Variable: modes_mrt_1_V3
  - Description: Mode(s) used on most recent trip – Combined, version 3
    1. Primary transportation mode used on the most recent trip
    2. Only for Most Recent Trip table
  - Process:
    1. Assign indicator values for primary transportation mode for the trip
       1. 1 = Personal or Rental Vehicle
       2. 2 = Bus or Rail
       3. 3 = Airplane
       4. 4 = Other

- Variable: purpose_1_5_x_V1
  - Description: Purpose of most recent trip - Other, reclassified
    1. Only for Most Recent Trip table
  - Process:
    1. Manually input values for main categories that should have been selected based on purpose_1_5_x
       1. 1 = Business
       2. 2 = Personal
       3. 3 = Medical
       4. 4 = Other

- Variable: purpose_mrt_1_V1
  - Description: Purpose of most recent trip – Combined, version 1
    1. Only for Most Recent Trip table
  - Process:
    1. Assign indicator values to each purpose class that could be chosen
    2. Include assignments from purpose_1_5_x_V1 above

3. Calculate new variable by purpose class as follows
    1. 1 = Work/Business
    2. 2 = Personal
    3. 3 = Medical
    4. 4 = Other
    5. 5 = Work/Business and Personal
    6. 6 = Work/Business and Medical
    7. 7 = Work/Business and Other
    8. 8 = Personal and Medical
    9. 9 = Personal and Other
    10. 10 = Work/Business and Personal and Medical
    11. 11 = Work/Business and Personal and Other

- Variable: occ_2_1
    - Description: Other adults (18 and over) on most recent trip, reclassified from occ_1_1
        1. Only for Most Recent Trip table
    - Process:
        1. Set variable equal to occ_1_1 minus 1, so that the number displayed is equal to the actual number of adults (the highest value of the new variable represents that number of adults <u>or more</u>)

- Variable: occ_2_2
    - Description: Children (under 18) on most recent trip, reclassified from occ_1_2
        1. Only for Most Recent Trip table
    - Process:
        1. Set variable equal to occ_1_2 minus 1, so that the number displayed is equal to the actual number of Children (the highest value of the new variable represents that number of children <u>or more</u>)

- Variable: vehicles_2_1
    - Description: number of vehicles in household, reclassified from vehicles_1_1
    - Process:
        1. Set variable equal to vehicles_1_1 minus 1, so that the number displayed is equal to the actual number of vehicles (the highest value of the new variable represents that number of vehicles <u>or more</u>)

- Variable: household_1_1_V1
    - Description: number of adults in the household, reclassified from household_1_1
    - Process:
        1. Set variable equal to household_1_1 minus 1, so that the number displayed is equal to the actual number of people (the highest value of the new variable represents that number of people <u>or more</u>)

- Variable: household_1_2_V1
  - Description: number of adults in household with drivers licenses, reclassified from household_1_2
  - Process:
    1. Set variable equal to household_1_2 minus 1, so that the number displayed is equal to the actual number of people (the highest value of the new variable represents that number of people <u>or more</u>)

- Variable: household_2_1_V1
  - Description: number of children under 18 in the household, reclassified from household_2_1
  - Process:
    1. Set variable equal to household_2_1 minus 1, so that the number displayed is equal to the actual number of people (the highest value of the new variable represents that number of people <u>or more</u>)

- Variable: household_2_2_V1
  - Description: number of children under 18 in household with drivers licenses, reclassified from household_2_2
  - Process:
    1. Set variable equal to household_2_2 minus 1, so that the number displayed is equal to the actual number of people (the highest value of the new variable represents that number of people <u>or more</u>)

- Variable: Internet_1
  - Description: Internet Access: Internet Service Levels version 1
  - Process:
    1. Write out all possible combinations of options
    2. Number options from least available to most available, assign these numbers to rows based on responses in internet_1_1 through 6 as follows:
        1. public only
        2. work or school
        3. mobile
        4. home
        5. work and school
        6. public and (work or school)
        7. mobile and public
        8. mobile and (Work or School)
        9. home and mobile
        10. home and public
        11. home and (Work or School)
        12. public, work, and school
        13. mobile, work, and school
        14. home, work, and school
        15. mobile, public and (work or school)

