Quantifying Tourists' Long-Distance Destination Search Geographies: Balancing Length of Stay and the Time It Takes to Get There

by

Thomas Cole Edwards

A thesis submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Master of Civil Engineering

Auburn, Alabama
May 2, 2026

Keywords: Long Distance Travel, Time Away, Time Spent Travelling, Travel Distance

Approved by

Dr. Jeffrey LaMondia, Chair, Professor of Civil Engineering
Dr. Huaguo Zhou, Professor of Civil Engineering
Dr. Benjamin Bowers, Professor of Civil Engineering

Abstract

While research pertaining to leisure trips and their many components, such as travel times, time at destination and travel distances exist, the relationships between them have not been studied. The academic community has also explored leisure trips and the specific factors that influence destination choice. However, these factors have not been studied together, and their relationships have not been tracked, specifically the relationships between travel times and time at destination (Long-Distance Trip Efficiency Ratio, or ER) and travel distance and time at destination (Long-Distance Trip Coverage Ratio, or CR). Therefore, this work aims to understand the relationship between these variables, along with the specific factors that influence travel distances. To properly study these relationships, data collected from the 2024 National Household Travel Survey along with the home and destination characteristics were used. In total, 4,148 trips were analyzed for this study. This data was used to prove that a ratio between these two relationships can be calculated, and individuals' travel time-related decisions are most statistically influenced by the characteristics of the destination and trip purpose. This research indicates that the tourism and travel industry can use travel characteristics, household characteristics, and origin and destination zone characteristics to accurately anticipate where and how far individuals will travel for vacation .

Artificial Intelligence (AI) Use Disclosure

In the preparation of this thesis, the following Artificial Intelligence (AI) Tools were used: Microsoft Copilot. These tools were used primarily to create complex Excel formulas and to write a code to determine flight times for airport pairs. The author fully acknowledges full responsibility for the intellectual content of this work and has ensured that all AI-assisted sections have been reviewed and revised for accuracy and appropriate academic style. All AI-generated content was reviewed and validated for relevance, appropriateness and accuracy before incorporation into the final document to maintain scholarly integrity of this research.

Digital Accessibility Disclosure Statement

In the preparation of this thesis, the following digital accessibility tools were used to ensure this document complies with federal requirements: Microsoft Word, Microsoft PowerPoint, Microsoft Excel, ArcGIS Pro, IBM SPSS Statistics and PyCharm. The author acknowledges full responsibility for the intellectual content of this work and has made a good faith effort to comply with digital accessibility requirements in publishing, wherein the nature of the content does not significantly change in order to do so. Furthermore, all content has been reviewed and revised to meet these requirements prior to final publication.

Acknowledgements

I would like to express my gratitude to my advisor, Dr. Jeffrey LaMondia for his guidance and support throughout this research. I would also like to thank Dr. Huaguo Zhou and Dr. Benjamin Bowers for their support and feedback during this research as well. I would also like to thank the Federal Highway Administration for providing the results of the National Household Travel Survey along with additional information used in this study.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

GDP	Gross Domestic Product
ER	Long Distance Efficiency Ratio
CR	Long Distance Coverage Rate
SR	Long Distance Search Radius
NHTS	National Household Travel Survey
FHWA	Federal Highway Administration
ATL	Hartsfield-Jackson Atlanta International Airport
LAX	Los Angeles International Airport
SAIPE	Small Area Income and Poverty Index
NCEI	National Center For Environmental Impact
NOAA	National Oceanic and Atmospheric Administration

Chapter 1: Introduction

A long-distance vacation or leisure trip is when an individual or group of individuals travel at least 50 miles (ca. 80 km) from their origin for non-work-related purposes. It should be noted that these trips can be day trips if the trip is at least 50 miles. This type of travel is important for both local and federal businesses and agencies since long distance travel generates around $3 trillion in economic output and is 2.5% of the GDP of the United States (1). Studying long distance travel and decision-making processes of tourists is important because it allows the most popular destinations visited by Americans to forecast the volumes of visitors they will have at any given point. This is crucial knowledge for the destination locations because it allows them to plan for their economic and transportation needs in the future.

Despite the research that has been done pertaining to long distance travel, there are a few gaps remaining. First, research has studied the travel times, travel distances, and time spent at destination using varying modes of travel individually, but not together. Furthermore, research focuses its attention on how these four individual components are influenced by other variables such as trip characteristics, household characteristics, or the demographics of a traveler. For example, Research in this field focuses heavily on specific destination types (2, 3, 4, 5, 6, 7, 8), types of socioeconomic characteristics that influence a reasonable travel radius from the traveler's origin such as age and income (5, 6, 8, 11, 13, 14, 15, 16, 17) and the types of modes most commonly used for a long

distance trip (as well as how distance influences these decisions (6, 18, 19, 20, 21, 22, 23, 24). However, there has been little research conducted that compares time spent travelling to and time at destination and expected travel radius from the traveler's origin and time at destination. While research focuses on the individual variables, it does not quantify these relationships and how they affect traveler decisions. Understanding these relationships are crucial to quantify the relationship between these variable sets and the ER, CR, and SR of a trip. This work, as mentioned before, should be used by local and state governments, travel agencies and businesses to predict the average length of stay and the expected travel distance for the individuals who visit their communities based on the home location of the visitor.

Therefore, this work will quantify three values: (a) the ratio between travel time and time at destination (referred to as the Long Distance Trip Efficiency Ratio, or ER for the rest of this paper) and (b) the ratio between expected travel radius and time at destination (referred to as Long Distance Trip Coverage Rate, or CR for the rest of this paper) and (c) the search radius (or SR, in miles) for the trip. This work will also measure the correlation of other variables such as mode choice, home location characteristics, trip attributes, and socioeconomic, demographic and household characteristics. To determine these ratios and to compare these ratios to other personal variables and trip characteristics, data from the 2024 National Household Travel Survey (NHTS) will be used. Both ratios will be calculated for each trip, then they will then be compared to other variables from the NHTS to observe how both ratios change as these variables

change. Finally, a linear regression will be conducted to observe which variables influence these ratios the most.

Chapter 2: Literature Review

This literature review aimed to understand the relationship between expected travel times, travel distances and time away from home in total. To do this, many topics were analyzed to determine the relationships between travel characteristics and times and the characteristics of the individuals themselves. Specifically, this literature review analyzed the components (not the values themselves) of the ER, CR and SR of a trip and compared them to the attributes of the traveler such as their socioeconomic and demographic status, home/origin location qualities and the destinations themselves. While research has not done any work pertaining to the ER, CR and SR of a trip, the components of all three are heavily studied in the literature.

