Advancing Understanding of Long-Distance and Intercity Travel with Diverse Data Sources

July 2018

A Research Report from the National Center for Sustainable Transportation

Jonathan Dowds, University of Vermont
Chester Harvey, University of California, Berkeley
Jeff LaMondia, Auburn University
Sarah Howerter, University of Vermont
Hannah Ullman, University of Vermont
Lisa Aultman-Hall, University of Vermont

National Center for Sustainable Transportation

The UNIVERSITY of VERMONT
TRANSPORTATION RESEARCH CENTER

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.

U.S. Department of Transportation (USDOT) Disclaimer

Acknowledgments

This study was funded by a grant from the National Center for Sustainable Transportation (NCST), supported by USDOT through the University Transportation Centers program. The authors would like to thank the NCST and USDOT for their support of university-based research in transportation, and especially for the funding provided in support of this project. The technical support and input from James Sullivan and Karen Sentoff is gratefully acknowledged. We appreciate the service and input of Drs. Pablo Bose and Jane Kolodinsky on graduate committees. We appreciate the opportunity to work with data collected by RSG Inc., AirSage, and the University of Vermont Center for Rural Studies. Data sharing by the Vermont Agency of Transportation and the California Department of Transportation is gratefully acknowledged.

NCST

Advancing Understanding of Long-Distance and Intercity Travel With Diverse Data Sources

A National Center for Sustainable Transportation Research Report

July 2018

Jonathan Dowds, Transportation Research Center, University of Vermont

Chester Harvey, College of Environmental Design, University of California Berkeley

Jeff LaMondia, Department of Civil Engineering, Auburn University

Sarah Howerter, Transportation Research Center, University of Vermont

Hannah Ullman, Transportation Research Center, University of Vermont

Lisa Aultman-Hall, School of Engineering, University of Vermont

[page left intentionally blank]

TABLE OF CONTENTS

List of Tables

List of Figures

Advancing Understanding of Long-Distance and Intercity Travel with Diverse Data Sources

EXECUTIVE SUMMARY

Long-distance travel is an ambiguous designation that is used to refer to an extremely diverse set of trips, differing from one another in mode, distance, and purpose. Long-distance travel encompasses everything from "short" long-distance surface trips between adjacent metropolitan areas through intercontinental air trips spanning thousands of miles. These trips serve a wide range of purposes including business travel, leisure travel, and travel to access essential services such as medical care. As such, long-distance travel is increasingly important for sustainable transportation planning both due to the environmental externalities associated with these trips and also because the benefits of access to long-distance travel are inequitably distributed throughout the population. This project drew on five survey datasets, a mobile-device based dataset from AirSage Inc., and semi-structured interviews to address research questions related to how best to measure long-distance travel, how long-distance travel influences well-being, and how access to long-distance travel varies among socio-demographic groups.

Adequately defining long-distance travel for data collection and modeling remains a challenge and long-distance travel behavior researchers are continuing to experiment with innovative data collection methods for measuring these trips. It is clear that the most common current long-distance data collection methods are suboptimal. Distance thresholds, such as trips over 50 or 75 miles, are a poor method for defining long-distance travel on a national or global scale and there is no consensus on the appropriate recall period for retrospective long-distance travel surveys. Several promising avenues for future research and data collection were identified through this project. Focusing on more memorable long-distance travel indicators within surveys, such as overnight trips, trips including air travel, or an individual's maximum distance from home, may reduce recall error and provide useful outcome measures for assessing individuals' overall long-distance travel tendencies. Project results indicated that expanded use of convenience samples may provide more cost-effective opportunities to measure long-distance trip length and destination distributions. Social network characteristics may be predictive of certain types of long-distance travel and additional methodological improvements are needed to understand how to more effectively collect social network information including network geography. Analysis of survey data in this project suggested that one-time, self-assessed travel frequency estimates (a common existing measure of long-distance travel) can provide only a crude approximation of the levels of long-distance travel and that self-assessment is most effective for identifying non-travelers and very infrequent travelers.

Historically, transportation equity research has focused on access to local goods and services but access to long-distance travel and to more distant destinations is increasingly important for

maintaining social networks and accessing economic opportunities and specialized services. Across multiple datasets in this project, there is ample evidence that lower-income individuals engage in less long-distance travel and have more unmet long-distance travel needs than their higher-income counterparts. Given both the theoretical and empirical evidence that long-distance and intercity travel is correlated with an individuals' own sense of well-being, especially for leisure or personal purposes, inequitable access to long-distance travel cannot be ignored. This finding suggests generally that lack of equity in long-distance access has been masked by lack of data and is a policy concern that must be considered in sustainable transportation planning moving forward.

Introduction

The travel activities that transportation planners have come to refer to as long-distance travel encompass a highly diverse set of trips ranging from surface travel between adjacent metropolitan areas to interregional rail travel and intercontinental air travel. Long-distance trips serve a wide range of purposes including business travel, travel to access essential services such as medical care, and leisure travel for social, recreational, or experiential ends. Access to long-distance travel can positively impact the economic opportunities that are available to an individual and improve well-being by facilitating strong connections with friends and family. In contrast to the very real and important benefits provided by long-distance travel, long-distance travel is also responsible for significant externalities, and increased congestion, energy consumption, and emissions. Long-distance travel also requires a significant time commitment and impacts individuals' time budgets. The role that time constraints play in long-distance travel may be changing as information and communication technologies increasingly enable travelers to engage in alternative activities while traveling. The move toward more automated travel and an improved ability to perform other activities while traveling may have profound implications on the frequency of long-distance travel, and therefore on system impedance between origins and destinations. Creating sensible sustainable transportation policy that maximizes the benefits and minimizes the harms associated with long-distance travel requires a detailed understanding of the motivators, purposes, and modes of long-distance travel.

Unfortunately, within both the academic and practitioner communities, data collection and analysis related to long-distance travel have been limited in comparison to analysis of routine, daily, home-based travel. The lack of focus on long-distance travel reflects a number of factors, including the historical dominance of daily travel in terms of number of trips and trip-miles, the scale disparities between transportation agency jurisdictions and long-distance travel, and the inherent difficulty of long-distance data collection. Data collection is challenging for several reasons. First, there is no widely agreed upon definition of long-distance travel in the transportation community. It has been defined based on a variety of arbitrary one-way trip distance thresholds (ranging from 50 to 250 miles), based on travel modes (specifically air travel), and by trip duration (e.g., an overnight trip). Second, because of the comparatively low frequency of long-distance trips, capturing a representative sample of long-distance travel using survey data collection requires either long recall periods, which can drastically reduce data quality, or very large sample sizes which significantly increase costs. In spite of these significant challenges, the growing prevalence of long-distance travel and its social and environmental impacts necessitates increased data collection and analysis to support intermodal planning, forecasting, and modeling efforts.

According to the US Travel Association, U.S. residents completed close to 2.2 billion trips in excess of 50 miles from home in 2016. Current long-distance travel has been estimated to account for in excess of 30% of total person-miles traveled in the U.S. (*1*). Long-distance travel miles are projected to increase significantly over the next century with air-miles increasing more rapidly than other types of long-distance travel (*2*). The rise in long-distance travel has prompted calls for better data collection on a national and international scale (*3*–*7*).

One response to the long-distance travel data gap has been the development of techniques for extracting travel patterns from mobile device location data (*8*–*12*). These passive approaches can capture large samples at much lower cost than traditional surveys and avoid the limitations associated with recall bias. While passive data collection methods are very promising for applications such as trip rate and O/D generation, the limited or non-existent information about trip makers, travel parties and trip purpose, reduces its utility for forecasting as well as many other research applications. Passive data offers very limited ability to understand who is traveling long-distances and why, thus making it difficult to assess equity concerns or to craft policies to control the environmental impacts of long-distance travel. For these types of research questions, other innovative data collection methods are required that facilitate capturing valid long-distance data at lower cost than a traditional, very large, burdensome household travel survey.

Latent demand for long-distance travel represents an additional gap in our understanding of long-distance travel behavior. Passive data collection methods are inherently limited to collecting data on realized travel and most travel surveys also neglect latent or unserved demand. In the long-distance and intercity realm, unmet needs to access economic opportunities, services or family networks is rarely measured. This creates a blind spot in our ability to understand travel equity and its relationship to quality of life, ultimately limiting our ability to plan for truly sustainable transportation systems.

This project aimed to evaluate different data collection methods and analysis types to begin to fill knowledge gaps in the area of long-distance travel. This report summarizes numerous papers and theses undertaken during this NCST project that document exploratory efforts for gathering long-distance data using: a) unique survey formats that deliver data at a lower cost than would be required for a full–year, randomized, household travel survey; b) face-to-face interviews; and c) passive mobile devices. The methods considered here include alternative classifications of long-distance travel, utilizing self-assessment of travel frequency, non-random sampling strategies, and measuring the geographic extent of individuals' social networks. The data for these analyses were collected with two original surveys, the Longitudinal Survey of Overnight Travel (LSOT) and the People in Your Life (PiYL) pilot survey, and three existing surveys, the California Household Travel Survey (CHTS), a Long Range Transportation Planning Survey (LRTPS) conducted by the Vermont Agency of Transportation, and the Vermonter Poll telephone survey collected at UVM. The motivation of these analyses was to improve data methods, evaluate evidence of inequitable long-distance travel access, and explore mode choice-related decision making to address the environmental impacts associated with the longest distance trips.

Report Organization

The remaining sections in this report are organized as follows:

Section 2. Background

Section 3. Survey Datasets

Section 4. Alternative Methods for Long-distance Data Collection
Section 5. Equity in Long-Distance Travel
Section 6. Conclusions

Section 2 provides a brief overview of the history of long-distance surveys in the United States, the challenges associated with long-distance data collection, the role of passive data collection research, and the need for innovative long-distance data collection methods. Section 3, *Survey Datasets,* describes two original survey efforts, the LSOT and the PiYL pilot, as well as two DOT surveys, the CHTS and the Vermont LRTPS, that form the basis of much of the analysis conducted for the project.

Four journal articles related to the challenges of collecting long-distance data are summarized in Section 4. These methodological articles explore differing approaches to lower-cost long-distance data that is sufficiently comprehensive to support effective long-distance travel planning and research. The first of these, Aultman-Hall et al. (*13*), tests a range of different distance-threshold definitions for long-distance travel, presents several tour generation regression models, and examines the spatial complexity of long-distance tours. The second paper, Dowds et al. (*1*), compares individuals' one-time estimates of their typical overnight travel frequency with the number of trips they reported in monthly surveys administered over the subsequent 12 months. The third paper, Harvey and Aultman-Hall (*14*), assessed the feasibility of using convenience sampling for long-distance data collection. These first three papers all draw on the LSOT data. The fourth paper, Aultman Hall et al. (*15*), explores the association between social network extent and long-distance travel using PiYL data. These four summaries are followed by a more extensive description of how to characterize social network geography that has not yet been published elsewhere.

Section 5 explores equity concerns related to access to long-distance travel. It includes a summary of the role of long-distance travel on well-being based on interviews and survey data and looks at the unmet travel needs documented in the Vermont LRTPs (*16*), work that was at the core of a Master's thesis funded from the project. In addition, this section documents socio-economic differences in long-distance travel from the CHTS and the Vermont LRTPS that are not yet published elsewhere.

The report ends with overarching conclusions about long-distance data collection and equity in access to long-distance travel in Section 6.

Background

The 1995 American Travel Survey (ATS) stands as the most recent, full year, long-distance data collection effort in the United States. This survey defined long-distance travel as any trip with a one-way distance of at least 100 miles and collected long-distance trip information in three month increments for a full year (*17*). The 1995 ATS had a sample size of 70,000 households and, in spite of the fact that the data are more than two decades old, continues to be used in long-distance travel research (*18*, *19*) because it is yet to be supplanted by a comparable, national data collection effort. The 2001 National Household Travel Survey (NHTS) also collected long-distance travel data but did so using a 50-mile one-way distance threshold and a limited, four-week data collection period. Because of its smaller sample size, only 26,000 households, and shorter data collection period, the 2001 NHTS had more limited utility and was not suitable for producing long-distance trip rates for smaller subnational regions (*20*).

There are a number of factors that have inhibited long-distance data collection. At the most basic level, there is no agreed upon definition of what constitutes a long distance trip. Long-distance travel has been defined in terms of distance thresholds (ranging 50 miles to several hundred miles), modes used, and overnight stays in non-home locations. Regardless of the definition, many long-distance trips cross regional and even national jurisdictional boundaries, reducing the incentive for agencies to collect data about these trips (*4*). When agencies are interested in collecting long-distance data, and have identified a long-distance definition that is meaningful in their jurisdiction, a number of methodological challenges remain. First, declining participation rates in surveys of all types has triggered concerns about the representativeness of household travel surveys in general (*21*). Second, individuals trip-length estimation skills are limited and variable (*22*), introducing errors when using a distance threshold definition of long-distance travel. Finally, because long-distance trips occur at lower frequencies than other trips, survey sample sizes must be larger than for daily travel or respondents must report long-distance trips for a longer recall period. This first approach drives up survey costs while the second approach increases the risk of significant underreporting due to recall errors.

Because of these challenges there has been a growing emphasis on passive data collection, especially using mobile device-based location data. Mobile device-based location data, either calculated from cell-tower triangulation or recorded directly from a device's GPS offer numerous advantages for collecting long-distance travel behavior. These data can be collected over long periods of time, which is necessary to capture lower frequency long-distance travel (*23*, *24*). Recent research has demonstrated the viability of detecting origins and destinations with these data (*25*–*27*) and to track intercity travel (*28*). Travel studies using passively collected data have provided evidence of significant underreporting in traditional travel surveys (*12*). Unfortunately, passive data collection methods are limited in terms of the information that they can capture with regard to trip purposes, travel modes, and travel party composition. These data sources all present new challenges in terms of representativeness as they do not allow for the collection of random samples, penetration rates vary among carriers and usage rates vary across users (*29*).

Given these important limitations in passive data collection, long-distance researchers are continuing to experiment with innovative survey data collection methods. These include the experimentation with new survey and data analysis methods as well as the creation of specialized survey Apps that merge some of the benefits of surveys and device-based passive data collection (*11*). Researchers are developing methods to extrapolate long-distance travel patterns based on the timing of a respondent's single most recent long-distance trip (*30*), seeking to improve recall by tying travel to life events (*31*, *32*), and asking respondents for estimates of their long distance travel frequency rather than a record of long-distance trips themselves (*33*). Non-random sampling has also been explored on the hypothesis that enthusiastic survey takers may be more willing to tolerate high survey burden (*34*).

Survey Datasets

This section outlines five survey datasets that were used in the analyses covered in this project. Three of these surveys were original data collection efforts lead by the research team (the LSOT and PiYL) or included questions developed by the team (the Vermonter Poll). The two remaining two surveys were conducted by public agencies in California (the CHTS) and Vermont (the LRTPS).