16. home, public and (work or school)
17. home, mobile and (work or school)
18. home, mobile, public, and (work or school)
19. mobile, public, work, and school
20. home, public, work, and school
21. home, mobile, work, and school
22. home, mobile, public, work, and school

- Variable: Internet_2
  - Description: Internet Access: Internet Service Levels version 2
  - Process:
    1. Distribute previous internet combination levels into 3 categories based on availability
    2. Assign numeric indicators to each category
    3. Choose combinations that would be equal in accessibility
    4. Use if statements for internet_1 to set indicators for each row in new variable, as follows:

Internet Levels	Previous Levels (internet_1)	Descriptions
Limited (1)	1, 2, 5, 6, 12	public
		work or school
		work and school
		public and (work or school)
		public, work, and school
Sometimes (2)	3, 4, 7, 8, 10, 11, 13, 14, 15, 16, 19, 20	mobile
		home
		mobile and public
		mobile and (Work or School)
		home and public
		home and (Work or School)
		mobile, work, and school
		home, work, and school
		mobile, public and (work or school)
		home, public and (work or school)
		mobile, public, work, and school
		home, public, work, and school
Always (3)	9, 17, 18, 21, 22	home and mobile
		home, mobile and (work or school)
		home, mobile, public, and (work or school)
		home, mobile, work, and school
		home, mobile, public, work, and school

- Variable: income_2_1
  - Description: Income reclassified from income_1_1
  - Process:
    1. Group income categories of income_1_1 and assign indicator values
       1. Set variable equal to 4 for income_1_1 categories including and above $200,000 in household income
       2. Set variable equal to 3 for income_1_1 categories between $100,000 and $200,000 in household income, including $100,000
       3. Set variable equal to 2 for income_1_1 categories between $50,000 and $100,000 in household income, including $50,000
       4. Set variable equal to 1 for income_1_1 categories below $50,000 in household income
       5. Set variable equal to 0 for missing values in income_1_1

- Variable: income_2_2, income_2_3
  - Description: income indicator for low income (<$50k, income_2_2) and high income (>=$200k, income_2_3)
  - Process:
    1. Set both variables to 0 for all rows
    2. set new variable income_2_2 to 1 if income_2_1 is equal to 1
    3. set new variable income_2_3 to 1 if income_2_1 is equal to 4

- Variable: city_id_#
  - Description: Destination City ID number
    1. Will be identical down the column since it is indicating the destination city for the following 4 columns
    2. Each city gets its own set of columns
  - Process:
    1. Set new variable equal to value in # position of variable name
       1. Set to 1 for city_id_1
       2. Set to 2 for city_id_2
       3. Set to 3 for city_id_3
       4. Set to 4 for city_id_4

- Variable: gcd_#
  - Description: Greater Circle Distance from Zip Code Coordinates to City Center
    1. straight line distance from center point of zip-code area to center point of city boundary
  - Process:
    1. Using variables origin_lat and origin_long from above
    2. In ArcGIS, add destination coordinates to point table used in origin_lat and origin_long calculation above
    3. Transfer full point table to R by using "Table to Excel" command in ArcGIS, then saving as to CSV format, then loading into R.

4. Calculate greater circle distance in R with latitudes and longitudes for both origins and destinations, origins being zip codes and destinations being destination city centers.
    1. Using the radius of the earth, set as 3,959 miles
    2. Convert the latitudes and longitudes to radians
    3. Calculate distance with 2*Radius*asin(sqrt(sin(change in latitude/2)^2 + cos(origin latitude) * cos(destination latitude) * sin(change in longitude/2)^2))

- Variable: drd_#
    - Description: Driving Distance from Zip Code Coordinates to City Center
        1. on road distance from center point of zip-code area to destination city center
    - Process:
        1. In R, use the same point file as was loaded for the gcd_# variable above
        2. Calculate distance in R with latitudes and longitudes for both origins and destinations using google distance API at uncongested time of day (future Sunday morning at 5am).