Searching for and selecting tourism destinations is a complex process of interest to researchers and practitioners in the transportation, tourism, marketing, and economic fields. To understand the decision-making process of travelers, two key components must be understood: destination searching and destination choice. Destination search is the process by which a traveler searches for a destination that meets a set of criteria. This criterion can be many things including the specific attributes of the destinations being considered, along with the distance and time needed to reach these destinations, along with the mode of transport used to get there. Destination choice is the process by which the destination is chosen for the long-distance trip due to the criteria the traveler has being met. Then, a mode is selected based on the distance and time availability of the traveler. These are two important distinctions that are crucial when studying the destination search and selection process of a long-distance traveler. This literature review summarizes how

the three interrelated components of tourism travel (i.e., distance, modes, time away) influence destination choices. Additionally, it describes how other factors (i.e. trip purpose, destination characteristics, travel party composition, and household socioeconomic characteristics) also play a role in destination searches.

This work will specifically study trips designated as “Long Distance Trips”. A long-distance travel study is research done on trips of at least 50 miles (ca. 80 km) distance in one direction. This is important to note because research on leisure trips of less than 50 miles were not considered for this review. Most of the literature that this review will focus on study leisure and tourism travel exclusively, with some other literature studying tourism and leisure travel along with other long distance travel purposes such as business and personal travel over 50 miles. It should also be noted that this study observed research from other continents such as Europe and Asia despite this study focusing exclusively on long distance travel data from the United States.

At its core, the most important factor influencing where tourists travel is the distance (and associated travel time) they are willing to complete to get there. Most literature pertaining to long distance travel studied are between 250 and 500 miles (2, 3, 4). Travel distances are a common research topic, and many factors that influence travel distance have been studied. However, the most common factor in influencing the distance travelled is the demographic characteristics of the traveler. Examples of these demographic factors commonly studied in the literature are age, income, gender, etc. According to the literature, the likelihood of taking a longer trip by distance decreases as age increases (5, 6, 7, 8). Furthermore, age also has an overall negative effect on taking

long distance trips in general (5, 9). Conversely, younger travelers do not view distance as negatively and are more likely to travel further to get to their destination (7). Therefore, older travelers travel less and travel further distances, while younger travelers are willing to travel further. Another demographic factor that influences trip distance is income, which the literature defines as a positive factor in increasing travel distance in long distance trips (2, 3, 5, 7, 8). Furthermore, the literature specifies that employed, married men with children and high incomes (around $100,000 or more) are the most likely to travel the furthest and the most often for their long-distance vacation (with or without the rest of their families) (2, 3). Also, these men are most likely to travel further if they have a higher education level (2, 3, 8). This has been attributed to the correlation between those who earn higher incomes and those who have a higher level of education. Together, these demographic factors play a role in the overall distance of a long-distance trip.

Of course, one must also consider how travel modes are inherently connected to how tourists consider distance and travel times. The mode choice of the traveler to their destination is based on a few factors including travel distance, travel time (or comparing the travel time of different modes) and travel costs. In general, research argues that as travel distance increases, so does the speed of the chosen mode (10, 11). For this research, speed is defined as the changing of modes as distance increases, as faster speeds indicate air travel and slower speeds indicate travel via personal vehicle (10, 11), since both will reach their destinations at the same time given their average speeds and distance of the trip falling within the threshold of an average trip taken by that mode (10). Despite this trend, airline passengers face more tradeoffs that personal vehicle travelers

do when taking a long-distance trip. These tradeoffs include increase in air fare (or air fare cost being significantly more expensive than personal vehicle utilization) the unpredictability of flight scheduling and delays, access to major airport hubs and the time lost if that traveler's flight involves a connection (10, 12, 13, 14, 15, 16, 17, 18, 19). These tradeoffs have led many travelers to choose travel via personal vehicle, since travelling in this way is more predictable and, in some cases, a cheaper option (13, 14). In general, however, travelling by air becomes more common if a trip exceeds 500 miles in distance (10, 11, 19, 20, 21). Conversely, travelling using a personal vehicle is more common if a trip is less than 500 to 600 miles. However, trips that are shorter than 250 miles in length are almost always taken via personal vehicle, while travel between 250 and 500 miles saw a slight increase in air travel usage (11, 19, 20, 21). Finally, research indicates that these factors are not isolated from each other and are all factors ultimately in mode choice decisions (22, 23). Furthermore, when a long-distance trip is planned, individuals in the literature reported having a more "feasible" mode that they used for their trip (23) However, this feasible mode is dependent on the "serviceability" or "access" to a closer airport, along with other factors mentioned previously. Thus, when a long-distance trip is being planned, modes are chosen based on distance, travel time and speed, costs, distance and serviceability of the nearest airport and predictability of flight times and delays.

Furthermore, and perhaps most relevant to this research, is the role that the total time away (also notated as the total travel time budget) has in destination choice. Travelers have limited time budgets for travel. The time they use to get to a destination

is time taken away from the activities they could do while at their destination. This time component is, not surprisingly, also connected to travel distances and mode choices. As mentioned before, one major deterrent from utilizing air travel is the lack of consistency in time variables. More specifically, many travelers are wary of flying because of the unpredictability of flight delays, fear of missing connecting flights if there are no other options for longer connections, and the lack of transparency of delays in advance (this time in advance is measured in weeks (10, 12, 13, 14, 15, 16, 17, 18, 19). Due to this uncertainty, many passengers will opt out of flying and utilize their personal vehicle instead. However, mode choice is less flexible as travel distance increases. As mentioned before, if a trip is over 500 to 600 miles (or around 1000 km) the likelihood of flying is much higher (11, 19, 20, 21, 23). This is primarily due to the travel time of flying being much lower than driving. In many cases, even if there are significant delays, travel time by air will still be lower travel time via ground transportation. Therefore, travelers will risk delays and uncertainty in total travel time if it means more time at their destination if they have access to a local airport with more direct flights and fewer connections.