Longitudinal Survey of Overnight Travel (LSOT)

The LSOT was a year-long, online, panel survey administered in 2013-2014. It was the first year-long survey of long-distance travel in the United States since the 1995 American Travel Survey. Given well-known short-comings with distance thresholds as a means of defining long-distance travel, most notably respondents' limited capacity to estimate distances accurately (*22*), the LSOT used overnight travel (defined in the survey as "a trip where you leave town AND spend the night somewhere other than home") as a means of identifying long-distance trips. In addition, because overnight travel is comparatively memorable, it was hypothesized that this definition would be less prone to recall errors than a distance threshold.

The LSOT consisted of an intake survey and twelve monthly follow-up surveys. The intake survey collected basic socio-demographic data as well as self-assessed estimates of overnight travel frequencies. Estimates of travel frequencies were recorded for five tour definitions and two travel purposes – work travel and personal travel. The five overlapping tour definitions were:

1) all overnight tours,
2) overnight tours that included air travel,
3) overnight tours that included intercity train travel,
4) overnight tours that included intercity bus travel, and
5) overnight tours that included a destination outside of North America.

Self-assessed travel frequencies were estimated on a five-point scale:

1) never,
2) very infrequent (less than once per year),
3) infrequent (once or twice per year),
4) frequent (multiple times per year), or
5) very frequent (multiple times per month).

In each subsequent month, participants received an email prompt linking them to a monthly survey collecting information on the travel that they had completed since the prior survey. For the monthly surveys, respondents recorded changes to their household structure, the number of day trips greater than 50-miles they had completed, as well as detailed information on their overnight travel. For each overnight tour, respondents provided tour start and end dates, travel modes(s), travel party composition, and overnight stop locations. Participants that missed a single monthly survey could provide information for two months of travel in the subsequent monthly survey but participants that missed two monthly surveys were dropped from the panel.

In addition to respondents recruited using employer-based email lists and community newsletters, the LSOT panel included transportation professionals and acquaintances of the research team invited by email on the hypothesis that these closely linked respondents would be more motivated than average to complete the survey. In total, 1,220 people started the panel and 628 completed the study, a retention rate of 51.5%. Participants known by the research team or working in transportation were only slightly more likely to complete the panel than other respondents. Additional information about the LSOT can be found in (*34*).

People in Your Life (PiYL) Pilot Survey

The PiYL pilot survey (Appendix A) was designed to gauge the geographic extent of a respondents' social network and to capture indicators of the level of their long-distance travel to facilitate modeling social network geography as a predictor of long-distance travel behavior. It was developed in response to focus group interviews with LSOT participants who indicated that many of their long-distance travel choices where influenced by the location of family, friends, and work activities. Comprehensive documentation of individuals' social networks and travel behaviors are both highly burdensome tasks. Thus, a primary goal of the PiYL pilot was to test the effectiveness of collecting more abbreviated social networks and travel data. After multiple rounds of testing and development, the pilot survey was administered in the winter of 2016-2017. It collected home locations for 13 individuals in each respondent's social network, travel frequency estimates for eight different trip types, and a limited slate of demographic information.

The pilot survey was administered to a total of 110 respondents recruited in Alabama, California, and Vermont. The Alabama-based respondents consisted of 65 engineering undergraduate and graduate students and several staff members at Auburn University. Twenty-one California-based participants living in greater Sacramento were recruited at the University of California Davis or from senior citizen participants in a University seminar program and twenty-four women were recruited from Burlington, Vermont. Additional information about the creation of the PiYL pilot survey and the demographics of the respondents can be found in Aultman-Hall et al (*15*).

The 13 individuals in each respondent's social network, referred to as "contacts," consisted of 10 contacts defined based on their relationship to the respondent (relation-based contacts) and three contacts selected based on home locations (location-based contacts). Respondents were asked to provide the home locations for 10 relation-based contacts according to the following criteria:

- three family members that did not live with the respondent;
- a person the respondent would go to for work or professional advice;

- a person the respondent would go to for personal advice;
- a good friend;
- a childhood friend;
- a person the respondent wishes they could spend more time with; and
- two people whom the respondent felt an obligation to visit.

In addition, respondents were asked to identify contacts with whom they had communicated with in the last one year that lived in specific, distant locations. The specified locations varied based on the respondent's home state and were selected based on the discussions with pre-test respondents. These three contacts are referred to as location-based contacts. Contacts were solicited in the following locations:

- New York, California, and Europe/Asia for Alabama-based participants;
- New York, Florida, and Europe/Asia for California-based participants; and
- Florida, California, and Europe/Asia for Vermont-based participants.

General long-distance travel behavior measures were collected by asking the respondents to estimate the frequency with which they undertook the following eight trips types:

Trips to destinations more than a 2-hour drive from home:

- To visit family or friends;
- For work; and
- For personal business such as a medical appointment, banking, or other services.

Trips meeting the following criteria:

- For vacation or leisure;
- That include air travel;
- With NO overnight stay that include air travel;
- With NO overnight stay and include 2 or more hours of driving EACH way; and
- That include a destination outside of North America.

Trip frequencies were recorded on a six-point scale:

- More than once per Month,
- Once per Month,
- Multiple Times per Year,
- Once per Year,
- Less than Once per Year, and
- Never.

These long-distance travel measures are relatively broad and recent analysis of both the LSOT and PiYL travel frequency estimates suggest that self-assessed travel frequency is not a reliable indicator of travel. Thus, these travel measures should be interpreted with caution and may be the reason for weak results associating social network location to travel level.

The Vermonter Poll

The Vermonter Poll is a telephone-based poll conducted annually by the Center for Rural Studies at the University of Vermont. The poll sample is drawn randomly from Vermont landline and cellular telephone numbers and uses computer-aided telephone interviewing (CATI) for data collection. The 2017 Vermonter Poll was conducted between the hours of 9:00 a.m. and 9:00 p.m. on weekdays beginning on February 21, 2017 and ending on February 28, 2017. The response rate for the 2017 Vermonter Poll was 20.1%, producing 613 valid responses.

A series of four questions about the respondents' most recent non-work long-distance trip were added to the 2017 Vermonter Poll as part of this project. Respondents were asked to think about their last overnight trip out-of-town for personal reasons (like to visit family or friends or take a vacation) and asked for the country, state, and place containing their travel destination as well as the primary mode of travel on the trip. Respondents were also asked two questions about their trip planning process:

Which of the following was most true about how you traveled?

- Travel mode (such as driving or flying) was decided first and destination was selected afterward.
- Destination was selected first and then multiple travel modes were considered.
- Destination was selected first and only one travel mode was considered.
- Destination was selected first and there was only one travel mode available.
- Don't Know.

Which of the following statements is most true about the trip's destination?

- My exact destination was known and I did not consider any other destinations.
- Multiple destinations in the same region were considered.
- Multiple destinations in different states, regions, or countries were considered.
- Don't Know.

Of the 613 total respondents, 552 provided at least the destination country for their most recent trip. The 61 non-respondents to this question may be non-travelers or simply been unable to recall the last trip meeting the criteria in the question.

California Household Travel Survey (CHTS)

The California Household Travel Survey (CHTS) is conducted every ten years by the California Department of Transportation (Caltrans). The 2010-2012 CHTS used address-based sampling and all households had the option of selecting from among multiple survey modes: Computer Assisted Telephone Interviewing (CATI), an online survey instrument, or a paper mail survey instrument (*35*). [1] Collecting long-distance travel was a point of emphasis in the 2010-2012 CHTS and was accomplished using a supplemental long-distance travel log (*35*). Respondents were instructed to use the long-distance log to report all trips that they had taken in the preceding eight weeks with a one-way travel distance of 50 miles or more. The CHTS included more than 42,000 completed household surveys and captured 77,000 long-distance trips (*35*) making it one of the best resources currently available describing long-distance travel in the United States.

Unfortunately, several challenges were encountered during the survey administration that, while contributing to increased knowledge regarding best practices for long-distance data collection, reduced the analysis options feasible with the CHTS long-distance data. After the pre-testing period, Caltrans and NuStats, the consulting company administering the survey, made numerous changes to the long-distance travel log. These changes included altering the layout of the paper version of the long-distance log without subsequently re-testing this instrument and changing the recall period for the log. In pre-testing, the recall period was limited to two-weeks and respondents with no long-distance travel in this time frame were asked to report their last long-distance trip regardless of the timing of that trip. During the administration of the full-survey the recall period was extended to eight weeks and the follow-up question for respondents that had not made a long-distance trip during the recall period was dropped. The pre-test version of the paper long-distance travel log was also determined to be to visually confusing and, consequently, several questions, including access and egress modes for air and transit trips, were dropped from the printed version of the long distance log. These questions were maintained in the CATI and online retrieval scripts for the respondents' most recent trip, resulting in differences in the data collected depending on the survey mode selected by the respondent. The final CHTS report also indicated a high degree of uncertainty about which household member had completed the long-distance log, indicating that in 42% of cases, the household member who completed the survey could not be determined (*35*). For these reasons, the analysis in this project is limited to single-person households in order to utilize person variables. Even with this challenge, the CHTS remains the largest recent long-distance travel survey in the United States to capture individual trips.

A public version of the CHTS dataset is available through the Transportation Secure Data Center through the National Renewable Energy Laboratory which redacts the latitude and longitude of trip origins and destinations. For this project, Caltrans shared these confidential location data for long-distance trip destinations, enabling us to calculate distances from home for these respondents.

The Long Range Transportation Planning Survey (LRTPS)

The Vermont Long Range Transportation Planning Survey (LRTPS) was administered by

[1] Note that there are multiple, similar but not identical versions, of the CHTS Final Report and Appendix available from Caltrans, the Transportation Secure Data Center, and other agencies. The Final Report/Appendix versions cited here are from the Caltrans website: http://www.dot.ca.gov/hq/tpp/offices/omsp/statewide_travel_analysis/chts.html.

Resource Systems Group, Inc. (RSG) for the Vermont Agency of Transportation (VTrans) in 2016. The LRTPS included 43 questions on travel behavior, customer satisfaction, policy and funding opinions, emerging trends in technology, and sociodemographic variables. It included five questions related to long-distance travel and to unmet travel needs. Specifically, respondents were asked how often they:

1. Made a trip that had a destination in Canada (Vermont borders the Canadian province of Quebec);
2. Made a trip that had a destination outside the U.S. or Canada;
3. Used commercial air services;
4. Needed to travel to a destination inside Vermont but could not due to lack of transportation options; and
5. Needed to travel to a destination outside Vermont but could not due to lack of transportation options.

Questions four and five measured unmet travel demand. For each question, frequency was measured using a 5-point Likert scale:

1. Never
2. Very infrequently (one time per year or less)
3. Infrequently (multiple times per year)
4. Frequently (multiple times per month)
5. Very frequently (multiple times per year)

The decision to use inside and outside the state and country as definitions of long-distance travel was a deliberate plan to avoid the distance-based thresholds which respondents have difficulty with. Ultimately, the respondents who live along the state boundaries made this wording a weak choice by our team if purely long-distance trips were of interest. The survey used random address-based recruitment and could be completed using either an online or paper mail-back survey instrument. It has a sample size of 2,232 (42 percent were completed online and 58 percent were paper surveys) among five Vermont regions, each containing a minimum of 347 surveys. Additional information about the LRTPS is available in (*36*).

Alternative Methods for Long-Distance Data Collection

Four journal articles related to addressing different aspects of long-distance data collection are summarized here. The first of these articles advances the discussion of appropriate definitions of long-distance travel, a prerequisite to effective data collection. The three subsequent articles test the potential utility of non-random sampling, using one-time, self-assessed estimates of travel frequency, and collecting data on social network geography to predict travel. Each of these methods could reduce the burden of long-distance data collection and facilitate greater understanding of this travel.

The interrelated research efforts in this project illustrate the continuing challenges in long-distance data collection while also illuminating promising avenues for future research. Analysis of LSOT overnight-based data strongly re-demonstrates that distance thresholds are a poor method for defining long-distance travel on a national scale and that tour structures are often more complex than mirrored out-and-back trips. The LSOT dataset also provides evidence that convenience samples may be suitable for studying the distribution of long-distance trip lengths and the spatial distributions of destinations though not rates of trip-making. It also suggested that self-assessed travel frequency estimates can provide only a crude approximation of long-distance travel and that the self-assessment is most effective for identifying non-travelers and very infrequent travelers. Assessing long-distance travel based on social network geography showed some promise but requires improved methods for quantifying social network extent (one such method is described below). Modeling the relationship between social network extent and long-distance travel remains a challenge, however, since datasets of significant size that capture the full extent of both individuals' long-distance travel and social networks are extremely rare.

Long Distance Trip Classification Schemes

Developing a defensible definition of long-distance travel is a necessary precursor to a national long-distance data collection. Current surveys often rely on single distance thresholds to define long-distance travel but threshold selection is highly variable and not well-grounded in empirical data. The most common 50-mile definition was established at a time when commutes were shorter and mega-regions had not grown beyond this length. Moreover, lumping "short" long-distance trips within a single mega-region together with transcontinental and international travel, complicates modeling efforts. This section of this report summarizes a paper aimed at developing a more comprehensive classification scheme for the highly complex long-distance travel using a year of overnight trips collected from 628 individuals with the LSOT. A more complete description of this effort is available in (*13*).

Each month, LSOT respondents were asked to document the start and end dates of the overnight tours they had completed since the last survey, the location of all overnight stops on the tour, stop purposes and modes of travel. The total number of tours, total travel distance in miles, and total number of days away from home were calculated for each respondent. Somewhat surprisingly, these measures of travel (number, distance, and duration of tours) were relatively independent of one another and there were no consistently identifiable differences between work and personal tours. Exploratory analysis of these tour logs provided insights into the sample's tour frequency, tour length distributions (as well as modal choice by tour length), and the spatial patterns of these tours. Negative binomial regression was used to model tour generation (in tours per person per year) for different ranges of tour length and by tour purpose to test different classification schemes.

Overall, the LSOT sample was composed of frequent travelers with respondents reporting a mean of 9.3 overnight tours per year. As would be expected, tour frequency declined as tour length (measured as maximum distance from home) increased. Surface travel was the dominant mode of travel for shorter trips, with a notable increase in air travel at 250 miles. A comparison of mode choice and trip distances suggested a possible four-tier trip classification scheme consisting of regional travel (50 – 249 miles), inter-regional travel (250-499), continental travel (500-2,999) and global travel (>3,000 miles).

Spatially, tours could be characterized as simple out-and-back tours or complex tours consisting of either a circular chain, where a series of destinations were visited in sequence without repeated visits to any single destination, or a hub-and-spoke pattern, where the traveler went to one destination and then made a series of out and back sub-tours from that location. Within the LSOT sample, close to 80% of tours were simple out-and-back tours while 16% were hub-and-spoke and 4.5% were circular chains.