- Variable: drtime_#
    - Description: Driving Time, uncongested, from Zip Code Coordinates to City Center
        1. on road uncongested travel time from center point of zip-code area to center point of city boundary
    - Process:
        1. Calculate travel time in R with latitudes and longitudes for both origins and destinations using google distance API at uncongested time of day(future Sunday morning at 5am).

- Variable: drstatus_#
    - Description: Driving Calculation Status from Zip Code Coordinates to City Center
        1. Indicates if the drd_# and drtime_# calculation was completed successfully or if there were errors
    - Process:
        1. Automatic output during the google distance API calculation for drd_# and drtime_# above

- Variable: city_id_mrt
    - Description: Most Recent Trip Destination City ID number
        1. Only for Most Recent Trip table
        2. Important for calculating the next 4 variables
    - Process:
        1. Set variable equal to city_1_1

- Variable: gcd_mrt
    - Description: Greater Circle Distance from Zip Code Coordinates to City Center: Most Recent Trip
        1. Only for Most Recent Trip table

- o Process:
    1. Set variable equal to gcd_1 if city_id_mrt is equal to 1
    2. Set variable equal to gcd_2 if city_id_mrt is equal to 2
    3. Set variable equal to gcd_3 if city_id_mrt is equal to 3
    4. Set variable equal to gcd_4 if city_id_mrt is equal to 4

- Variable: drd_mrt
    - o Description: Driving Distance from Zip Code Coordinates to City Center: Most Recent Trip
        1. Only for Most Recent Trip table
    - o Process:
        1. Set variable equal to drd_1 if city_id_mrt is equal to 1
        2. Set variable equal to drd_2 if city_id_mrt is equal to 2
        3. Set variable equal to drd_3 if city_id_mrt is equal to 3
        4. Set variable equal to drd_4 if city_id_mrt is equal to 4

- Variable: drtime_mrt
    - o Description: Driving Time, uncongested, from Zip Code Coordinates to City Center: Most Recent Trip
        1. Only for Most Recent Trip table
    - o Process:
        1. Set variable equal to drtime_1 if city_id_mrt is equal to 1
        2. Set variable equal to drtime_2 if city_id_mrt is equal to 2
        3. Set variable equal to drtime_3 if city_id_mrt is equal to 3
        4. Set variable equal to drtime_4 if city_id_mrt is equal to 4

- Variable: drstatus_mrt
    - o Description: Driving Calculation Status from Zip Code Coordinates to City Center: Most Recent Trip
        1. Only for Most Recent Trip table
    - o Process:
        1. Set variable equal to drstatus_1 if city_id_mrt is equal to 1
        2. Set variable equal to drstatus_2 if city_id_mrt is equal to 2
        3. Set variable equal to drstatus_3 if city_id_mrt is equal to 3
        4. Set variable equal to drstatus_4 if city_id_mrt is equal to 4

- Variable: MTTime_#, MTTicket_#, MTAccess_#, MTFreq_#, MTTransfers_#
    - o Description: outputs of minimum time air trip from Zip Code Coordinates to City Center
        1. MTTime – estimated minimum total travel time (access, flying, waiting, transfers, + egress)
        2. MTTicket – ticket price associated with MT airport pairs
        3. MTAccess – time to drive to origin airport from origin coordinates
        4. MTFreq – average frequency of this MT route available per day
        5. MTTransfers – average number of transfers, weighted by number of passengers, for this MT route

- Process:
  1. Take latitude and longitudes of zip code centerpoints and destination centerpoints from ArcGIS (calculated above) and export to an Excel table
  2. Download inputs tables required according to Anuar Onayev's instructions
     1. Download all 4 quarters of the DB1B input set
  3. In R, combine the DB1B quarters into 1 table and save as a csv
  4. In R, remove rows from DB1B where the "miles flown" category equals 0
  5. In R, read in other input tables, transfer only required columns to new table and save as csv
  6. In R, bring in ArcGIS coordinates from step 1, rearrange so that there is a row for each origin-destination pair (4 rows per zip code, 1 for each destination), save as a csv
  7. Run the air trip calculation according to Anuar Onayev's Air Travel Algorithm instructions
  8. Bring in both the final output table (trip_air_road.csv) and the precursor to that file with more columns (trips_wairs_notformatted.csv), transfer significant variables from each to a new table ("trip_id","alt_id","O_airport","D_airport","num_transfers", "zip_id","dest_id")