Similarly, research has shown that destination choices are broadly tied to the purpose of the tourism trip. These purposes can include visiting friends and relatives, outdoor activities, shopping, relaxation, exploration, and even personal business. As mentioned previously, these factors are not the only factors that influence destination search but may lay a larger role in the destination selection for certain travelers. Other factors such as mode choice and distance have varying effects on the destination types. For example, the literature discusses how variance in trip purpose affects the mode a

traveler will choose. According to the literature, individuals travelling on business are more likely to fly to their destinations, while leisure travelers are more likely to drive to their destinations (24, 25, 26, 27). This can be attributed to several factors, but the most prominent in the literature are travel costs (or willingness of a company versus an individual to pay for air fare) and time budgets (or the value of time). Another characteristic of long-distance travel is distance. According to the literature, trips classified as leisure trips have a higher likelihood of being longer (in time spent at the destination) and in total distance, while work trips are typically shorter in length of stay and travel distance (28, 29, 30, 31). This is due to the varying time constraints and desire to be at the destination, since work trips rarely consist of leisure activities. Additionally, the literature observes that destination choice is more likely to change before the distance away from home does (32). This also means that, if a traveler is choosing between several destinations that are similar distances, they are more likely to choose their destination based on criterion such as trip purpose rather than trip distance (32).

More specifically, tourists evaluate and select destinations based on the characteristics of those destinations as they relate to their trip purpose. As mentioned, a major component of destination selection are the specific attributes of the destination. Specifically, this component consists of the activities, events, social opportunities, and opportunities for rest and relaxation, among other reasons that are at the location. According to the literature, the most popular vacation type for long distance travelers is to tropical locations (i.e. beach or locations with warmer weather) (2, 9, 33, 34, 35, 36, 37, 38). However, other types of destinations such as large cities, heritage and cultural

destinations, non-tropical coastal areas, locations with lots of green spaces and areas where outdoor sports such as hunting, fishing and hiking can take place (37, 39). Travelers will also select destinations that are like their homes (40). This is because these environments can provide a sense of familiarity and comfort to the traveler without the presence of everyday tasks such as work or school. Finally, other research has been done regarding the presence of novelties (or activities), and this research concludes that the most visited types of destinations are those that have a variety and abundance of activities that are all close together (41, 42).

Travel party composition also influences where tourists are most likely to visit, the modes they choose, and total time away. In general, most literature that explores average long distance trip length research does not consider the travel time. This literature also concludes that these trip lengths are around 5 to 12 days (44, 45, 46, 47, 48, 49, 50, 51, 52). Other literature argues that, while these trends are true, travelers are more likely to travel for weeklong (or 7 day) intervals, meaning that trip lengths of 7, 14, and 21 days are also common (49). This is an important trend to note because some length-of-stay studies observe how travel party size influences trip length. These studies specify that trip length increases as party size increases (44, 45, 47, 49, 50, 51, 54, 55). This has been linked to reasons such as the party wanting to see more places and do more activities while on the trip, thus making it longer. The literature further specifies how certain types of travel parties choose their length of stay. The literature indicates that families and large groups of friends are most likely to travel the longest and are also more likely to fly (52, 53, 54). This is primarily due to the structure that flying provides, and

many travelers feel that this makes travelling easier on themselves and the travel group (53). However, families and larger groups are also more sensitive to the changes that air travel provides, such as delays, and are also more likely to change modes if needed (54). Additionally, smaller travel groups (or solo travelers) have been linked to more trips, but the literature classifies these primarily as business travelers, which are not considered in this work. Therefore, families and larger parties are most likely to fly and stay longer at their destination but are more vulnerable to changes in travel commute itineraries.

Finally, tourism and travel researchers have long shown the importance of individual preferences on tourism decisions, including destinations, modes, and total time away. The primary variables that the literature identifies as significant in influencing trip length and mode are marital and family status, income and age. Research indicates that the most likely individuals to travel for a long-distance leisure trip are married men with multiple children (2, 55, 56). Additionally, these men tend to have higher incomes. However, these studies are the only studies that determined the travel patterns of men and women, as other works do not explicitly state the travel characteristics of men and women. Furthermore, the literature argues that higher income earners are more likely to take more long-distance trips. However, this does not mean that they stay longer or travel further, according to the literature (2, 55, 56). These trends, however, are not always true for older travelers. Research concludes that older travelers spend more time on vacation but tend to travel less and will travel shorter distances to get to their destination (2, 5, 6, 7, 8, 9, 56, 57), as mentioned previously. There are, however, a few gaps in the effects demographic characteristics have on long distance travel decision-making. First, there is little literature that aims to quantify long distance travel trends based on home location (or census region).

More research needs to be done in this field to accurately quantify the effects of the home state and home region of the traveler(s). Additionally, race and ethnicity are not deeply studied in the research. While both and variables included in most studies, like gender, few, if any conclusions about their impact on long distance travel decisions are made.

Chapter 3: Hägerstrand Space Time Prisms and Long-Distance Travel

To understand the movements of a traveler throughout space and time, it is important to understand Hägerstrand Space Time Prisms. A Hägerstrand Space Time Prism is a measure of an object's location in space (and distance away from its origin) at certain points in time (60, 61, 62, 63). This method of study has been used in many fields and is commonly used in the field of transport geography. the intended use of these prisms by Dr. Torsten Hägerstrand was to understand the basic movements and time uses of individuals in European city centers. Furthermore, space-time prisms were invented to understand how factors such as land use and human factors affect the everyday movement of an individual. While this work did not measure the movements of individuals in their everyday lives, it will be used to model a traveler's time uses in the travel times and corresponding distances to their vacation destinations and times at these vacation destinations.

When describing how each space time prism, three relationships can be quantified. The first is the actual trip that the traveler took. This trip can be taken using either a personal vehicle or air travel. The next two relationships that space time cubes can model are the maximum travel time and distance that the same individual is willing to endure using both air and ground transportation. These two prisms typically have larger travel distances and longer travel times but have the exact same time at destination as the actual trip. As illustrated in Figure 1, these relationships can be modelled in two- or three-dimensional space. First, each trip has two components. They are (a) travel distance and (b) time spent away from destination or travelling.

These positions are compared to the time spent away from home to observe how far away the traveler is away from home.