Regression modeling was conducted by tour purpose and by distance thresholds. Model fit was relatively weak across all models with models of shorter tours performing particularly poorly. While no single predictor variable was significant across all models, education and income were significant in several models and linked to higher travel frequency, as expected.

This paper provides further evidence that distance thresholds are a flawed metric for defining long-distance travel on a national scale. The distance thresholds examined performed poorly at defining a unique type of travel and showed evidence of regional variability.

Consideration of Non-Random Samples for Long-distance Trip Attributes

Because of the heavy burden of long-term long-distance travel surveys, the LSOT relied on a convenience sample, a group we considered "keen survey takers" after most participants were willing to attend post-survey focus groups. While this non-random sample was not demographically representative of the population, it is possible that the travel behavior of keen survey takers does not differ substantially from the general public. Using convenience samples for some research would open new avenues for collecting long-distance travel data. Comparing the travel behavior captured with the LSOT convenience sample to the travel recorded in the 1995 ATS and to a very large sample of passively collected travel data from phone locations processed by AirSage, Inc. in 2013 provides evidence that convenience samples may provide useful information on trip distance and trip destination distributions but is not accurate for estimating trip generation rates. Additional information about this comparison is described in (*14*).

AirSage contracts with two cellular companies to acquire location data for devices with cellular connections (phone, tablets, hotspots, etc.). Device locations are calculated based on cell tower triangulation whenever a device transmits any type of information (voice, text, data usage). Anonymized device records, referred to for convenience as respondents, were assumed to represent single individuals, though multiple devices may be associated with a single individual in actuality. The AirSage data used in this analysis included respondent records from all devices with home locations in Chittenden County, Vermont and Lee County, Alabama in May and October of 2013, the year that the LSOT survey was conducted. These data captured approximately 15% to 30% of the population in Chittenden and Lee counties. ATS data were collected quarterly from a panel of survey respondents from April 1995 through March 1996. For the ATS, respondents were to report on all trips to a destination at least 100 miles from home.

Because of differences in the data collection goals and methods among the ATS, the LSOT, and AirSage, one of the main research challenges was determining a common measure of long-distance travel that could be used for comparing the data. The best common unit of travel that could be calculated for all three datasets was a count of the number of days that an individual spent all or part of the day at a long-distance destination, defined in this analysis as any county (AirSage and LSOT) or MSA (ATS) at least 250 miles from the respondent's home county. The research team termed this measure the respondents' *destination-days*. Development of the destination-days measure was driven primarily by the structure[2] of the AirSage data which consisted of a device ID, its home county, all other counties in which the device had reached a destination, and the number of days the devices was located at each destination county. Each non-home county destination was treated as a destination day. Because AirSage's algorithm for identifying destinations is proprietary, it was not possible to assess the magnitude of any errors in destination identification, but it is likely that the data includes some erroneous en-route stops as destinations and thus overstates destination-days to some degree. For the LSOT, destination-days were assumed to equal one more than the number of nights at all over-night destinations on each long-distance tour. For the ATS, destination days were derived from the Person Trips and Person Side Trips tables at the MSA level. After these calculations, destination-days served as a common measure of long-distance travel across all three datasets.

Visual comparisons of the frequency distribution of destination-days spent at various distances from home (for trips over 2250 miles, binned in 100-mile intervals), showed relatively similar travel patterns across the three datasets. Unsurprisingly, the greatest proportion of trips fell into the 250-450 miles range with trip frequency tending to diminish at longer distances. While the frequency distributions did differ, they showed many commonalities. Vermont respondents, for example, showed a spike in travel frequency that was apparent in all datasets

[2] Note AirSage constructed this data format specifically for our research project and this is not the typical origin destination matrix format produced by AirSage for transportation planning agencies.

to destinations between 1,150 and 1,350 miles away. Overall, this comparison provided evidence that reasonable trip length distributions may be obtainable from convenience samples.

To look at the distribution of destinations in the AirSage and LSOT datasets, destination-days were aggregated into equally sized hexagonal areas, to prevent bias due to variable county sizes. Visual assessment of the proportion of total destination-days spent in each destination hexagon showed similar patterns with between the two dataset. For Chittenden County respondents, travel was heaviest along the coasts and to major cities in the interior of the United Sates in both the LSOT and AirSage datasets. The correlation between travel volumes to different destination hexagons in the LSOT and AirSage datasets was further explored using Ordinary Least Squares (OLS) regression models, estimating the total number of days spent at each destination hexagon in the AirSage data as a function of LSOT number of days spent in the hexagon in the LSOT dataset. Separate models were created for Chittenden County and Lee County respondents. Both models also included the distance from the home hexagon to the destination hexagon as an independent variable. The Lee County model also included distance squared and distance cubed. All predicator variables in both models were significant based on a P = 0.05. Both the Chittenden and Lee County models showed a significant positive relationship between the LSOT and AirSAge data. The adjusted R^2 for these models was 0.75 and 0.59 respectively and the samples sizes for the two models were 352 and 344. Because of the inherent impact of population and destination clustering on travel behavior, investigating the influence of spatial autocorrelation is an important step in OLS analysis. The LSOT data, for example, includes a high proportion of trips from the east coast to the west coast implying distance was not necessarily a barrier to travel. However, the relatively lower population of the middle non-coastal states dictates relatively lower levels of attraction for these destinations. Global Morin's I was used to examine spatial autocorrelation for the Chittenden and Lee County OLS models but did not find significant evidence of autocorrelation after the distance terms were added (I = -0.004, p=0.356 for Chittenden County; I=-0.004, p=0.283 for Lee County).

Unlike trip distance and trip destination distributions, long-distance trip-making rates differed substantially between the LSOT and AirSage samples, especially for Lee County respondents. Trip-making was more than twice as common among Lee County LSOT respondents as opposed to AirSage respondents for May, 68% to 30%, and October, 56% to 26%. Overall trip-making rates were more similar for Chittenden County respondents but still differed substantially for distance thresholds. It is unclear why the Chittenden sample and Lee County LSTO samples differ in their effectiveness at capturing trip-making rates as both samples were recruited with the same methods and showed similar demographic biases. This may simply reflect the relatively small sample size in the LSOT or may reflect regional differences in how individuals engage in long-distance travel.

At least with respect to the Lee County sample, the keen survey takers that completed the LSOT were also keen travelers, taking long-distance trips at twice the rate of their AirSage counterparts. The LSOT and AirSage respondents did exhibit similar patterns in terms of distribution of long-distance trip lengths and destination spatial distribution. These findings suggest that online surveys utilizing lower cost convenience samples could be appropriate for destination choice analysis. Complementing passive datasets that lack demographic data with smaller survey datasets that capture this information would support modeling and forecasting.

Self-Assessed Travel Frequency

If individuals are able to accurately estimate their long-distance travel frequency, multiple, burdensome long-distance travel logs could potentially be replaced by one-time questions related to estimated travel frequency for many long-distance research applications. The ability to capture typical long-distance travel behavior in this manner would dramatically reduce the cost of long-distance data collection and expand long-distance research opportunities. To explore whether or not self-assessed travel-frequency is a viable proxy for long-distance travel, this analysis draws on the unique LSOT data set which included both one-time, self-assessed long-distance travel frequency estimates and monthly long-distance travel logs. Consistency between self-assessed travel frequencies and the number of tours recorded in the monthly long-distance logs would provide evidence for the accuracy of self-assessed travel frequency estimates. This work is described in more depth in Dowds, Aultman-Hall and LaMondia (in review).

As described in the *Survey Datasets* section of this report, LSOT respondents estimated their general long-distance travel frequency in month one of the survey and reported their realized travel, based on monthly recall, in 12 subsequent monthly surveys. For the LSOT, long-distance travel was defined as overnight, out-of-town trips. Using the self-assessed travel frequencies and the total number of recorded tours from the monthly surveys, three specific research questions were examined:

1. Are self-assessed travel frequencies for overnight, air-travel, and non-North American tours consistent with the number of total recorded tours for these tour types?
2. Does the degree of consistency differ between work and personal overnight tours?
3. Can the tendency to over- or underestimate travel frequency be modeled using information about the respondents' socio-demographic and travel behavior?

Overall, the level of consistency between self-assessed travel frequency and total recorded tours was under 70%, 69.7% for personal travel and 68.5% for work travel. The level of consistency in these estimates, in part, reflects the relatively large bin size for the "frequent" travel estimate which spanned 3 – 12 trips per year as well as high consistency among non-travelers and very infrequent travelers. Consistency was highest for tours to non-North American destinations (83.4% for personal travel and 93.9% for work travel), reflective of very high levels of consistency among those respondents who reported in the initial survey that they rarely or never took trips of these types. While prior research has shown the people tend to underestimate their trip-making, due to recall errors and other factors, the self-assessed travel frequencies reported in the LSOT included significant over- and underestimation relative to the number of total recorded tours in the monthly surveys. Tours involving air travel and tours to non-North American destinations were more prone to overestimation than tours that did not meet these criteria, perhaps reflecting that the perceived prestige of these types of travel introduced bias in respondents' self-assessments.

For the 576 respondents who were employed at the start of the survey, and therefore provided estimated travel frequencies for both personal and work travel, only 47.4% were consistent estimators for both travel purposes. Chi-square testing showed no evidence that respondents who consistently estimated travel for one purpose were more likely to consistently estimate travel for the other purpose. Respondents who misestimated their travel frequency for both work and personal travel, showed a tendency to systematical over or under-estimate for both tour types.

Multinomial logit models showed limited ability to predict the tendency to over- or underestimate travel frequency based on respondents' socio-demographics and travel behavior. Respondents who reported higher levels of trip making for all purposes in the monthly logs were more likely to underestimate (and less likely to overestimate) their annual travel for all overnight personal travel, personal travel including air transportation, and overnight work travel. Men were more likely than women to overestimate personal travel, personal travel involving air, and work travel involving air, possible suggesting greater positivity bias in male respondents.

Overall, these analyses suggested that self-assessed travel frequency estimates can provide only a crude approximation of long-distance travel. Self-assessment was most effective for non-travelers and very infrequent travelers and for highly memorable trip types, especially travel to non-North American destinations. We recommend the development of alternative measures of long-distance travel for future surveys, beyond a total frequency of travel in a recall period.

Surveying Social Network Geography to Model Long-Distance Travel

The PiYL social network survey was developed in response to focus groups with LSOT participants who indicated that the residential locations of family and friends had a major influence on their long-distance travel choices. Like long-distance travel surveys, comprehensive surveys of an individual's social network are quite burdensome and the completeness of the results can be diminished by respondent fatigue and recall errors. In order to explore the relationship between social network geography and long-distance travel, the PiYL survey needed to collect information on both of these challenging topics. The social network and long-distance questions in the PiYL pilot survey (Appendix A) were necessarily limited in their scope to reduce survey burden but were intended to be representative of broader social network extent and travel behavior patterns. Collection of these pilot data allowed for preliminary analysis of the relationship between social network geography and long-distance travel. A more extensive description of this dataset and analysis is available in (*15*).

While the PiYL survey participants were not recruited randomly, the 110 respondents did have significant variability in age, gender, education and income. The sample provided home locations for 992 relation-based contacts and 142 location-based contacts. Respondents tended to have fairly close emotional relationships with their relation-based contacts, reporting an emotional closeness of 7 or higher on an 11 point scale, where 0 indicated "not at all close" and 10 indicated "very close," for 66% of the relational-contacts. Respondents were also in relatively frequent communication with these contacts, reporting face-to-face interaction within the last month with 51% of the contacts and telecommunications with 75% of the contacts in the same time frame. The physical distance between the respondents and their relational-contacts was highly variable and not significantly correlated with emotional closeness. Unsurprisingly, the emotional closeness for the location-based contacts was generally lower than then the relational-contacts, though once again emotional closeness and geographic closeness were not correlated.

Social network extent, measured by the logged average distance to each respondent's contacts, was modeled against typical socio-demographic predictor variables: gender, age, presence of children in the household, household income, educational attainment, telecommuter status, and state of residence. The resulting models were quite weak but some variables, such as income, were significant. Further exploration with a larger sample is merited to understand the relationship (or lack of relationship) between social network extent and socio-demographics.

Travel frequency data collected for eight different trip types, described in Section 0 of this report, showed little relationship between the travel frequencies of different trip types. In other words, respondents who traveled frequently for one type of trip were not notably more likely to travel frequently for other types of trips. Moreover, the respondents tended to show a similar overall level of long-distance travel across all trip types, though this finding may be an artifact of the small sample size and non-random sampling. Overall, many respondents reported relatively infrequent travel for several trip types indicating that the 2 – 8 week recall periods used in many traditional surveys would be inadequate to capture the full variety of long-distance trips that individuals make over the course of a year.

Ordered Probit models for travel frequency by trip type were created to test the relationship between long-distance travel behavior and social network extent. Predictor variables included socio-demographics as well as the maximum distance to contacts, the number of locational contacts provided, the average distance to emotionally close relational contacts, counts of at-distance emotionally close contacts, and dummy variables indicating that no locational contacts or all 3 locational contacts were provided. The model for travel frequency to visit friends and family performed the best of any of the eight trip type models though it had a relatively low McFadden R^2 of 0.06. The inability to generate stronger models may be the result of small sample size and weak trip frequency measures, or may reflect an underlying lack of connection between social network geography and long-distance travel.

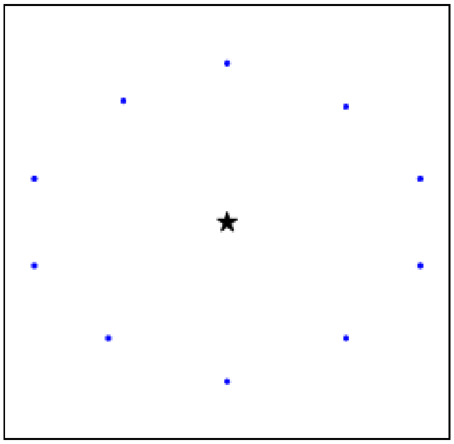
Though the relationship between social network geography and travel was tenuous in this small data set, the PiYL dataset did show significant variation in social network extent, suggesting that the social network survey questions may be effective at capturing the general extent of social network geography. The relatively weak measures of long-distance travel used in the survey, along with high variability in both network geography and long-distance travel behaviors did not produce strong models but nevertheless provided evidence that this avenue of inquiry merits further study.