- Variable: MTTime_MRT
  - Description: minimum time air trip from Zip Code home origin Coordinates to Destination City Center
    1. Only for Most Recent Trip table
  - Process:
    1. Set variable equal to MTTime _1 if MRT destination city is Boston (city_id_mrt is equal to 1)
    2. Set variable equal to MTTime _2 if MRT destination city is New York City (city_id_mrt is equal to 2)
    3. Set variable equal to MTTime _3 if MRT destination city is Philadelphia (city_id_mrt is equal to 3)
    4. Set variable equal to MTTime _4 if MRT destination city is Washington DC (city_id_mrt is equal to 4)

- Variable: MTTicket_MRT
  - Description: ticket price for minimum time air trip from Zip Code Coordinates to Destination City Center
    1. Only for Most Recent Trip table
  - Process:
    1. Set variable equal to MTTicket _1 if MRT destination city is Boston (city_id_mrt is equal to 1)
    2. Set variable equal to MTTicket _2 if MRT destination city is New York City (city_id_mrt is equal to 2)
    3. Set variable equal to MTTicket _3 if MRT destination city is Philadelphia (city_id_mrt is equal to 3)

4. Set variable equal to MTTicket _4 if MRT destination city is Washington DC (city_id_mrt is equal to 4)

- Variable: MTAccess_MRT
  - Description: time to drive to origin airport from origin coordinates for minimum time air trip from Zip Code Coordinates to Destination City Center
    1. Only for Most Recent Trip table
  - Process:
    1. Set variable equal to MTAccess _1 if MRT destination city is Boston (city_id_mrt is equal to 1)
    2. Set variable equal to MTAccess _2 if MRT destination city is New York City (city_id_mrt is equal to 2)
    3. Set variable equal to MTAccess _3 if MRT destination city is Philadelphia (city_id_mrt is equal to 3)
    4. Set variable equal to MTAccess _4 if MRT destination city is Washington DC (city_id_mrt is equal to 4)

- Variable: MTFreq_MRT
  - Description: average frequency of this route available per day for minimum time air trip from Zip Code Coordinates to Destination City Center
    1. Only for Most Recent Trip table
  - Process:
    1. Set variable equal to MTFreq _1 if MRT destination city is Boston (city_id_mrt is equal to 1)
    2. Set variable equal to MTFreq _2 if MRT destination city is New York City (city_id_mrt is equal to 2)
    3. Set variable equal to MTFreq _3 if MRT destination city is Philadelphia (city_id_mrt is equal to 3)
    4. Set variable equal to MTFreq _4 if MRT destination city is Washington DC (city_id_mrt is equal to 4)

- Variable: MTTransfers_MRT
  - Description: average number of transfers, weighted by number of passengers, for this route for minimum time air trip from Zip Code Coordinates to Destination City Center
    1. Only for Most Recent Trip table
  - Process:
    1. Set variable equal to MTTransfers _1 if MRT destination city is Boston (city_id_mrt is equal to 1)
    2. Set variable equal to MTTransfers _2 if MRT destination city is New York City (city_id_mrt is equal to 2)
    3. Set variable equal to MTTransfers _3 if MRT destination city is Philadelphia (city_id_mrt is equal to 3)
    4. Set variable equal to MTTransfers _4 if MRT destination city is Washington DC (city_id_mrt is equal to 4)

- Variable: Amtrak_dist
  - Description: Greater Circle Distance from home Zip Code Coordinates to Closest Amtrak station
  - Process:
    1. In ArcGIS, load point file of Amtrak station locations from the Federal Railroad Administration database and convert to geodatabase
    2. In ArcGIS, run the near function to find the nearest Amtrak station to the centerpoints of the zip codes, include the coordinates of the Amtrak station in the information added to the zip code point table.
    3. Transfer full point table to R by using "Table to Excel" command in ArcGIS, then saving as to CSV format, then loading into R.
    4. Calculate greater circle distance in R with latitudes and longitudes for both origins and destinations, origins being zip codes and destinations being Amtrak stations.
       1. Using the radius of the earth, set as 3,959 miles
       2. Convert the latitudes and longitudes to radians
       3. Calculate distance with 2*Radius*asin(sqrt(sin(change in latitude/2)^2 + cos(origin latitude) * cos(destination latitude) * sin(change in longitude/2)^2))