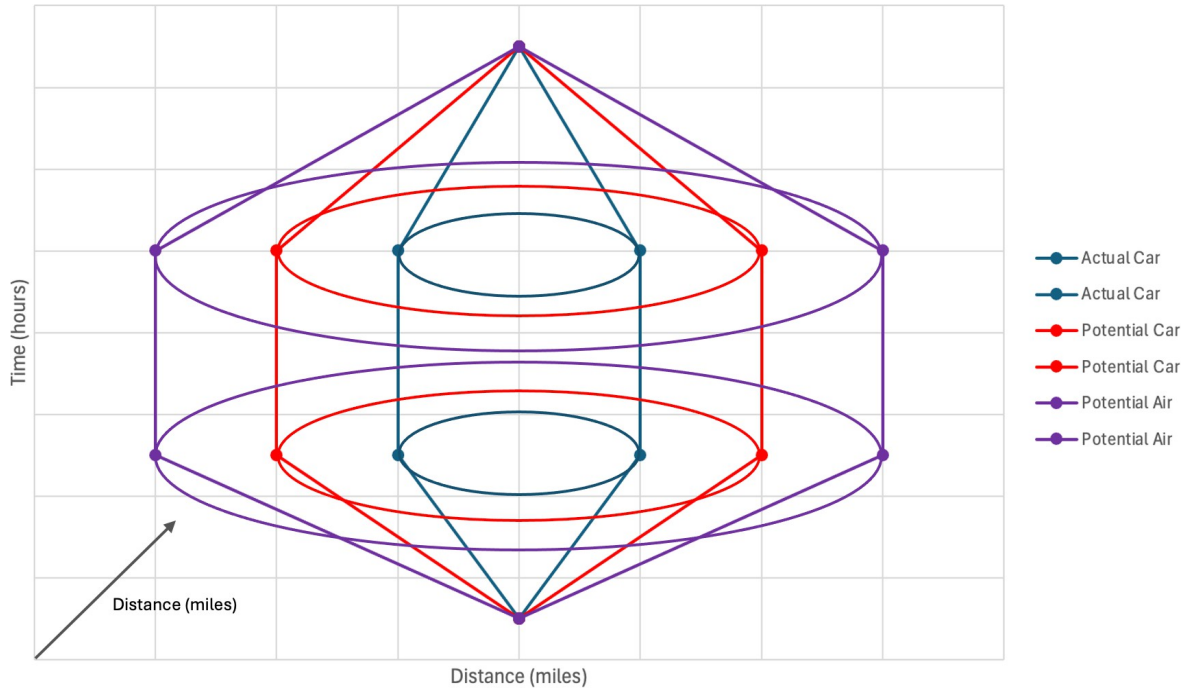


Figure 1, Example Hägerstrand Space-Time Prisms

These three prisms model the relationships observed in this study. Those relationships are the Long-Distance Efficiency Ratio (ER), Coverage Rate (CR), and Search Radius (SR). The first of three major components is the Efficiency Ratio of a traveler. This component measures how the travel time (or time spent travelling to and from the destination) and the total time spent away (or hours spent away from home) increase and decrease together. This component of space-time prisms is important because it is a measure of how travelers utilize their time when they travel. In terms of quantifying the Efficiency Ratio of a trip, this value will be between 0 and 1 for any given trip. This range of values can be due to travel times being smaller than the total time spent

away, since travel times cannot be larger than total time spent away. Additionally, an Efficiency Ratio of 1 is possible, but very unlikely, since most trips will not have the same travel time and time spent away from home. However, this is entirely possible.

The second and third relationship measured in space time prisms are the Search Radius and Coverage Rate of a trip. The search radius of a trip can be defined simply as the maximum distance the traveler was from the origin, or how far the destination was from home. This value is important because it is an easy way of identifying how far a traveler is willing to travel from home. Unlike the efficiency ratio, however, the search radius is only influenced by distance and is not influenced by length of time spent away from home. The third and final component of a space time prism are the coverage rates of a traveler. This value compares the Search Radius of a trip to the total time spent away from home in days. Coverage rates are intended to achieve a similar result as an efficiency ratio. This is done by illustrating how the search radius and the total time spent away from home change together. This value is not bounded by a set of values like the Efficiency Ratios are, and can exceed 1, but are always larger than 0.

To calculate each value, three different processes were used. In this model, the Efficiency Ratio (ER) can be calculated by the time spent travelling to the destination (or the sloped parts of the diagram) and the time spent away from home (or the vertical lines representing the time at the destination). This process is outlined by Formula 1.

$$ER = \frac{t_{depart} + t_{return}}{d_{hours}} \quad (1)$$

Where:

ER = Efficiency Ratio

t = Travel Time to and From Destination (hours)

d = Total Time Spent Away From Home (hours)

An ER ratio does not consider distance as a factor in, and only considers travel time (in hours), along with time at destination (in days). Along with modeling this relationship, ER also aims to quantify the maximum travel time using both air travel and personal vehicle given the length of stay. Specifically, ER aims to quantify the maximum travel time for both modes for specific stay intervals. These stay intervals include trip lengths such as weekends (1-3 days), weeklong trips, or trips longer than a week. For this study, every trip studied will have their own specific ER that will be used to determine the average ER for travelers who stay for certain lengths of time and use certain modes. The efficiency ratio of a trip is calculated using the travel times (for both the departing and returning trips) divided by the total time spent away from home. This time combines the total time spent travelling and time at destination, which is derived from the NHTS's "Total Number of Nights Away" variable. The variable in the NHTS was collected in nights, but this variable was calculated by counting the number of hours away if one night away was 24 hours. Additionally, the travel time was calculated differently for each mode. For car trips, the travel time was computed using an assumed speed of 65 mph and the calculated distance (which is discussed in the next section). For airplane trips, the travel time was calculated

using the time spent travelling on the flight and time spent commuting between airports (also using an assumed speed limit of 65 mph).

The second relationship this work observes is CR, or the Coverage Rate. A CR works similarly to ER, however, travel times are not considered for CR calculations. Instead, only travel distances (in miles), are considered. These travel distances, like travel times for ER calculations, were used to determine the maximum distance a traveler is willing to go for a trip with a certain trip length. CR studies used the same length of stay ranges (weekend, week, etc.) as ER. This relationship can be seen in Formula 2.

$$CR = \frac{SR}{d_{days}} \quad (2)$$

Where:

CR = Coverage Rate

SR = Search/Travel Radius (miles)

d = Time at Destination (days)

This relationship contains two primary components: the Search Radius (SR) and the time at destination in days. The time at destination variable is calculated in the same manner as the hours spent from home in the ER calculation. However, this variable observes days away as opposed to hours away. The variable was calculated by adding one day to the number of nights away, since there are always one more day in a trip than nights spent away (or days spent travelling, but not necessarily gone for the full day).