Social Network Sizes and Shapes: Classifying the Geography of People in Space

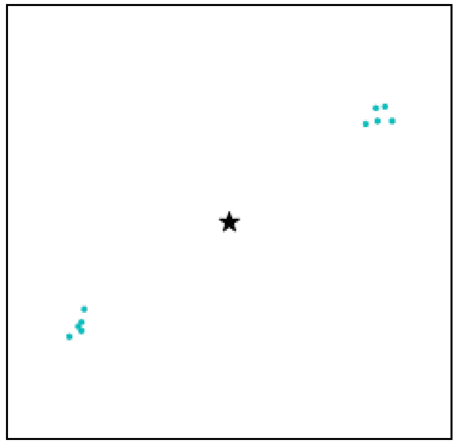
As activity-based approaches to travel modeling have grown in prominence, there has been an increasing recognition that the spatial characteristics of individuals' social networks – that is where individuals' family, friends and other important contacts live – influence their travel behavior (*37*). With social networks expanding in size at the global level due to migration, telecommunications, and other factors, social networks will be an increasingly important driver of long-distance travel. Effective methods for quantifying social network extent are necessary to effectively connect social network characteristics to long-distance travel. Current social network measures, such as average distance and confidence ellipses, have limitations and better ways to measure the spatial extent of these networks are still being pursued. A new social network classification method utilizing the distances from the respondent to all of their social network contacts as well as the distances between all contact pairs in the respondent's social networks is presented here using the social network data collected in the PiYL survey.

The quantification of egocentric social networks – meaning networks that consist of a set of ties, or contacts surrounding a sampled individual, or "ego" – is not simple or straightforward. The earliest work that sought to measure social networks were only concerned with the number of network members and not the spatial distribution of these members (*38*, *39*). Utilizing measures of social network geography as a predictor of long-distance travel behavior requires methods that characterize social network location, accounting for network spatial complexity, rather than simply measuring ego-contact pairings. The simplest way to capture this would be to sum the distances of all ties in the network, but this fails to capture spatial distribution patterns such as clustering or ego isolation (*40*). The common method for the measurement of network spatial "size" is the confidence ellipse method. The confidence ellipse method is a parametric method defined by a fixed percentage confidence region, first presented for the measurement of a person's activity space by Schönfelder (*41*). This has become the standard measure for egocentric social networks because it is easily computable and has been found to correlate with other more difficult to calculate methods (*40*, *42–45*). The area of the ellipse, centered on the ego's home location, represents the network size and the calculation of the ellipse works under the assumption that the locations are normally distributed. Concerns have been raised by many who have utilized this method that the ellipse area as the network size measure is an over-representation of space, partially due to the ellipses being necessarily symmetrical and the assumption of continuity of space. The original use of the confidence ellipse was for the measurement of a person's daily activity space (*41*), which tends to have a smaller localized spatial distribution. Adapting this method for globally distributed, egocentric personal network extent further diminishes the accuracy of this tool since more of the area captured by the ellipse is likely to consist of empty space such as bodies of water or deserts.

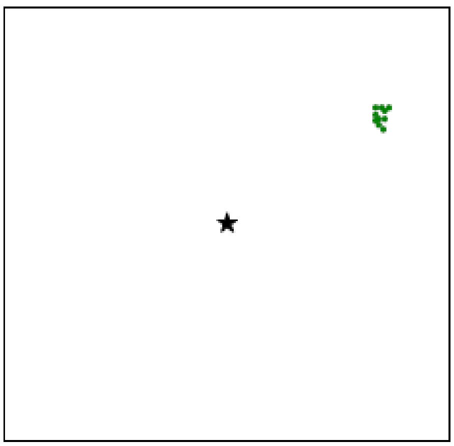
Several conceptual scenarios, demonstrating different extremes in network extent, were developed in this project as an initial step towards classifying the geographic distribution of egocentric social networks collected in the PiYL survey. These six conceptual ("small-world") social network types are shown in Figure 1. In each plot, the ego is shown at the center as the black star and the closest contacts are distributed around them, indicated by colored dots. Images A and C in Figure 1 show two different extreme scenarios, where all of a person's closest contacts are at a far distance. In the first case (A) the contacts are distributed uniformly around the ego and in the second case (C) they are clustered in one direction with respect to the ego. Whether contacts are all in one direction or in many different directions might have an effect on travel by the ego, considering that if the contacts are all in one location then one trip could allow the ego to interact with all of its closest contacts. Four distance-based measures were considered as the basis for the classifying these networks quantitatively. These measures are the mean distance and variance in distance from the ego to each contact (referred to as the "ego-to-contact" or ETC measures) and the mean distance and variance in distance from each contact to every other contact, (referred to as the "contact-to-contact" or CTC measures). Conceptual ETC and CTC measures are also provided in Figure 1. While real-world social networks are expected to have more variability than those shown in these six conceptual examples, different typologies of social network may be identifiable using cluster analysis of ETC and CTC measures.



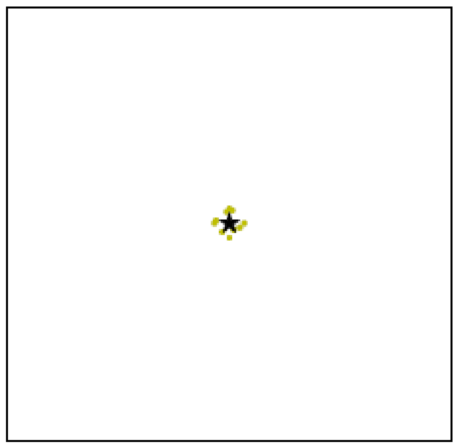
A. Isolated Ego - Disparate:
ETC: mean = HIGH, variance = LOW
CTC: mean = HIGH, variance = HIGH



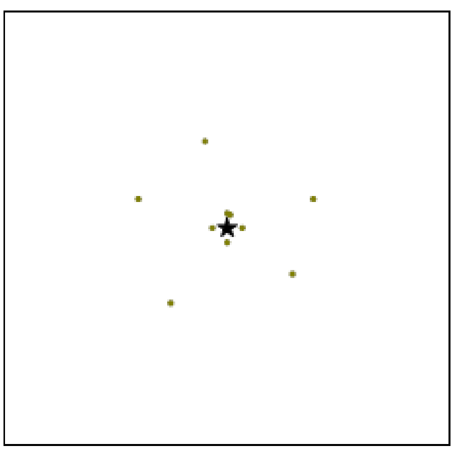
B. Isolated Ego - Polar:
ETC: mean = HIGH, variance = LOW
CTC: mean = HIGH, variance = VERY HIGH



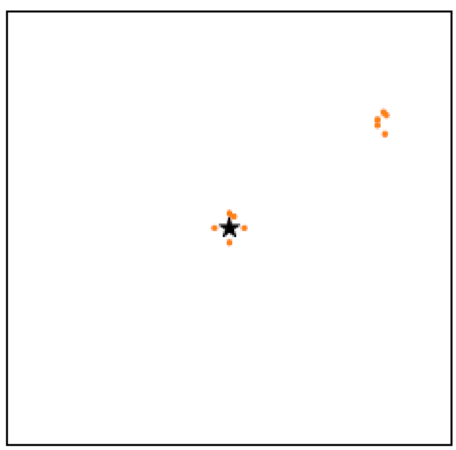
C. Isolated Ego - Clustered:
ETC: mean = HIGH, variance = LOW
CTC: mean = LOW, variance = LOW



D. Tight-and-Close:
ETC: mean = LOW, variance = LOW
CTC: mean = LOW, variance = LOW



E. Near-and-Less-Far:
ETC: mean = MED, variance = MED
CTC: mean = MED, variance = MED



F. Near-and-Clustered-Far:
ETC: mean = MED, variance = HIGH
CTC: mean = MED, variance = HIGH

Figure 1. Conceptual social network examples

Empirical ETC and CTC distance statistics for the social networks of PiYL respondents were calculated using the latitudes and longitudes for each respondent's home location and that of their contacts using the great circle distance method. The average, standard deviation, and coefficient of variance of the ETC distances and CTC distances were used as descriptor variables for the respondents' networks. A summary of the ETC and CTC variables for the PiYL dataset can be found in Table 1.

Table 1. Summary of Distance Variables for all respondents' social networks

DISTANCE (MILES)		MEAN	STD. DEV.	MIN.	25%	50%	75%	MAX.
Ego-to-Contact	Average	523.5	699.3	24.2	121.6	329.6	619.1	5250.9
	Standard Deviation	585.7	663.7	21.9	149.9	355.6	782.1	4206.2
	Coefficient of Variance	24.0	13.4	4.5	13.4	20.5	33.7	73.4
Contact-to-Contact	Average	708.3	800.6	39.0	205.2	383.3	974.0	5006.6
	Standard Deviation	656.6	704.8	27.3	184.2	407.5	934.9	4084.7
	Coefficient of Variance	22.6	12.2	4.4	12.9	19.6	30.3	65.7
	Number of Contacts	9.1	1.2	5	8	10	10	10

K-Means clustering was performed on four candidate sets of ETC and CTC distance variables, shown in Table 2. The candidate sets compared the effectiveness of using the standard deviation to the coefficient of variance for the ETC/CTC as well as the inclusion/exclusion of a number of contacts as a clustering variable. The score distributions for these four candidate sets can be seen in **Error! Reference source not found.**. Set 4 included the average and coefficients of variance of the ETC and CTC distances for each respondent and was selected for the final clustering criteria because it achieved a higher score than clustering with the standard deviation. Inclusion or exclusion of the number of contacts variable had limited importance on cluster score and did not change how respondents were clustered.

Table 2. Candidate Clustering Variable Set

Variables	Variable Sets			
	Set 1	Set 2	Set 3	Set 4*
ETC Average Distance	X	X	X	X
ETC Coefficient of Variance		X	X	X
ETC Standard Deviation	X		X	
CTC Average Distance	X	X	X	X
CTC Coefficient of Variance		X	X	X
CTC Standard Deviation	X		X	

Number of Contacts	X	X

* Final Variable Set

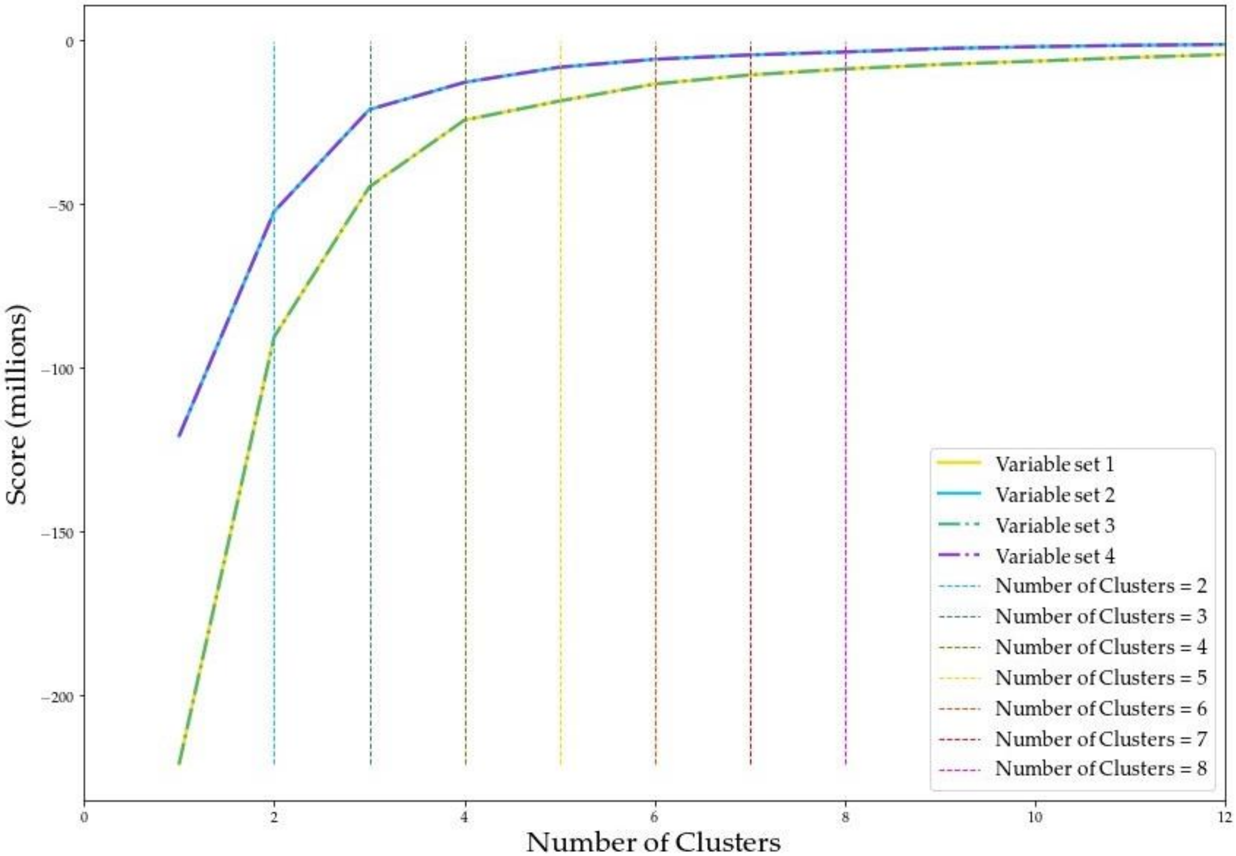


Figure 2. Scores of k-means cluster candidate variable sets number of clusters

The clustering distributions produced by clustering into 1 to 20 total clusters with the final clustering variable set is shown in Figure 3. The number of clusters used for the final analysis should result in meaningfully sized clusters – that is clusters that are small enough to distinguish among respondents based on important differences but not so small that minor differences between respondents separates them into different groups. At the most extreme scenarios, using a single cluster would group all respondents together while using as many clusters as respondents would result in each respondent having their own group – neither of which provide useful information about the respondents. Based on the clustering distribution results, we assessed that six clusters of respondents succeeded in creating unique groupings, reflecting significant differences in terms of the social network geography variables, while the additional groupings created when using more than six clusters were very small in size and not appreciably different from the groups produced with six clusters. For this reason, the final analysis was conducted with six clusters.

Once the PiYL respondents had been clustered into six groups, each respondent's social network was mapped and visually inspected. The six clusters were categorized based on the characteristics of the spatial distributions of the social network. Categorizations incorporated the distance from the ego to other contacts (regional, continental, or global) as well as the degree of concentration among the contacts (dispersed vs. polar – only two or three unique locations). The ETC and CTC variables for each of the six clusters are summarized in Table 3.