Appendix C: Socioeconomic Variable Data Description

Table C-1. Description of Children in Household

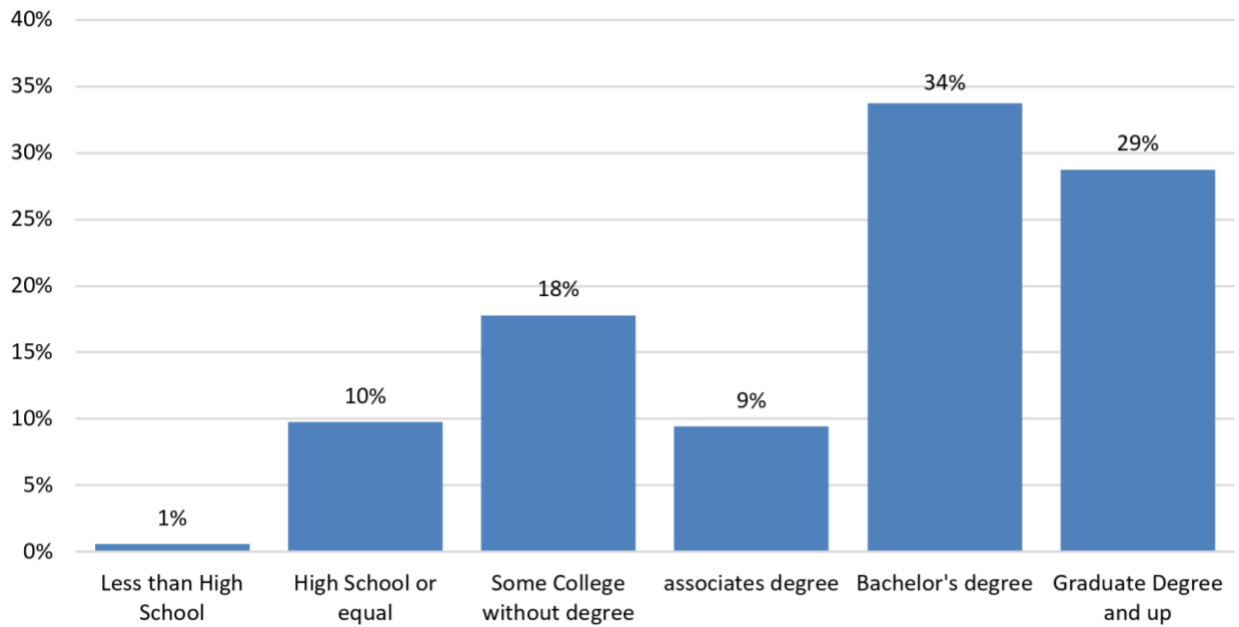|	Number of Children in HH
Mean	0.4
St Dev	0.82
Minimum	0
Maximum	8