A final component of this analysis is the Search Radius of the traveler. This variable

is one of two components in the coverage rate formula and merely considers how far a traveler is willing to go. This value is independent of any travel time or time at destination variables and is used as both a calculation component in this work along with a third Hägerstrand characteristic that is observed in this study.

Chapter 4: Data

Data for this study was gathered from several sources, including the 2024 National Household Travel Survey (NHTS), County to County distance data, total time spent on flight (if the chosen mode for the trip was an airplane) and the individual attributes of the origin and destination counties themselves. Variables pertaining to travel distance/search radius and travel time were used to calculate ER, CR, and SR, while other attribute variables were used to measure their significance in determining ER, CR, and SR.

4.1 National Household Travel Survey

The National Household Travel Survey (NHTS) (64) was a survey collected by the Federal Highway Administration (FHWA) in 2024 to assist transportation planners and policymakers in planning for travel patterns in the United States. The survey aims to study these travel patterns by understanding the trip characteristics of the survey respondents. While this study contains sections that pertain to the vehicle and trip (non-long distance) characteristics, the only three sections from the study used were the household, person, and long-distance sections of the data. This information was gathered via digital survey information pertaining to demographics, socioeconomic characteristics, and characteristics of the long-distance trip in question. The NHTS provides many variables pertaining to the respondent's trip, including mode choice and time at destination. In total, 16,997 responses were recorded for the long-distance data file. However, as described later, this was not the final sample size, which would be 4,148. A

comprehensive description of how this number of entries was achieved will be discussed later.

4.2 County to County Distance Data

Unfortunately, the NHTS does not provide variables pertaining to travel time. This variable is crucial for determining the Efficiency Ratio, or ER, which compares the travel time of the respondent to the respondent's time at destination. To calculate these travel times, the mode of the traveler needed to be determined. This information is, fortunately, provided in the NHTS. The NHTS also provides the origin county and destination county of every long-distance trip conducted in the United States. Since the origins, destinations and mode choice of the traveler are known, a reasonable travel distance and time can be calculated between all county pairs. To calculate the travel distance and time for the trips via a personal vehicle, the distance between the origin and destination county centroids were calculated. The US Census Bureau provides a dataset of all counties in the United States (65). However, the dataset does not provide the specific coordinates of the center of each county, which were used for this study. Therefore, the centroid of each county needed to be found, along with the coordinates of each county. Using this information, the distance between the origin and destination county could be calculated. It should be noted that counties in Alaska, Hawai'i, and all territories of the United States were not used for this study, since a traveler cannot choose to drive to these places. Also, even though they are not considered "counties", the parishes of Louisiana and the independent cities of Baltimore, Maryland; St. Louis, Missouri; Carson City, Nevada; Washington, DC; and the independent cities of Virginia were also classified as counties for this work. While these are not considered counties by the US Census Bureau, they

were considered counties for this study. This means that a total of 3,108 counties were used for this analysis, and 9,659,664 county-to-county distance pairs. This distance calculation was done using the Haversine formula. The Haversine formula is a formula that uses the latitude and longitude of two points and calculates the distance between them. This formula can be found below.

$$h = 2r \sin^{-1}(\sqrt{\sin^2(\frac{\varphi_2-\varphi_1}{2}) + \cos(\varphi_1)\cos(\varphi_2)\sin^2(\frac{\lambda_2-\lambda_1}{2})} \quad (3)$$

Where:

r = Radius of the Earth in Miles (3958.8 miles)

ϕ = Latitude in Radians

λ = Longitude in Radians

The distances between every county in the United States (aside from the counties of Hawai'i and Census Boroughs of Alaska) were calculated. Then, all pairs of county distances were assigned to the trip in the NHTS that had the same origin counties and destinations. This would be used to estimate the travel time and travel distance via personal vehicle.

This dataset was calculated entirely for this project and previously did not exist. Therefore, no cleaning of the data was necessary since the distance between county pairs and centroid of each county were the only data that were used for this section of data collection. In short, this dataset included the distances between all 3,108 out of 3,143 counties in the continental United States and were later paired with the origin and

destination counties of each trip taken by personal vehicle in the NHTS.

4.3 Flight Time Data

Another mode that was used in the NHTS was air travel. Like the trips recorded in the NHTS that used a personal vehicle, trips that utilized air travel were assigned an origin and destination county. However, the distance calculations used for origin and destination counties of personal vehicle trips were not used to calculate travel times for air travel. To calculate the travel time for air travel (and later ground transportation from home location to origin airport and from destination airport to the final travel location), two separate calculations were conducted. The first two legs of each trip were calculated first. These two legs of the trip are distances to and from each airport. To calculate these distances, each origin and destination county needed an assigned airport. This was done because the NHTS does not provide information pertaining to the airports the traveler used. Therefore, this study used the closest airports to each origin and destination. However, only a certain number of airports were used. For this study, the only airports that were used were airports that had at least 500,000 enplaned passengers in 2023. This information was provided by the US Bureau of Transportation Statistics (66). This was done since most air passengers are likely to choose a larger airport to fly out of and in to, according to the literature, because they offer more flight options, more direct flights, and lower costs. The list provided by the Bureau of Transportation Statistics provides the top 200 airports in the United States in 2023. In total, 132 airports met these criteria (serving 500,000 passengers or more in 2023). However, the airport also had to be in the continental United States, eliminating airports in Alaska, Hawai'i, Puerto Rico, Guam, and the United States Virgin Islands (who all had at least one airport within this passenger

volume). These airports were eliminated for the same reasons that personal vehicle trips were not studied. Therefore, the final total number of airports used for this study was 122 airports. Those airports can be found below in Table 1. Additionally, a map of these airports with their respective airport codes can be found in Figure 2.