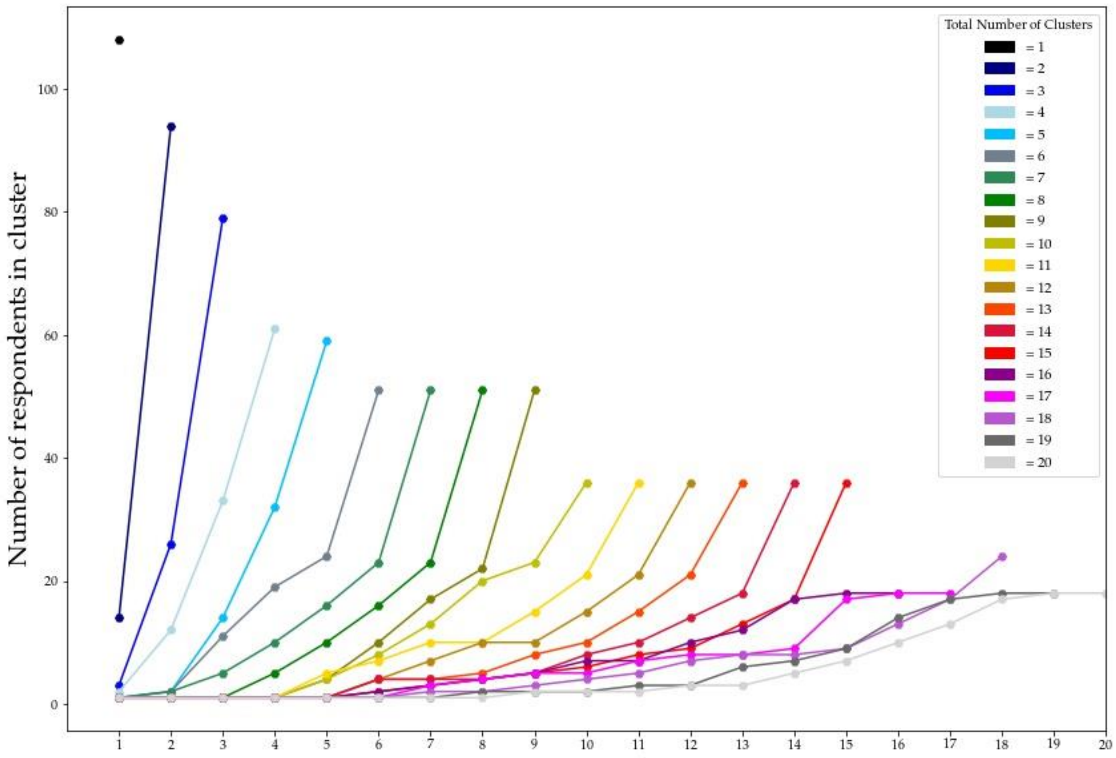


Figure 3. Clustering Distribution for Variable Set 4

The largest group, Cluster 1, was characterized as "regional", since these social networks were dominated by contacts living in the same region as the ego. The 51 respondents whose social network geographies were regional had fairly low average long-distance trip frequencies using air and to international destination, but the highest average frequency for visiting family and friends. This group was 66% male and dominated (75%) by respondents between the ages of 21 and 24 years old. The second largest cluster, Cluster 6, consisted of 24 respondents with "polar continental" social networks, meaning they were contained within the country or continent of the ego, and that most contacts lived very close to the ego while a small number lived very far away in only one or two unique locations. This cluster had fairly high average long-distance trip frequencies in general, though not for international travel. It was 58% female and had an average age of 35 years. Cluster 3 was categorized by "dispersed global" social networks and included 11 respondents. This cluster was predominantly male, highly educated, older, and had high long-distance travel frequencies across all but one of the respondents. The last of the larger clusters is cluster 4, which was predominantly "dispersed continental" networks. These respondents were 70% female, had the oldest average age, 50 years old, and had very similar average long-distance trip frequencies to Cluster 6, the other "continental" cluster. Clusters 2 and 5 contained only three individuals in total and consisted of networks where contacts homes were on the opposite side of the world from the ego. These clusters were uncommon in this small sample, possibly due to sample size, and need to be assessed further, both in terms of our clustering techniques and our calculation of distances when the contacts are halfway around the world.

Table 3. Summary of Cluster variables by Cluster Type

Cluster Type	Distance (Miles)	Mean	Std	Min	Max
Cluster 1: Regional	ETC Average	125.8	59.6	24.2	272.7
	ETC Coefficient of Variance	15.1	6.6	4.5	37.7
	CTC Average	201.4	97.4	39.0	377.7
n=51	CTC Coefficient of Variance	14.3	6.2	4.4	35.7
Cluster 2: Polar Global	ETC Average	2682.6	597.6	2260.0	3105.2
	ETC Coefficient of Variance	57.6	2.3	56.0	59.2
n=2	CTC Average	3783.1	748.6	3253.7	4312.5
	CTC Coefficient of Variance	50.3	1.2	49.4	51.1
Cluster 3: Dispersed Global	ETC Average	1411.9	372.4	1026.9	2380.9
	ETC Coefficient of Variance	39.6	16.0	19.4	73.4
	CTC Average	1752.9	352.1	1312.7	2376.5
n=11	CTC Coefficient of Variance	38.2	13.0	23.8	65.7
Cluster 4: Dispersed Continental	ETC Average	746.3	183.1	516.5	1192.3
	ETC Coefficient of Variance	31.2	7.1	19.9	42.6
	CTC Average	1073.2	144.4	827.2	1353.7
n=19	CTC Coefficient of Variance	27.5	6.2	19.5	42.2
Cluster 5: Polar Global	ETC Average	5250.9	N/A	5250.9	5250.9
	ETC Coefficient of Variance	58.0	N/A	58.0	58.0
n=1	CTC Average	5006.6	N/A	5006.6	5006.6
	CTC Coefficient of Variance	57.7	N/A	57.7	57.7
Cluster 6: Polar Continental	ETC Average	408.1	71.6	262.7	583.4
	ETC Coefficient of Variance	25.9	10.0	8.5	41.6
	CTC Average	582.5	153.1	298.5	889.1
n=24	CTC Coefficient of Variance	25.3	9.4	10.3	39.9

A breakdown of travel frequency for the four larger social network clusters is presented in Table 4. Trips involving air travel and to non-North American destination (note that these are overlapping categories) were more common among respondents with continental and global social networks than those with regional clusters. Conversely, respondents with regional social network had the high frequency of visiting family and friends. Both of these results are consistent with the hypothesis that social network extent influences personal travel decision-making.

Table 4. Travel Frequency by Social Network Cluster Type

Trip Type:	Trip Frequency:	SOCIAL NETWORK CLASSIFICATION CLUSTERS			
		Regional	Polar Continental	Dispersed Continental	Dispersed Global
Trips to Visit Family/Friends[1]	once per month or more	45%	25%	11%	27%
	multiple times per year	43%	63%	74%	45%
	once a year or less	10%	13%	16%	18%
	never	2%	0%	0%	9%
Personal Business Trips[1]	once per month or more	8%	8%	0%	9%
	multiple times per year	22%	13%	0%	0%
	once a year or less	29%	38%	26%	36%
	never	39%	42%	68%	45%
Work Trips[1]	once per month or more	10%	13%	0%	27%
	multiple times per year	20%	17%	26%	18%
	once a year or less	25%	25%	42%	18%
	never	43%	46%	26%	27%
Vacation or Leisure Trips	once per month or more	10%	17%	5%	18%
	multiple times per year	65%	54%	79%	55%
	once a year or less	24%	29%	16%	27%
	never	2%	0%	0%	0%
Trips Including Air Travel	once per month or more	0%	0%	0%	9%
	multiple times per year	14%	54%	47%	36%
	once a year or less	65%	46%	47%	55%
	never	22%	0%	0%	0%
Air Trips with No Overnight Stay	once per month or more	0%	0%	0%	0%
	multiple times per year	4%	0%	0%	9%
	once a year or less	12%	25%	26%	18%
	never	82%	75%	63%	73%
Driving Trips With No Overnight[2]	once per month or more	6%	13%	5%	9%
	multiple times per year	33%	33%	21%	18%
	once a year or less	33%	29%	47%	45%
	never	27%	25%	21%	27%
Trips Out of North America	once per month or more	0%	0%	0%	0%
	multiple times per year	0%	0%	0%	18%
	once a year or less	39%	67%	79%	64%
	never	61%	33%	16%	18%
Total respondents in cluster		51	24	19	11

[1] Trips to a destination more than 2 hours from where the respondent currently lives.
[2] Including 2 or more hours of driving each way.

This new approach to categorizing social networks that takes into account not only the distances from the respondent to their contacts, but the distances between each contact in the social network is able incorporate the geographic extent of the networks when compared to the more basic approach (e.g. the average distance to contact method). Preliminary examination of the PiYL dataset shows coherent patterns in the estimated travel behavior for the respondents in the larger clusters. Further work should be conducted with a larger sample size to analyze this method of categorization of social network geography against other continuous methods such as confidence ellipse area, and the topic of spatial measures of these small collections of globally distributed locations should continue to be investigated and discussed. As discussed above, the future research should also use a different, more accurate measure of level of long-distance travel.

Equity in Long Distance Travel

Historically, transportation equity research has focused on the basic necessities obtainable by access to local goods and services within a person's home community. Long-distance transportation equity has received scant attention both because of the lack of long-distance data and because it is often considered non-essential in comparison to local travel. As global mobility and inter-connection increases, however, access to long-distance travel is likely to play an increasingly important role in supporting well-being. Long-distance travel serves multiple purposes that cannot be satisfied with local travel or with increasingly ubiquitous telecommunication options. These include maintenance of one's social network as families and friend groups become more widely dispersed, and the ability to access economic opportunities, cultural opportunities, and specialized services (e.g. some types of medical care) that tend to be available only in specific (often large metropolitan) locations. Underprivileged groups with limited access to long-distance travel may experience reduced well-being relative to groups that have greater access to long-distance travel.

This section provides an overview of the influence of long-distance travel on well-being based on interviews and of unmet long-distance travel needs documented in the Vermont LRTPS (Ullman, 2017). In addition, it presents exploratory analysis of evidence of disparities in realized long-distance travel behavior in the 2010-2012 CHTS and the 2017 Vermonter Poll. Across these multiple datasets, there is evidence that lower-income individuals travel less, have more unmet long-distance travel needs, and that long-distance travel is positively correlated with an individuals' own sense of well-being.

Disparities in Long-Distance Travel Access and Influence on Well-Being

Transportation equity can be studied by looking at disparities in realized travel among different sociodemographic groups but also by asking individuals explicitly about their unmet travel needs. Analysis of questions about unmet travel needs in the Vermont LRTPS, as well as of 24 original interviews with Vermont women that explored how long-distance travel impacts well-being and what barriers prevent individuals from accessing long-distance travel, revealed that latent demand for long distance travel in Vermont is linked to socio-demographics and impacts individuals' sense of well-being. A more complete description of this work can be found in a recent UVM graduate thesis (Ullman, 2017).

The LRTPS asked respondents how often they needed to travel to destinations inside and outside of Vermont but were unable to due to lack of transportation options. This provided an unusual opportunity to look at unmet demand for long-distance travel. (Note that travel to destinations outside of Vermont was considered a proxy for long-distance travel in this context). The LRTPS data was used to assess the following research questions:

1. What is the extent of unmet demand for long-distance travel?
2. Is unmet long-distance travel need correlated with the unmet need for local travel?
3. How does long-distance travel need vary with socioeconomic, geographic, and household characteristics?

The LRTPS sample demonstrates that a subset of Vermonters suffer from an inability to travel to destinations both within and outside of the state due to a lack of transportation options. Overall 22% of the sample reported some level of unmet travel need on an annual basis. Focusing on travel to destinations outside of Vermont, close to 3% of respondents (n=2,070) expressed that they needed to travel outside the state at least multiple times per month but were unable to do so because of a lack of transportation options. An additional 5% of respondents reported unmet long-distance travel need multiple times per year.

A logistic regression model of the likelihood of having unmet long-distance travel needs was created using a backwards stepwise method with employment, geographic, household, and personal predictor variables. The final model (n = 1,395) had a McFadden R^2 of 0.08 so was not strongly predictive of the likelihood of expressing unmet travel needs outside of Vermont. Several socio-demographic variables were statistically significant, however, including household size, number of children in household, household income, and the number of vehicles in the household. Households with incomes less than $50,000 were nearly twice as likely to have unmet long-distance travel needs as the $50,000 - $150,000 reference group (odds ratio = 1.91). Household size (odds ratio 1.26), belonging to the 18 – 34 year old age cohort (odds ratio 1.98), and strong support for passenger rail (odds ratio 2.61), were also strongly positively correlated with unmet long-distance travel needs. Unsurprisingly, as the number of household vehicles increased, the likelihood of reporting unmet travel needs decreased (odds ratio 0.74).

As suggested by the relatively weak model for unmet travel needs, the determinants of the need/desire for long-distance travel are complex and not easily captured in survey datasets. Moreover, non-travelers may be more likely to opt out of transportation surveys, further reducing the ability to understand who is experiencing unmet long-distance travel demand and the impact of these unmet needs' impacts on well-being. To explore this question in more depth, 24 women living in Chittenden County, Vermont were recruited to take part in semi-structured interviews on the relationship between long-distance travel and well-being. Interview data collection allows for follow-up and clarifying questions that are not possible with survey instruments. Prior to taking part in the interview process, participants completed the PiYL survey and, after the interview ended, completed the Satisfaction with Life Scale subjective well-being measure developed by psychologist Ed Diener (Appendix B), and were also asked which of the follow statements about travel was most true for them:

4. Increasing travel would increase my overall well-being.
5. Decreasing travel would increase my overall well-being.
6. Changing my level of travel would not affect my overall well-being.

The information collected through this three-stage process was used to explore the following questions:

1. How does the ability to participate in long-distance travel affect well-being?
2. What are the barriers to accessing long-distance travel?

A majority of the interviewees said that long-distance travel was very important or even essential to their sense of well-being. The interviews suggested five primary mechanisms by which long-distance travel impacted respondents' sense of well-being. These mechanisms were by facilitating face-to-face time with loved ones, providing a break from one's routine, providing access to adventures and new experiences, providing access to non-leisure or non-social needs (e.g. medical services), and (negatively) long-distance social status (e.g. travel envy). More frequent long-distance travel to visit friends and family and for vacation or leisure trips were both positively associated with well-being as measured on the Satisfaction with Life Scale.

Overall, the interviewees expressed a desire to increase their long-distance travel, though those with higher incomes had less unmet long-distance travel need than their lower income counterparts. Unsurprisingly, financial constraints were the most commonly cited barrier to long-distance travel. Other barriers included work obligations, household obligations (e.g. working around school schedules and family routines), and emotional stressors.

Measuring Long-Distance Travel Equity in the CHTS

Because access to long-distance destinations represents access to social capital and opportunities of value to well-being, considering differential trip making by sub-population group in large long-distance travel surveys is important. Although rarely considered, differences in the number of long-distance trips taken, the length of these trips, and the use of air travel among socio-demographic groups can be useful indicators of differential access to long-distance travel. The 2010-2012 CHTS is among the largest survey to include a long-distance travel log in recent years. Analysis of single-person households in this dataset indicates that income as well as gender, race, and age are important predictors of long-distance travel.