Figure C-1. Distribution of Education in Sample

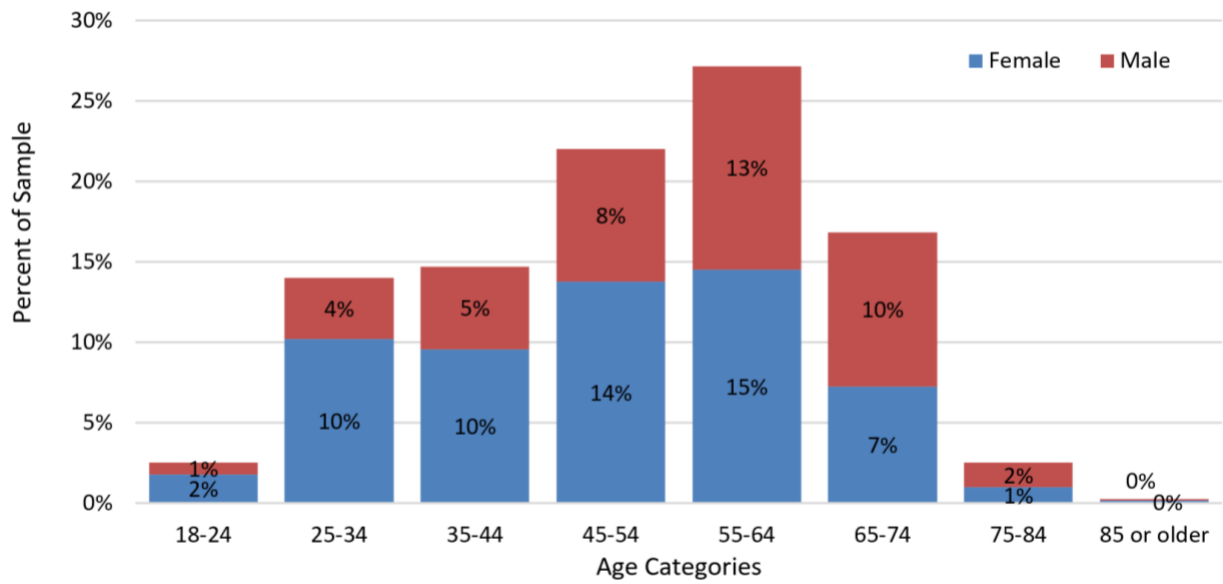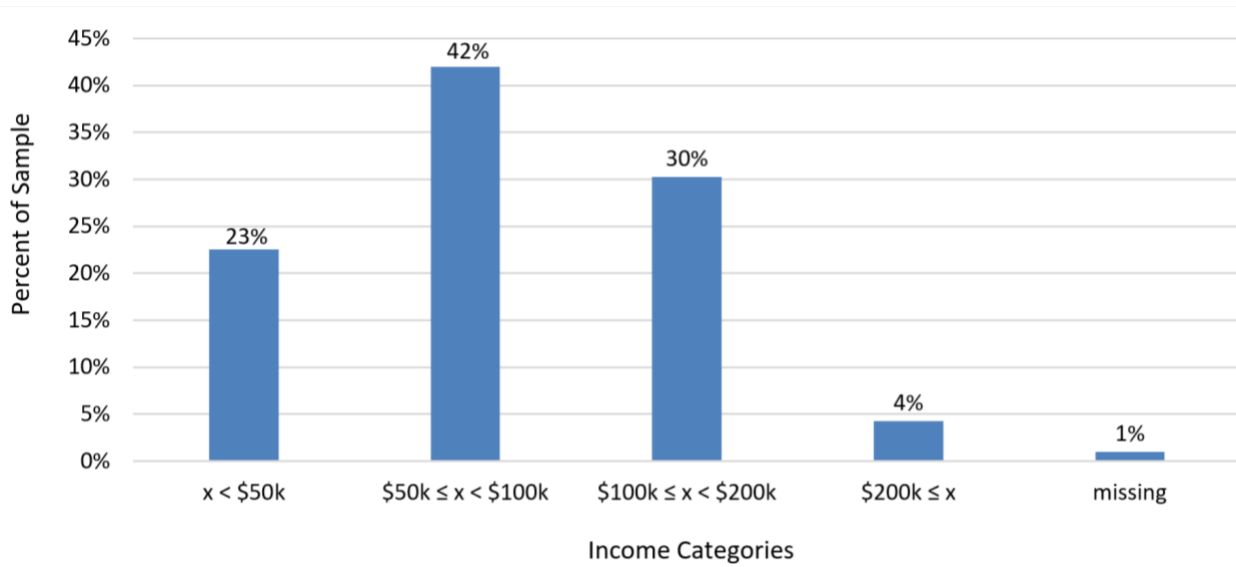99% of sample have a driver's license.

Figure C-2. Distribution of Age by Gender in Sample

Figure C-3. Distribution of Income in Sample