Table 1: List of Airports Used in This Study Based on Bureau of Transportation Statistics

Airport	Airport Code	2023 Enplaned Passengers
Atlanta, GA: Hartsfield-Jackson Atlanta International	ATL	50,925,000
Dallas/Fort Worth, TX: Dallas/Fort Worth International	DFW	39,226,000
Denver, CO: Denver International	DEN	37,840,000
Los Angeles, CA: Los Angeles International	LAX	36,611,000
Chicago, IL: Chicago O'Hare International	ORD	35,828,000
New York, NY: John F. Kennedy International	JFK	30,512,000
Orlando, FL: Orlando International	MCO	28,019,000
Las Vegas, NV: Harry Reid International	LAS	27,642,000
Charlotte, NC: Charlotte Douglas International	CLT	25,881,000
Seattle, WA: Seattle/Tacoma International	SEA	24,551,000
Miami, FL: Miami International	MIA	24,544,000
Newark, NJ: Newark Liberty International	EWR	24,476,000
San Francisco, CA: San Francisco International	SFO	24,174,000
Phoenix, AZ: Phoenix Sky Harbor International	PHX	23,814,000
Houston, TX: George Bush Intercontinental/Houston	IAH	22,210,000
Boston, MA: Logan International	BOS	19,950,000
Fort Lauderdale, FL: Fort Lauderdale-Hollywood International	FLL	17,037,000
Minneapolis, MN: Minneapolis-St Paul International	MSP	16,985,000
New York, NY: LaGuardia	LGA	16,172,000
Detroit, MI: Detroit Metro Wayne County	DTW	15,350,000
Philadelphia, PA: Philadelphia International	PHL	13,629,000
Salt Lake City, UT: Salt Lake City International	SLC	12,892,000
Baltimore, MD: Baltimore/Washington International Thurgood Marshall	BWI	12,765,000
Washington, DC: Ronald Reagan Washington National	DCA	12,347,000
San Diego, CA: San Diego International	SAN	12,156,000
Washington, DC: Washington Dulles International	IAD	12,049,000
Tampa, FL: Tampa International	TPA	11,656,000

Table 1 Continued: List of Airports Used in This Study Based on Bureau of Transportation Statistics

Airport	Airport Code	2023 Enplaned Passengers
Nashville, TN: Nashville International	BNA	11,195,000
Austin, TX: Austin - Bergstrom International	AUS	10,807,000
Chicago, IL: Chicago Midway International	MDW	10,646,000
Dallas, TX: Dallas Love Field	DAL	8,454,000
Portland, OR: Portland International	PDX	8,116,000
St. Louis, MO: St. Louis Lambert International	STL	7,297,000
Raleigh/Durham, NC: Raleigh-Durham International	RDU	7,101,000
Houston, TX: William P Hobby	HOU	6,763,000
Sacramento, CA: Sacramento International	SMF	6,364,000
New Orleans, LA: Louis Armstrong New Orleans International	MSY	6,295,000
San Jose, CA: Norman Y. Mineta San Jose International	SJC	5,944,000
Santa Ana, CA: John Wayne Airport-Orange County	SNA	5,659,000
Kansas City, MO: Kansas City International	MCI	5,639,000
Oakland, CA: Metro Oakland International	OAK	5,464,000
San Antonio, TX: San Antonio International	SAT	5,305,000
Fort Myers, FL: Southwest Florida International	RSW	4,955,000
Cleveland, OH: Cleveland-Hopkins International	CLE	4,791,000
Indianapolis, IN: Indianapolis International	IND	4,764,000
Pittsburgh, PA: Pittsburgh International	PIT	4,461,000
Cincinnati, OH: Cincinnati/Northern Kentucky International	CVG	4,257,000
Columbus, OH: John Glenn Columbus International	CMH	4,075,000
West Palm Beach/Palm Beach, FL: Palm Beach International	PBI	3,805,000
Jacksonville, FL: Jacksonville International	JAX	3,609,000
Ontario, CA: Ontario International	ONT	3,175,000
Hartford, CT: Bradley International	BDL	3,116,000
Charleston, SC: Charleston AFB/International	CHS	3,029,000
Burbank, CA: Bob Hope	BUR	3,010,000
Milwaukee, WI: General Mitchell International	MKE	2,920,000

Table 1 Continued: List of Airports Used in This Study

Airport	Airport Code	2023 Enplaned Passengers
Albuquerque, NM: Albuquerque International Sunport	ABQ	2,602,000
Omaha, NE: Eppley Airfield	OMA	2,458,000
Richmond, VA: Richmond International	RIC	2,390,000
Memphis, TN: Memphis International	MEM	2,385,000
Boise, ID: Boise Air Terminal	BOI	2,366,000
Norfolk, VA: Norfolk International	ORF	2,305,000
Buffalo, NY: Buffalo Niagara International	BUF	2,276,000
Louisville, KY: Louisville Muhammad Ali International	SDF	2,259,000
Reno, NV: Reno/Tahoe International	RNO	2,215,000
Sarasota/Bradenton, FL: Sarasota/Bradenton International	SRQ	2,142,000
Oklahoma City, OK: Will Rogers World	OKC	2,126,000
Spokane, WA: Spokane International	GEG	2,000,000
El Paso, TX: El Paso International	ELP	1,975,000
Tucson, AZ: Tucson International	TUS	1,907,000
Savannah, GA: Savannah/Hilton Head International	SAV	1,904,000
Grand Rapids, MI: Gerald R. Ford International	GRR	1,895,000
Long Beach, CA: Long Beach Airport	LGB	1,832,000
Providence, RI: Rhode Island TF Green International	PVD	1,726,000
Myrtle Beach, SC: Myrtle Beach International	MYR	1,654,000
Palm Springs, CA: Palm Springs International	PSP	1,623,000
Tulsa, OK: Tulsa International	TUL	1,568,000
Des Moines, IA: Des Moines International	DSM	1,512,000
Birmingham, AL: Birmingham-Shuttlesworth International	BHM	1,491,000
Sanford, FL: Orlando Sanford International	SFB	1,435,000
Syracuse, NY: Syracuse Hancock International	SYR	1,408,000
Knoxville, TN: McGhee Tyson	TYS	1,375,000
Albany, NY: Albany International	ALB	1,373,000

Table 1 Continued: List of Airports Used in This Study

Airport	Airport Code	2023 Enplaned Passengers
Pensacola, FL: Pensacola International	PNS	1,365,000
Rochester, NY: Frederick Douglass Greater Rochester International	ROC	1,324,000
Greer, SC: Greenville-Spartanburg International	GSP	1,252,000
Bozeman, MT: Bozeman Yellowstone International	BZN	1,227,000
St. Petersburg, FL: St. Pete Clearwater International	PIE	1,227,000
Fresno, CA: Fresno Yosemite International	FAT	1,198,000
Colorado Springs, CO: City of Colorado Springs Municipal	COS	1,182,000
Asheville, NC: Asheville Regional	AVL	1,124,000
White Plains, NY: Westchester County	HPN	1,123,000
Valparaiso, FL: Eglin AFB Destin Fort Walton Beach	VPS	1,116,000
Portland, ME: Portland International Jetport	PWM	1,114,000
Little Rock, AR: Bill and Hillary Clinton Nat Adams Field	LIT	1,094,000
Madison, WI: Dane County Regional-Truax Field	MSN	1,029,000
Fayetteville, AR: Northwest Arkansas National	XNA	968,000
Punta Gorda, FL: Punta Gorda Airport	PGD	949,000
Phoenix, AZ: Phoenix - Mesa Gateway	AZA	937,000
Greensboro/High Point, NC: Piedmont Triad International	GSO	896,000
Wichita, KS: Wichita Dwight D. Eisenhower National	ICT	840,000
Eugene, OR: Mahlon Sweet Field	EUG	837,000
Panama City, FL: Northwest Florida Beaches International	ECP	816,000
Huntsville, AL: Huntsville International-Carl T Jones Field	HSV	723,000