As noted in Section 0, the identity of the household member who completed the long-distance log was known for only 58% of the CHTS logs. In addition, the paper long-distance log, as presented in the CHTS final report Appendix (*46*), is ambiguously designed. It is unclear how many respondents recorded only their own trip and how many recorded the trips of all household members. In addition, the fields for which of the household members were traveling on each trip were not completed for most long-distance trips. Consequently, the analysis here is limited to single-person households since long-distance trips can be attributed with certainty to a specific individual. There are a total of 9,140 single-person households in the CHTS.

Both Bierce and Kurth (*47*) and Goulias et al. (*48*) observed that there was substantial variability in how individual respondents completed the long-distance log. Some respondents provided information for each trip leg within a complete long-distance tour, providing, e.g., separate records for the trip from home to a destination and the trip from that destination back home (in some cases even including layover airport locations), while others only provided a single record of a trip from home to a destination with no return trip. To minimize the impact of this inconsistency, Bierce and Kurth opted to focus only on outbound travel, eliminating return to home trips (*47*) while Goulias et al. undertook a more extensive effort to link individual trip legs

into tours to create a more consistent representation of long-distance travel across respondents (*48*). For the purposes of examining long-distance travel access equity, these inconsistencies are only important if they are correlated with the socio-demographic variables of interest. For this reason, the more simplistic approach used by Bierce and Kurth is replicated here, and trips flagged as return home in the CHTS trip table were eliminated from the analysis. The final dataset included 7,353 long-distances trips with non-home destinations.
The socio-demographic breakdown of this sample in terms of gender, income, race and age as well as several indicators of long-distance travel behavior are summarized in Table 5. The number of observations is different from the number of long-distance travelers and long-distance trips because many individuals did not make a long-distance trip in the 8-week period. For each demographic factor, Table 5 provides a count of the number of respondents in each category as well as the percentage of these respondents who reported at least one long-distance trip. The average number of trips taken by these long-distance travelers as well as their maximum distance from home across all reported trips, and the percent of long-distance travelers who utilized air travel on one or more trips are also shown.

The clearest relationship between long-distance travel and socio-demographics is apparent in Table 5 when considering respondents by income. All three measures of long-distance travel – number of trips, maximum distance from home, and air travel – increase nearly monotonically as income increases. When considering race, the percent of long-distance travelers and number of trips taken by these travelers is highest for White and Asian respondents. Note that racial demographics were originally recorded in five separate fields, race1 through race4 and a respondent-specified “other” category. For this analysis, respondents who specified more than one race or who selected “other” and specified multiracial are classified as multiracial. In some instances, respondents specified multiple races in the other category so some respondents in that category could be classified as multiracial.

Three basic regression models were created to consider the relationship between these demographic factors and long-distance travel behavior. This first model is a binary logistic regression model to estimate the likelihood that an individual reported at least one trip in the long-distance log.
Respondents who did not report any long-distance trips fall into one of three groups: non-travelers, lower-frequency long-distance travelers who did not take a long-distance trip during the recall period, and inaccurate reporters who made one or more long-distance trip during the recall period but failed to record it. Thus, as noted by Goulias et al. (*48*), it is most accurate to consider these respondents as either reporting or not reporting a long-distance trip since it is impossible to differentiate between respondents who did not travel from those that failed to report long-distance travel. For ease of reference, we simply refer to these non-reporters as non-travelers. The results of this binary logistic model are presented in Table 6. All of the demographic variables presented in Table 5 were significant predictors of whether or not an individual was a long-distance traveler. Respondents in all income categories below the $50,000 to $75,000 reference category were substantially less likely to be long-distance travelers, while those households earning more than the reference category were more likely to travel (though this relationship was not significant for the highest two income categories). In this model, men

were 16% less likely than women to report long-distance travel and American Indians/Alaska Natives, African Americans and respondents belong to other races were between 29% and 52% less likely to travel than their white counter parts.

Table 5. Demographic Breakdown and Long-Distance Travel Behaviors of Single Person Households

Demographic Variable		Total Count	Long-Distance Travelers[3]	Behavior of Long-Distance Travelers: Number of Trips	Behavior of Long-Distance Travelers: Max Distance From Home (miles)	Behavior of Long-Distance Travelers: Air Travelers
Gender n=9,123	Male	3,846	34.2%	2.22	660	18.4%
	Female	5,277	34.6%	2.42	604	20.5%
Income n=8,344	< $10k	742	19.4%	1.74	489	9.7%
	$10k - 25k	1,889	24.8%	1.99	494	10.7%
	$25k - 35k	1,016	30.8%	2.09	570	16.6%
	$35k - 50k	1,281	36.4%	2.20	508	15.0%
	$50k - 75k	1,549	40.5%	2.50	612	22.5%
	$75k - 100k	888	46.1%	2.75	663	22.5%
	$100k - 150k	675	47.1%	2.68	867	29.9%
	$150k - 200k	173	50.3%	2.75	598	29.9%
	$200k - 250k	59	50.8%	2.00	1,135	43.3%
	$250k or more	72	40.3%	2.45	1,847	41.4%
Race n=8,930	White	7,108	36.6%	2.36	618	19.8%
	Black/African American	517	20.7%	1.58	723	19.6%
	American Indian/Alaska Native	178	23.6%	2.26	334	16.7%
	Asian	288	37.2%	2.80	862	22.4%
	Hawaiian/Pacific Islander	14	7.1%	1.00	39	0.0%
	Multiracial	256	31.3%	2.38	520	17.5%
	Other	569	27.1%	2.08	637	15.6%
Age (years) n=9,140	18 to 29	569	34.8%	2.29	552	19.7%
	30 to 39	465	42.6%	2.72	703	27.8%
	40 to 49	903	37.8%	2.20	674	23.2%
	50 to 59	2,489	36.1%	2.48	661	18.4%
	60 to 69	2,840	35.4%	2.30	566	18.7%
	70 or older	1,874	27.2%	2.09	656	17.8%

[3] Made at least one long-distance trip in the 8-week period

The second model was a binary logistic model of the likelihood that those who were deemed long-distance travelers made a trip using air-travel. Non-travelers were excluded from this analysis. Race was not a significant predictor of air-travel among long-distance travelers, but gender, income and age were significant in the model. As with the model of long-distance travelers, the likelihood of traveling by air was higher for women than for men and increased with income. An alternative modeling approach for this outcome variable would be to use zero-inflated regression techniques that simultaneously estimate a binary travelers/non-traveler model and a model of the specific outcome variable. This approach is explored in Goulias et al. (*48*).

Table 6. Likelihood of Making a Long-Distance Trip and of Using Air for Long-Distance Travel

	Traveler (n=7,940)			Flyer (n=2,792)		
	Estimate	P	Odds Ratio	Estimate	P	Odds Ratio
Intercept	0.44			-0.34		
Male vs Female	-0.15	***	0.86	-0.27	***	0.76
Income < $10k vs $50k - 75k	-0.98	***	0.38	-1.00	***	0.37
Income $10k - 25k vs $50k - 75k	-0.67	***	0.51	-0.86	***	0.43
Income $25k - 35k vs $50k - 75k	-0.38	***	0.68	-0.40	**	0.67
Income $35k - 50k vs $50k - 75k	-0.18	**	0.84	-0.48	***	0.62
Income $75k - 100k vs $50k - 75k	0.22	**	1.25	-0.02		0.98
Income $100k - $150k vs $50k - 75k	0.21	**	1.23	0.32	**	1.38
Income $150k - 200k vs $50k - 75k	0.40	**	1.49	0.40		1.49
Income 200k - 250k vs 50k - 75k	0.47	*	1.60	1.18	***	3.26
Income 250k or more vs 50k - 75k	0.02		1.03	0.83	**	2.30
American Indian /Alaska Native vs White	-0.37	**	0.69			
Asian vs White	-0.22		0.80			
Black or African American vs White	-0.73	***	0.48			
Multiracial vs White	-0.26	*	0.77			
Hawaiian/Pacific Islander vs White	-1.91	*	0.15			
Other Race vs White	-0.35	***	0.71			
Age (years)	-0.01	***	0.99	-0.01	***	0.99
Nagelkerke R^2	0.07			0.06		

*** = P < 0.01, ** = P < 0.05, * P < 0.1

The final model was an ordinary least squares regression model of the maximum distance that a respondent traveled from home (Table 7). As with the model for use of air travel, this model was limited to the long-distance travelers in the dataset. The maximum distance was log transformed to improve model fit and the normality of model residuals and thus the dependent variable for this model is the log of the respondents' maximum distance from home. Once again, though the model's explanatory power is small (R^2 = 0.03), income and age were highly significant with higher levels of income associated with long-distance destinations that were farther from home and increasing age associated with decreasing distance from home. Neither race nor gender was a significant predictor of respondents' maximum distance from home.

Table 7. OLS Regression – Log of the Maximum Distance from Home (n=2,783)

	Int.	Income (In thousands of dollars vs. \$50 - 75k)									Age (yrs)
		<10	10-25	25-35	35-50	75-100	100-150	150-200	200-250	>=250	
Est.	5.59	-0.48	-0.41	-0.20	-0.30	-0.04	0.26	0.04	0.56	0.58	-0.005
P		<0.001	<.0001	0.05	0.00	0.70	0.01	0.82	0.04	0.04	0.027
R^2	0.03										

Consistent with previous analysis of the CHTS long-distance travel (*48*), all three models show a very strong, positive relationship between long-distance travel and income. Interestingly, while race is an important predictor of whether or not individuals are long-distance travelers, it was not a significant predictor of whether or not an individual traveled by air or their maximum distance from home. Likewise, gender was significant for travel status and air-travel status but not distance from home. Further exploration of the role of race and gender in access to long-distance travel is merited. Though none of the models were particularly strong, an unsurprising result given the simplicity of the models and the complexity of long-distance travel, they strongly reinforce existing concerns about a lack of equity in long-distance travel access.

Decision-Making for Personal Long-Distance Travel from the Vermonter Poll

Among long-distance travelers, trip decision-making, in terms of mode choice and destination selection, is likely to be influenced by socio-economic factors. The set of travel choices that is financially feasible for wealthier travelers is inherently larger than that of their less well-off counterparts. The Vermonter Poll captured the mode choice as well as information about the trip planning process for each respondent's most recent overnight, out-of-town trip for personal reasons. With respect to trip planning, respondents were asked whether they had selected their travel mode first and then their destination, selected their destination first and then their travel mode, or selected their destination first and only considered a single mode (referred to here as a joint mode/destination choice). Disparities in mode choice, especially the selection of air travel, and in trip planning processes may be indicative of inequitable access to long-distance destinations.

Since respondents' home locations were collected at the county level and trip destinations were captured using country, state, and place names, calculating the distance between home and destination location required that a latitude and longitude be attached to each location. Latitudes and longitudes were extracted using a Google geocoding plug-in for Microsoft Excel. Thereafter, the great circle distance between these locations was calculated in a Python script using the Haversine method. This script also extracted the on-road distance between these locations using the Google Maps API. These calculations were performed for all domestic trips that included a city/town destination name and for all international destinations that included a country entry. In total, trip distances were calculated for 498 respondents. The demographics of these respondents are summarized in Table 8 and their travel destinations are summarized in

Table 9. As might be expected, destinations were heavily skewed towards Vermont and the immediately adjacent states (218 of the 498 trips) but the remaining trips cover a wide range of other US and international destinations. The vast majority of trips were made by automobile or airplane, 363 and 115 respectively, with an additional 11 trips made by train or bus (the mode was not recorded for 9 trips).

Table 8. Demographics of Vermonter Poll Sample (N=498)

Categorical variables	Category	Count	Percent
Gender	Male	218	43.8%
	Female	262	52.6%
	Don't Know/Refused	18	3.6%
Income	$100,000 or more	127	25.5%
	$75,000-$100,000	73	14.7%
	$50,000-$$75,000	87	17.5%
	$25,000-$50,000	84	16.9%
	< $25,000	50	10.0%
	Missing	77	15.5%
Household Size	1	101	20.2%
	2	176	35.3%
	3	77	15.5%
	4	77	15.5%
	5	23	4.6%
	6	22	4.4%
	7	3	0.6%
	10	1	0.2%
	Missing/Invalid	18	3.2%
Home Ownership	Own	392	78.7%
	Rent	79	15.9%
	Other	8	1.6%
	Missing/Refused	19	3.8%
Education	Less than High School (no diploma)	14	2.8%
	High School graduate (incl. GED)	73	14.7%
	Some college (no degree)	93	18.7
	Associates/technical	40	8.0%
	Bachelor	142	28.5%
	Post graduate/professional	119	23.9%
	Missing/Refused	17	3.4%
Race	White or Caucasian	449	90.2%
	Black or African American	5	1.0%
	American Indian or Alaskan Native	5	1.0%
	Asian or Pacific Islander	4	0.8%
	Mixed	5	1.0%
	Missing/Refused	30	6.0%
Response Method	Landline	293	58.8%
	Cell Phone	205	41.2%
Continuous variables	**Mode**	**Maximum**	**Minimum**
Total Age (years)	59	92	19

Table 9. Long-distance Travel Destinations in the Vermonter Poll

Destination State	Number of Respondents	Number of Destinations Within State
Alaska	2	2
Arizona	2	2
California	15	7
Colorado	4	2
Connecticut	19	15
Washington DC	7	1
Delaware	1	1
Florida	40	22
Georgia	4	3
Hawaii	4	3
Iowa	1	1
Illinois	5	1
Kentucky	1	1
Louisiana	2	2
Massachusetts	61	21
Maryland	4	3
Maine	41	22
Michigan	3	2
Missouri	2	1
North Carolina	6	5
New Hampshire	46	29
New Jersey	11	10
Nevada	2	1
New York	68	24
Ohio	5	3
Oklahoma	1	1
Pennsylvania	15	8
Rhode Island	4	4
South Carolina	3	2
Tennessee	3	3
Texas	3	2
Utah	2	2
Virginia	5	4
Vermont	43	26
Washington	2	1
Wisconsin	3	3
West Virginia	3	3
Wyoming	3	3
Outside US	52	N/A

As shown in Table 10, most respondents, 379 out of 481, selected their destination first and only considered a single mode of travel. In many cases, this joint destination/mode decision may reflect that only a single mode was available or that alternative mode choices were too impractical to be realistically considered. For example, it is possible to travel by air from Burlington, VT to Boston, MA (a straight line distance of approximately 180 miles) but there are no direct flights between these destinations. Another 81 respondents selected their destination first and then considered multiple modes of travel while only 21 respondents selected their travel mode first and then their destination.