Table 1 Continued: List of Airports Used in This Study

Airport	Airport Code	2023 Enplaned Passengers
Cedar Rapids/Iowa City, IA: Eastern Iowa	CID	686,000
Midland/Odessa, TX: Midland International Air and Space Port	MAF	683,000
Lexington, KY: Blue Grass	LEX	661,000
Key West, FL: Key West International	EYW	655,000
Wilmington, NC: Wilmington International	ILM	654,000
Sioux Falls, SD: Joe Foss Field	FSD	651,000
Burlington, VT: Burlington International	BTV	649,000
Harrisburg, PA: Harrisburg International	MDT	648,000
Manchester, NH: Manchester Boston Regional	MHT	644,000
Islip, NY: Long Island MacArthur	ISP	640,000
Santa Barbara, CA: Santa Barbara Municipal	SBA	636,000
Springfield, MO: Springfield-Branson National	SGF	630,000
Jackson/Vicksburg, MS: Jackson Medgar Wiley Evers International	JAN	626,000
Dayton, OH: James M Cox/Dayton International	DAY	589,000
Columbia, SC: Columbia Metropolitan	CAE	586,000
Bend/Redmond, OR: Roberts Field	RDM	558,000
Lubbock, TX: Lubbock Preston Smith International	LBB	539,000
Fargo, ND: Hector International	FAR	516,000
Mission/McAllen/Edinburg, TX: McAllen International	MFE	509,000

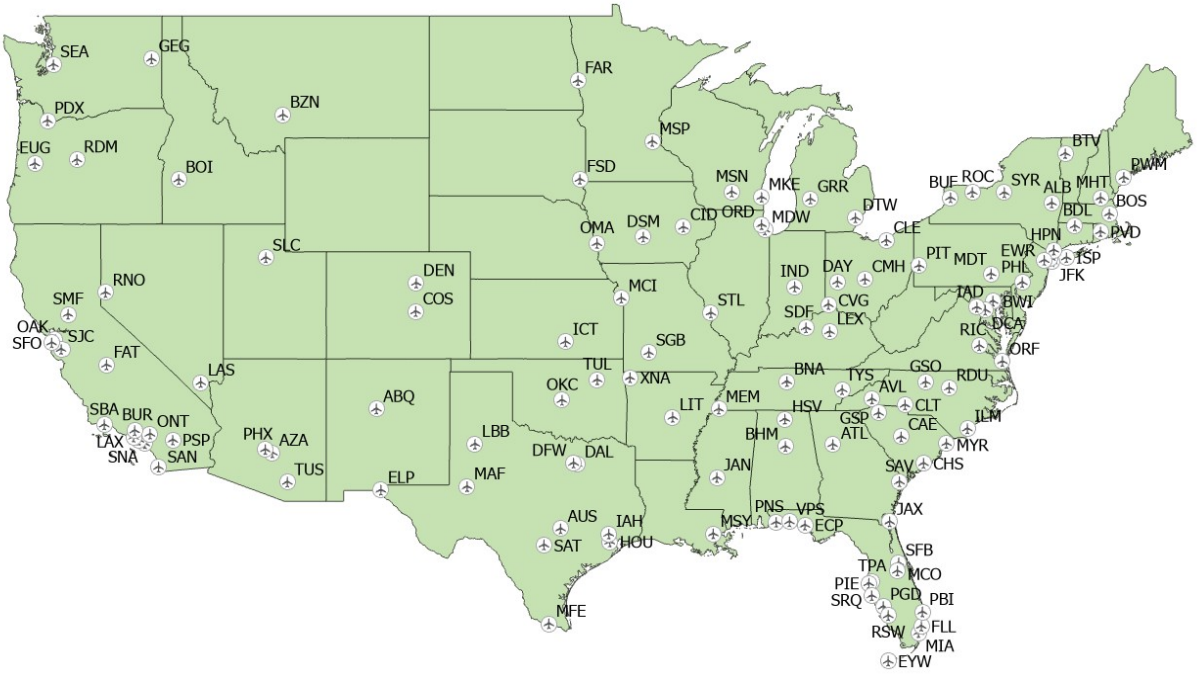


Figure 2: Map of Airports Used in this Study

It is important to note that some states did not have an airport large enough to be considered for this study. Those states were Delaware, West Virginia and Wyoming, while all other states had at least one airport. It should also be noted that this work assumed all trips used the closest airport to their county, regardless of what state the airport is in. Once these airports were determined, a data set with a list of all airport pairs in the dataset was created. This resulted in 14,884 pairs of airports to choose from.

Next, the distance was calculated between the origin/destination county and its closest airport. To do this, at the latitude and longitude of all 122 airports was identified. This was done in the same way that the distance between origin and destination counties were calculated. To calculate travel time and travel distance, the Haversine formula (Formula 1) was used. Once this distance was known, the first step in calculation the travel time for air travel was complete. It should be noted that all calculations performed

in this section were used for both the commute to and from the airport.

The second step in calculating travel times by air travel was calculating the flight time of each air trip. Since the origin airport and destination airport have been determined, an analysis on specific flight times between the airport pairs was performed. Since there is no way to calculate these times by hand, a python code was written which scrapped each flight time from flightroutes.com, a website that catalogs the characteristics of every flight completed each year (67). This code used these online sources to extract the total flight time between each airport pair, many of which included layovers. The chosen source of data (67) factors in layover times, along with time spent taxiing to and from gates, time spent on the runway, and of course, time spent in the air. This was found to be true since a sample of airport pairs were checked to determine if the total travel time provided by flightroutes.com considers these additional times in the calculations.