Table 10. Mode/Destination Selection Sequence by Destinations Considered

		Destination Selection		
		Exact destination was known and no other destinations were considered	Multiple destinations in the same region were considered	Multiple destinations in different states, regions or countries were considered
Mode/Destination Selection Sequence	Travel mode was decided first and destination was selected afterward.	17 (3.5%)	3 (0.6%)	1 (0.2%)
	Destination was selected first and then multiple travel modes were considered.	71 (14.8%)	4 (0.8%)	6 (1.2%)
	Destination was selected first and only one travel mode was considered or available.	330 (68.6%)	39 (8.1%)	10 (2.1%)

It is conceivable that the selection of the travel mode before selection of the destination is an indication of economically constrained travel. For example, someone might be decided to travel by bus or by car if flying was cost prohibitive and then choose a destination that was accessible by that mode. Similarly, an individual who wanted to fly somewhere but was cost constrained might choose their mode first and then look for low cost destinations. Alternatively, the choice of travel mode first might reflect the desire for a particular type of experience, e.g. a desire to take a road trip. Of the 21 respondents who selected their mode of travel first, 17 traveled by car, three by airplane and one by bus. Two automobile travelers and two air travelers considered multiple destinations while the remaining 17 individual only considered one destination. A chi-square test showed ($\chi^2(2)=4.74$, p = 0.09) some evidence of an association between household income and mode/destination selections sequence, with households with incomes below $75,000 being more likely to select their travel mode first than household with higher incomes. Homeownership, gender and the presence of children in the household did not show any evidence of a relationship to mode/destination selection sequences as summarized in Table 11. Relationships between demographics and trip decision-making processes could influence the structure of long-distance travel models.

Table 11. Mode/Destination Selection Sequence by Socio-Demographic Characteristics

Mode/Destination Selection Sequence	Income		Own Home		Gender		Children in HH	
	<75K	≥75K	Yes	No	Female	Male	Yes	No
Mode then Destination	74%	26%	71%	29%	66%	33%	41%	59%
Destination then Mode	46%	54%	85%	15%	56%	44%	30%	70%
Joint Destination/Mode	53%	47%	82%	18%	54%	46%	34%	66%
Chi-Square P-value	0.09		0.36[1]		0.51		0.54	

[1]Results for Fisher's Exact Test since expected values less than 5

A binary logistic regression of the likelihood of selecting air travel was created for domestic long distance trips (n = 446). Air travel was used in 19.7% of these trips. The coefficients and odds ratios for the final model are shown in Table 12. As would be expected, trip length was an extremely significant predictor of mode choice with the probability of traveling by air increasing with the distance from home. Again, as expected, lower levels of educational attainment were predictive of a lower likelihood of selecting air travel with those without a college degree only 30 – 40% as likely to travel by air as those with a post graduate degree. (Note that educational attainment and income were significantly associated with one another and that education was included in the final model because it had a lower rate of missing/refused values than the income variable.) Other demographic variables including, race, gender, the presence of children in the household, and household size were considered but were not individually significant and did not improve model performance.

Table 12. Likelihood of Selecting Air Travel

		β		Odds Ratio
	Intercept	-1.43		
Education	Less than High school vs Post graduate/professional	-1.33		0.27
	High School graduate vs Post graduate/professional	-1.21	**	0.30
	Some College vs Post graduate/professional	-0.91	*	0.40
	Associates/technical vs Post graduate/professional	-0.65		0.52
	Bachelor vs Post graduate/professional	-0.83	*	0.44
	Age (years)	-0.03	**	0.97
	Distance from Home (miles)	0.003	***	1.003
-2 Log Likelihood: 243.9				
Nagelkerke R^2: 0.65				

*P<0.1; **P<0.05; ***P<0.01

Looking both at the trip planning process and at mode selection provides, not unexpected, evidence that long-distance travel choices are influenced by socio-demographics. Higher income, more highly educated households have access to a broader set of long-distance travel choices. Moreover, it is likely that this analysis understates these effects for two reason. First, this analysis looked exclusively at individuals who had completed a long-distance trip. A totally of 61 households did not report a long-distance trip and these household were more than twice as likely (59% vs 27%) to report an income of less the $50,000 than those household that did report a long-distance trip. Second, the poll only asked respondents about their most recent trip. As summarized previously and discussed in (*49*), the frequency of trip-making decreases with trip length, and thus shorter long-distance trips are more likely to be reported when respondents are asked only about a single trip, even if their annual long-distance travel patterns includes much longer trips.

Conclusions

Cost effectively collecting high-quality, long-distance data collection remains a challenge. It is clear from research in this project and prior work by others that the most common current long-distance data collection methods are suboptimal. Distance thresholds are a poor method for defining long-distance travel on a national scale and there is no consensus on the appropriate recall period for retrospective long-distance travel surveys. Analysis of the LOST suggested that self-assessed travel frequency estimates can provide only a crude approximation of long-distance travel and that self-assessment is most effective for identifying non-travelers and very infrequent travelers.

On a promising note, the LSOT dataset also provides evidence that convenience samples may be suitable for studying the distribution of long-distance trip lengths and destination spatial distributions which could reduce the cost of obtaining these data. Assessing long-distance travel based on social network geography also showed some promise though modeling the relationship between social network extent and long-distance travel effectively will require larger and more comprehensive datasets than captured with the PiYL pilot survey.

Consistent with previous findings, there is ample evidence across multiple datasets in this project that lower-income individuals travel less and have more unmet long-distance travel needs. Given both the theoretical and empirical evidence that long-distance travel is correlated with individuals' own sense of well-being, the social inequitable access to long-distance travel cannot be ignored. This finding suggests generally that equity in long-distance access is a policy concern that must be considered in transportation planning. Moreover, as long-distance data collection and modeling methods are developed, including methods that rely on mobile device data, the ability to assess which groups do travel and do not travel engage in long-distance should be required in order to appropriately consider all aspects of sustainability including social welfare.

References

1. Dowds, Jonathan, Aultman-Hall, Lisa, and LaMondia, Jeffrey. Evaluating the Utilitiy of Self-Asssessed Travel Frequency Estimates For Measuring Overnight and Long-Distance Travel. in review.
2. Schäfer, A. W. Long-Term Trends in Domestic US Passenger Travel: The Past 110 Years and the next 90. *Transportation*, Vol. 44, No. 2, 2017, pp. 293–310. https://doi.org/10.1007/s11116-015-9638-6.
3. LaMondia, J. J., and C. R. Bhat. A Conceptual and Methodological Framework of Leisure Activity Loyalty Accommodating the Travel Context. *Transportation*, Vol. 39, No. 2, 2012, pp. 321–349.
4. Transportation Research Board. *Interregional Travel: A New Perspective for Policy Making*. Transportation Research Board, 2016.
5. Zhang, L., F. Southworth, C. Xiong, and A. Sonnenberg. Methodological Options and Data Sources for the Development of Long-Distance Passenger Travel Demand Models: A Comprehensive Review. *Transport Reviews*, Vol. 32, No. 4, 2012, pp. 399–433. https://doi.org/10.1080/01441647.2012.688174.
6. Larsen, J., J. Urry, and K. W. Axhausen. Networks and Tourism. *Annals of Tourism Research*, Vol. 34, No. 1, 2007, pp. 244–262. https://doi.org/10.1016/j.annals.2006.08.002.
7. Kuhnimhof, T., R. Collet, J. Armoogum, and J.-L. Madre. Generating Internationally Comparable Figures on Long-Distance Travel for Europe. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2105, No. 1, 2009, pp. 18–27. https://doi.org/10.3141/2105-03.
8. Calabrese, F., G. Di Lorenzo, L. Liu, and C. Ratti. Estimating Origin-Destination Flows Using Opportunistically Collected Mobile Phone Location Data from One Million Users in Boston Metropolitan Area. *IEEE Pervasive Computing*, Vol. 10, No. 4, 2011, pp. 36–44.
9. Järv, O., R. Ahas, and F. Witlox. Understanding Monthly Variability in Human Activity Spaces: A Twelve-Month Study Using Mobile Phone Call Detail Records. *Transportation Research Part C: Emerging Technologies*, Vol. 38, 2014, pp. 122–135. https://doi.org/10.1016/j.trc.2013.11.003.
10. Çolak, S., L. P. Alexander, B. G. Alvim, S. R. Mehndiratta, and M. C. González. Analyzing Cell Phone Location Data for Urban Travel: Current Methods, Limitations, and Opportunities. *Transportation research record: Journal of the transportation research board*, No. 2526, 2015, pp. 126–135.
11. Ritter, C., and E. Greene. Have Smartphone, Will Travel: Long Distance Travel Surveys with Smartphone GPS. Presented at the Transportation Research Board Annual Meeting, 2017.

12. Janzen, M., M. Vanhoof, Z. Smoreda, and K. W. Axhausen. Closer to the Total? Long-Distance Travel of French Mobile Phone Users. *Travel Behaviour and Society*, Vol. 11, 2018, pp. 31–42. https://doi.org/10.1016/j.tbs.2017.12.001.
13. Aultman-Hall, L., C. Harvey, J. Sullivan, and J. J. LaMondia. The Implications of Long-Distance Tour Attributes for National Travel Data Collection in the United States. *Transportation*, Vol. 45, No. 3, 2018, pp. 875–903. https://doi.org/10.1007/s11116-016-9754-y.
14. Harvey, Chester, and Aultman-Hall, Lisa. Comparing the Distance Distribution of Long-Distance Trips Between Online Surveys and Passive Mobile Devices. in review.
15. Aultman-Hall, Lisa, LaMondia, Jeffrey, H. Ullman, and M. Suender. Surveying Social Network Geography to Model Long-Distance Travel. *Transportation Research Procedia*, 2018.
16. Ullman, Hannah. *EXAMINGING DISPARITIES IN LONG-DISTANCE TRAVEL ACCESS*. The University of Vermont, Burlington, 2017.
17. Hu, P. S., and T. R. Reuscher. *Summary of Travel Trends: 2001 National Household Travel Survey*. Federal Highway Administration, 2004, p. 135 p.
18. FHWA. *Traffic Analysis Framework Part IIA -- Establishing Multimodal Interregional Passenger Travel Origin Destination Data*. Federal Highway Administration, Washington, D.C., 2013.
19. Outwater, M., M. Bradley, N. Ferdous, S. Trevino, H. Lin, Federal Highway Administration, and I. RSG. *Foundational Knowledge to Support a Long-Distance Passenger Travel Demand Modeling Framework: Implementation Report*. 2015, p. 92p.
20. Sharp, J., and E. Murakami. Travel Surveys: Methodological and Technology-Related Considerations. *Journal of transportation and statistics*, Vol. 8, No. 3, 2005, pp. 97–113.
21. Son, S., A. Khattak, X. Wang, P. Agnello, and J.-Y. Chen. Quantifying Key Errors in Household Travel Surveys: Comparison of Random-Digit-Dial Survey and Address-Based Survey. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2354, 2013, pp. 9–18.
22. Bricka, S., and E. Sabina. Evaluation of Key Design Elements of Long-Distance Survey of Front Range Travel Counts. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2285, 2012, pp. 110–118. https://doi.org/10.3141/2285-13.
23. Axhausen, K. W., M. Löchl, R. Schlich, T. Buhl, and P. Widmer. Fatigue in Long-Duration Travel Diaries. *Transportation*, Vol. 34, No. 2, 2007, pp. 143–160. https://doi.org/10.1007/s11116-006-9106-4.
24. Ortúzar, J. D. D., J. Armoogum, J. Madre, and F. Potier. Continuous Mobility Surveys: The State of Practice. *Transport Reviews*, Vol. 31, No. 3, 2011, pp. 293–312. https://doi.org/10.1080/01441647.2010.510224.

25. Bekhor, S., Y. Cohen, and C. Solomon. Evaluating Long-Distance Travel Patterns in Israel by Tracking Cellular Phone Positions. *Journal of Advanced Transportation*, Vol. 47, No. 4, 2013, pp. 435–446.

26. Rokib, S., A., M. A. Karim, T. Z. Qiu, and A. Kim. Origin-Destination Trip Estimation from Anonymous Cell Phone and Foursquare Data. 2015.

27. Bindra, S. Using Cellphone O-D Data for Regional Travel Model Validation. Presented at the TRB 95th Annual Meeting, 2016.

28. Wang, M.-H., S. D. Schrock, N. Vander Broek, T. Mulinazzi, Mid-America Transportation Center, and Research and Innovative Technology Administration. *The Use of Cell Phone Network Data in Traffic Data Collection and Long-Haul Truckshed (Geographic Extent) Tracking*. 2012, p. 37p.

29. Chen, C., J. Ma, Y. Susilo, Y. Liu, and M. Wang. The Promises of Big Data and Small Data for Travel Behavior (Aka Human Mobility) Analysis. *Transportation Research Part C: Emerging Technologies*, Vol. 68, 2016, pp. 285–299. https://doi.org/10.1016/j.trc.2016.04.005.

30. Linda Christensen. Consideration on a Method to Analyse Long Distance Travelling When Collecting Only Data about the Latest Journey. Presented at the 11th Internattion Conferences on Transport Survey Methods, Quebec, 2017.

31. Schoenduwe, R., M. G. Mueller, A. Peters, and M. Lanzendorf. Analysing Mobility Biographies with the Life Course Calendar: A Retrospective Survey Methodology for Longitudinal Data Collection. *Journal of Transport Geography*, Vol. 42, 2015, pp. 98–109. https://doi.org/10.1016/j.jtrangeo.2014.12.001.

32. Frändberg, L. International Mobility Biographies: A Means to Capture the Institutionalisation of Long-Distance Travel? *Current Issues in Tourism*, Vol. 9, No. 4–5, 2006, pp. 320–334. https://doi.org/10.2167/cit262.0.

33. Berliner, R., L. Aultman-Hall, and G. Circella. Exploring the Self-Reported Long-Distance Travel Frequency of Adult Californians. Presented at the The 97th Transporation Research Board Annual Meeting, Washington, D.C., 2018.

34. Aultman-Hall, L., C. Harvey, J. J. LaMondia, and C. Ritter. Design and Response Quality in a One-Year Longitudinal Survey of Overnight and Long-Distance Travel. *Transportation Research Procedia*, Vol. 11, 2015, pp. 136–153.

35. Kunzmann, M., and V. Daigler. *2010 - 2012 California Household Travel Survey Final Report*. 2013.

36. RSG. LRTP PUBLIC OPINION SURVEY. http://vtrans.vermont.gov/sites/aot/files/planning/documents/planning/VTrans%20LRTP%20Survey%20Report%20Final%202016.pdf.

37. Carrasco, J. A., B. Hogan, B. Wellman, and E. J. Miller. Collecting Social Network Data to Study Social Activity-Travel Behavior: An Egocentric Approach. *Environment and Planning B: Planning and Design*, Vol. 35, No. 6, 2008, pp. 961–980. https://doi.org/10.1068/b3317t.

38. Huang, G., and M. Tausig. Network Range in Personal Networks. *Social Networks*, Vol. 12, No. 3, 1990, pp. 261–268.

39. Killworth, P. D., E. C. Johnsen, H. R. Bernard, G. A. Shelley, and C. McCarty. Estimating the Size of Personal Networks. *Social Networks*, Vol. 12, No. 4, 1990, pp. 289–312.

40. Axhausen, K. W., and A. Frei. Size and Structure of Social Network Geographies. *Arbeitsbericht Verkehrs-und Raumplanung*, Vol. 444, 2007.

41. Schönfelder, S. Urban Rythms: Modelling the Rythms of Individual Travel Behaviour. Department Bau. *Umwelt und Geomatik, ETH, Zürich*, 2006.

42. Schönfelder, S., and K. W. Axhausen. Structure and Innovation of Human Activity Spaces. *Arbeitsberichte Verkehrs-und Raumplanung*, Vol. 258, 2004, pp. 1–40.

43. Schönfelder, S., and K. W. Axhausen. Activity Spaces: Measures of Social Exclusion? *Transport policy*, Vol. 10, No. 4, 2003, pp. 273–286.

44. Kowald, M., and K. W. Axhausen. The Structure and Spatial Spread of Egocentric Leisure Networks. No. 651, 2010, pp. 65–66.

45. Guidon, S., M. Wicki, K. W. Axhausen, and T. Bernauer. Investigating the Relationship between Individuals' Social Networks, Mobility Behaviour, and Social Capital-Survey Method and First Results. 2017.

46. Kunzmann, M., and V. Daigler. *2010 - 2012 CHTS Final Report Appendix*. 2013.

47. Kurth, D. The Use of Three Surveys for Long Distance Travel Estimates in California. 2014, p. 16.

48. Goulias, K., A. W. Davis, E. C. McBride, K. Janowicz, and R. Zhou. *Long Distance Travel in CHTS and Social Meida Augmentation*. University of California Center for Economic Competitiveness in Transportation, 2017.

49. Aultman-Hall, L., C. Harvey, J. Sullivan, and J. J. LaMondia. The Implications of Long-Distance Tour Attributes for National Travel Data Collection in the United States. *Transportation*, 2016. https://doi.org/10.1007/s11116-016-9754-y.

Appendix A – People in Your Life Survey

People and Travel Survey – Alabama Version A

Thank you for participating in the People and Travel Survey. Many things affect the places we go and what we do there. This survey is designed to help researchers understand how the people you know influence your travel choices to allow planners to better serve future travel needs.

The survey will take approximately 10 minutes to complete; it has 3 sections but most answers only require a check mark or a circle.

All responses that you provide will be confidential and anonymous. You may stop the survey or skip questions at any time. If you have any questions about this survey please contact Hannah Ullman or Lisa Aultman-Hall at the University of Vermont – (802) 656-1245.

Section 1 – People in your Life

The following questions describe people who may live in your city/town OR somewhere else in the world.

1. Think about a family member you don't live with. (optional nickname:______________)

□ No such person (go to #4 page 3)

a) When was your last IN-PERSON FACE-TO-FACE contact with this person?

□ within 1 MONTH □ within 1 YEAR □ more than 1 YEAR ago □ never

b) In the LAST MONTH, have you exchanged an email, phone call, text, video-chatted or similar with this person?

□ Yes □ No

c) How close is your relationship with this person? (circle a number)

10	9	8	7	6	5	4	3	2	1	0
Very Close				Somewhat close				Not close at all		

d) Where does this person live? City/town:________________________

State or Country: __________________

2. Think about another family member you don't live with. (optional nickname:_________)

□ No such person (go to #4 next page)

a) When was your last IN-PERSON FACE-TO-FACE contact with this person?

□ within 1 MONTH □ within 1 YEAR □ more than 1 YEAR ago □ never

b) In the LAST MONTH, have you exchanged an email, phone call, text, video-chatted or similar with this person?

□ Yes □ No

c) How close is your relationship with this person? (circle a number)

10	9	8	7	6	5	4	3	2	1	0
Very Close				Somewhat close				Not close at all		

d) Where does this person live? City/town:________________________

State or Country: __________________

3. Think about one last family member you don't live with. (optional nickname:___________)

□ No such person (go to #4 next page)

a) When was your last IN-PERSON FACE-TO-FACE contact with this person?

□ within 1 MONTH □ within 1 YEAR □ more than 1 YEAR ago □ never

b) In the LAST MONTH, have you exchanged an email, phone call, text, video-chatted or similar with this person?

□ Yes □ No

c) How close is your relationship with this person? (circle a number)

10	9	8	7	6	5	4	3	2	1	0
Very Close				Somewhat close				Not close at all		

d) Where does this person live? City/town:________________________

State or Country: __________________

4. **Think about a person you would go to for work or professional advice. (optional nickname:___)**

□ No such person (go to #5 below)

a) When was your last IN-PERSON FACE-TO-FACE contact with this person?

□ within 1 MONTH □ within 1 YEAR □ more than 1 YEAR ago □ never

b) In the LAST MONTH, have you exchanged an email, phone call, text, video-chatted or similar with this person?

□ Yes □ No

c) How close is your relationship with this person? (circle a number)

10	9	8	7	6	5	4	3	2	1	0
Very Close				Somewhat close				Not close at all		

d) Where does this person live? City/town:________________________

State or Country: __________________

5. **Think about a person you would go to for personal advice. (optional nickname:_______)**

□ No such person (go to #6 next page)

a) When was your last IN-PERSON FACE-TO-FACE contact with this person?

□ within 1 MONTH □ within 1 YEAR □ more than 1 YEAR ago □ never

b) In the LAST MONTH, have you exchanged an email, phone call, text, video-chatted or similar with this person?

□ Yes □ No

c) How close is your relationship with this person? (circle a number)

10	9	8	7	6	5	4	3	2	1	0
Very Close				Somewhat close				Not close at all		

d) Where does this person live? City/town:________________________
State or Country: __________________

6. Think about a good friend. (optional nickname:____________)

□ No such person (go to #7 below)

a) When was your last IN-PERSON FACE-TO-FACE contact with this person?

□ within 1 MONTH □ within 1 YEAR □ more than 1 YEAR ago □ never

b) In the LAST MONTH, have you exchanged an email, phone call, text, video-chatted or similar with this person?

□ Yes □ No

c) How close is your relationship with this person? (circle a number)

10	9	8	7	6	5	4	3	2	1	0
Very Close					Somewhat close					Not close at all

d) Where does this person live? City/town:________________________
State or Country: __________________

7. Think of a childhood friend. (optional nickname:___________)

No such person (go to #8 next page)

a) When was your last IN-PERSON FACE-TO-FACE contact with this person?

within 1 MONTH □ within 1 YEAR □ more than 1 YEAR ago □ never

b) In the LAST MONTH, have you exchanged an email, phone call, text, video-chatted or similar with this person?

□ Yes □ No

c) How close is your relationship with this person? (circle a number)

10	9	8	7	6	5	4	3	2	1	0
Very Close					Somewhat close					Not close at all

d) Where does this person live? City/town:________________________
State or Country: __________________

8. Think about a person you wish you could spend more time with. (optional nickname:________)

□ No such person (go to #9 below)

a) When was your last IN-PERSON FACE-TO-FACE contact with this person?

□ within 1 MONTH □ within 1 YEAR □ more than 1 YEAR ago □ never

b) In the LAST MONTH, have you exchanged an email, phone call, text, video-chatted or similar with this person?

□ Yes □ No

c) How close is your relationship with this person? (circle a number)

10	9	8	7	6	5	4	3	2	1	0
Very Close					Somewhat close					Not close at all

d) Where does this person live? City/town:________________________

State or Country: __________________

9. Think of a person for whom you feel an obligation to visit. (optional nickname:________)

□ No such person (go to #11 next page)

a) When was your last IN-PERSON FACE-TO-FACE contact with this person?

□ within 1 MONTH □ within 1 YEAR □ more than 1 YEAR ago □ never

b) In the LAST MONTH, have you exchanged an email, phone call, text, video-chatted or similar with this person?

□ Yes □ No

c) How close is your relationship with this person? (circle a number)

10	9	8	7	6	5	4	3	2	1	0
Very Close					Somewhat close					Not close at all

d) Where does this person live? City/town:________________________

State or Country: __________________

10. Think of another person for whom you feel an obligation to visit. (optional nickname:_______)

□ No such person (go to #11 below)

a) When was your last IN-PERSON FACE-TO-FACE contact with this person?

□ within 1 MONTH □ within 1 YEAR □ more than 1 YEAR ago □ never

b) In the LAST MONTH, have you exchanged an email, phone call, text, video-chatted or similar with this person?

□ Yes □ No

c) How close is your relationship with this person? (circle a number)

10	9	8	7	6	5	4	3	2	1	0
Very Close				Somewhat close					Not close at all	

d) Where does this person live? City/town:_______________________
State or Country: _________________

11. Think of a person you have communicated with in the LAST YEAR who lives in NEW YORK. (optional nickname:________________)

□ No such person (go to #12 next page)

a) When was your last IN-PERSON FACE-TO-FACE contact with this person?

□ within 1 MONTH □ within 1 YEAR □ more than 1 YEAR ago □ never

b) In the LAST MONTH, have you exchanged an email, phone call, text, video-chatted or similar with this person?

□ Yes □ No

c) How close is your relationship with this person? (circle a number)

10	9	8	7	6	5	4	3	2	1	0
Very Close				Somewhat close					Not close at all	

d) Where does this person live? City/town:_______________________
State or Country: _________________

12. Think of a person you have communicated with in the LAST YEAR who lives in CALIFORNIA. (optional nickname:________________)

□ No such person (go to #13 below)

a) When was your last IN-PERSON FACE-TO-FACE contact with this person?

□ within 1 MONTH □ within 1 YEAR □ more than 1 YEAR ago □ never

b) In the LAST MONTH, have you exchanged an email, phone call, text, video-chatted or similar with this person?

□ Yes □ No

c) How close is your relationship with this person? (circle a number)

10	9	8	7	6	5	4	3	2	1	0
Very Close					Somewhat close					Not close at all

d) Where does this person live? City/town:________________________
State or Country: __________________

13. Think of a person you have communicated with this past year who lives in EUROPE/ASIA. (optional nickname:________________)

□ No such person (go to Section 2 next page)

a) When was your last IN-PERSON FACE-TO-FACE contact with this person?

□ within 1 MONTH □ within 1 YEAR □ more than 1 YEAR ago □ never

b) In the LAST MONTH, have you exchanged an email, phone call, text, video-chatted or similar with this person?

□ Yes □ No

c) How close is your relationship with this person? (circle a number)

10	9	8	7	6	5	4	3	2	1	0
Very Close					Somewhat close					Not close at all

d) Where does this person live? City/town:________________________
State or Country: __________________

Section 2 –Travel Frequency

1. Check approximately how often you make a trip to a **destination more than 2 hours** from where you currently live…

	More than once per Month	Once per Month	Multiple Times per Year	Once per Year	Less than Once per Year	Never
To visit family or friends	☐	☐	☐	☐	☐	☐
For work	☐	☐	☐	☐	☐	☐
For personal business such as a medical appointment, banking or other services	☐	☐	☐	☐	☐	☐

2. Check approximately how often you make a trip …

	More than once per Month	Once per Month	Multiple Times per Year	Once per Year	Less than Once per Year	Never
For vacation or leisure	☐	☐	☐	☐	☐	☐
That includes air travel	☐	☐	☐	☐	☐	☐
With NO overnight stay that includes air travel	☐	☐	☐	☐	☐	☐
With NO overnight stay and includes 2 or more hours of driving EACH way	☐	☐	☐	☐	☐	☐
That includes a destination outside of North America	☐	☐	☐	☐	☐	☐

Section 3 – About you

1. **What is your Home Zip Code?**

2. **Does your workplace allow you to work from home and other locations? (check all that apply)**

☐ No
☐ Yes, and I often work from home or other locations
☐ Yes, and I occasionally work from home or other locations
☐ Yes, but I never work from other locations

3. **What is your gender?**

☐ Male ☐ Female ☐ Other

4. **What is your birth year?**

5. **What is your employment status? (check all that apply)**

☐ employed full-time
☐ employed part-time
☐ fulltime student
☐ retired
☐ not currently employed

6. **Circle the number of people including yourself who live in your residence.**

8+ 7 6 5 4 3 2 1

7. **With whom do you live? (check ALL that apply)**

☐ spouse or significant other (married or unmarried)
☐ roommate(s) (unrelated adult(s))
☐ child(ren) under 5 yrs old
☐ child(ren) between 5 and 12 yrs old
☐ child(ren) between 12 and 18 yrs old
☐ child(ren) over 18 yrs old
☐ other extended family
☐ other, please specify:

8. **What is your highest level of education?**

☐ high school or some high school
☐ some college
☐ bachelor's or associate's degree
☐ graduate or professional degree

9. **Do you have pets? (check ALL that apply)**

☐ large dog(s) ☐ small dog(s) ☐ cat(s) ☐ other pets ☐ no pets

10. How often do your pets affect your travel choices?

☐ always ☐ frequently ☐ infrequently ☐ never

11. Circle the number of registered motor vehicles (passenger cars, pick-up trucks, sport utility vehicles, vans/minivans, and motorcycles) you have in your household. (circle one)

8+ 7 6 5 4 3 2 1 0

12. Do you have a cell phone?

☐ Yes ☐ No

13. Where do you access the internet? (Check ALL that apply)

☐ at home
☐ at school
☐ at work
☐ at a public space (e.g. library)
☐ cell phone
☐ I do not have access to the Internet

14. Which of the following categories best describes your 2016 household income before taxes? Please include income from all sources for all persons with whom you share income.

☐ less than $15,000
☐ $15,000 to $24,999
☐ $25,000 to $49,999
☐ $50,000 to $99,999
☐ $100,000 to $149,999
☐ $150,000 to $199,999
☐ $200,000 or more
☐ prefer not to answer

Appendix B – Satisfaction with Life Scale

The Satisfaction with Life Scale score is calculated by adding up the rating assigned to each of the 5 statements and attempts to gauge one's overall satisfaction with life. The range of possible scores is from 5 to 35, with each statement being assigned a rating of 1 for strongly disagree to 7 for strongly agree.

The five statements presented are the following:

1. In most ways my life is close to my ideal.
2. The conditions of my life are excellent.
3. I am satisfied with my life.
4. So far I have gotten the important things I want in life.
5. If I could live my life over, I would change almost nothing.

The overall score is divided into 7 groups which represent different levels of life satisfaction:

1. 31-35 Extremely Satisfied
2. 26-30 Satisfied
3. 21-25 Slightly Satisfied
4. 20 Neutral
5. 15-19 Slightly Dissatisfied
6. 10-14 Dissatisfied
7. 5-9 Extremely Dissatisfied

The original development of the Satisfaction with Life Scale is described in: Diener, E. D., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of personality assessment, 49*(1), 71-75.