This calculation not only considers time spent in the air, but also other flying time variables such as layover/connection times, time spent taxiing, etc. It should be noted that each flight between origins and destinations were calculated twice for each set of airports. For example, if an individual was travelling between ATL and LAX airports, the departure from ATL and LAX, along with the trip from LAX to ATL were both calculated. This was done to account for factors such as jet streams and other environmental factors that may increase or decrease flight times. Much like the distances and travel times between county pairs, this data was calculated exclusively for this study and is not a standalone dataset that can be found. Therefore, little cleaning, aside from removing certain airports from the list of considered airports was done. However, total travel time by air differs from travel by personal vehicle in the travel time calculation. Travelling by air involves both

variables: time travelling to/from airports and total flight times. These two travel times were combined into one travel time variable.

4.4 County Attributes and Characteristics

Next, data pertaining to the specific county of residence of the survey respondent was gathered. This data not specific to any NHTS respondent, but are summary statistics collected for all counties in the United States. This data primarily focuses on economic and labor data, traffic related statistics, and environmental data. The data collected has a variety of sources, including data from the United States Census Bureau and other federal agencies.

The first set of county level data pertained to the number of establishments for 18 different sectors at the origin location and 4 at the destination location. For the home location, this included jobs in the accommodation, administration, agriculture, construction, education, entertainment, finance, health, information, management, manufacturing, mining, professional/technical/science, real estate, retail, transportation, utility, and wholesale sectors. For the destination counties, information regarding the accommodation, entertainment, health and retail sectors were kept, since these are the only fields a tourist is likely to encounter on their trip. This data was gathered by the US Census Bureau (64). Additionally, information gathered by the US Census Bureau pertaining to the total population and median household income (Small Area Income and Poverty Index, or SAIPE dataset) were also added to the study (64). This was done for both origin and destination counties. Next, the mean home rent and mean home value were added to the dataset for both the home and destination location. These variables were included to measure the impact that the economic status of both the origin and

destination locations on ER, CR, and SR. This information was gathered from Zillow's Observed Rent Index (Rent) and general trends observed by the company in 2021 for housing costs (68). Next, data pertaining to average precipitation in each origin and destination county was inputted into the dataset. This dataset was collected by the National Center for Environmental Impact, or NCEI and the National Oceanic and Atmospheric Administration (NOAA) (69). This data was collected for all 12 months of the year, but for this study, each average temperature by month was combined for an average yearly temperature. Another variable used for this study was average temperature of the origin and destination county, which were also collected by NOAA. This data was collected in a similar process by NOAA, and each month of data was combined, and the average yearly temperature of the origin and destination counties were computed.

4.5 Efficiency Ratio, Coverage Rate, and Search Radius

This study aimed to understand the relationship between travel times and times at destination for various modes (ER) and the relationship between travel distance (or radius) and time at destination for various modes (CR and SR). Based on the literature and available data that have used the components of ER and CR, and have considered SR, reasonable ER and CR ratios can be determined at the individual level, state level, and national level for both air travel and personal vehicle travel. Also, based on the literature, there will be significant differences in ER and CR for varying household types, demographics, socioeconomic traits, and origins of the traveler. Therefore, the ER, CR and SR for every trip in the data set was calculated. To accurately quantify ER, CR and SR, they first had to be calculated for each trip entry in the NHTS. Both ER and CR were calculated separately using a mixture of variables provided in the NHTS and calculated exclusively for this study. The NHTS provides data pertaining to mode choice, length of

stay, and origin and destination counties. All these variables were crucial variables in determining both ratios. However, as mentioned earlier, the NHTS does not provide data pertaining to travel time and travel distances, so these needed to be calculated. For a more in-depth description of these variable calculations, see the previous section. Once each trip's travel distance and travel time (only travel time calculated was for each trips specified mode) were calculated, they could be compared to the time spent at destination and ER and CR could be calculated.

For this study, a few assumptions were made. Firstly, the airports the traveler used were assumed. The NHTS provide the origin and destination county of the trip, but no information regarding the airports were given. As mentioned in the previous section, an analysis was run to determine the closest airport to the origin and destination airports of each traveler. This analysis includes calculating the flight times between the airport pairs and the time spent traveling between the origin county and airport and the destination airport and county. These travel times and travel distances were calculated using the steps described in the previous section, and the centroids of both counties were used for calculation. The origin and destination counties were the smallest unit of analysis for origin and destination and no information regarding home address and destination address were provided. Therefore, the distance between origin and destination counties was the most exact location that was available for this study. While these distances may vary from the exact distance travelled on the trip, it is the most accurate based on available information.

4.6 Final Dataset

As mentioned before, data was collected from many sources. These sources varied in data purpose and source but were combined into one dataset to quantify and summarize all long-distance trips in the NHTS. Also, data was combined from each of the four categories of data described in this section. This was done to further specify and assign extra trip characteristics to each trip in the NHTS that were not asked about in the survey. For this study, the variables being analyzed were classified as one of five variable types: respondent, household, origin/home location, trip, and destination characteristics. Together, these four variable types created the data for each trip that can determine each respondent's ER, CR and SR. Additionally, the NHTS provides a weight for each variable value. This weight allows for a more accurate count of survey data responses if a larger sample (for this study, the sample size would be around 96,360,540 responses if scaled up).

The final dataset had to have several entries removed to meet the criteria of this study. The initial sample size provided in the NHTS is 16,997 trips. The first trip entries eliminated were those that had missing origin or destination counties. Since the travel time and travel distance are used to compute ER, CR, and SR, trips that provided no locations could not be used. This elimination of data brought the total data set down to 12,804 entries. Next, trips that were considered "work trips" were removed. Since this study was only concerned with leisure trips, business and work trips were removed, with 11,913 trips remaining. Next, trips that were international and were not taken using either a personal vehicle or airplane were removed, bringing the dataset to 11,609 entries. Next, trips that did not have a provided number of nights away were removed. This action left 5,763

entries remaining. Finally, trips that were considered “impossible” were removed from the dataset. These trips were trips where the travel time exceeded the total time away. For example, a trip entry in the dataset travelled from California to Virginia using a personal vehicle was observed to only be one night away. Since the travel time using a personal vehicle would far exceed the total time away, this trip was considered “impossible”. Since trips like this are not logical, they were removed, bringing the final dataset to 4,148 trips. These trips would be the trips used in this work. This set of actions resulted in the weighted sample sizes and statistics shown on the following pages in tables 2 and 3. Additionally, a diagram of the data merging process can be found below in Figure 3.

Figure 3: Data Merging Process

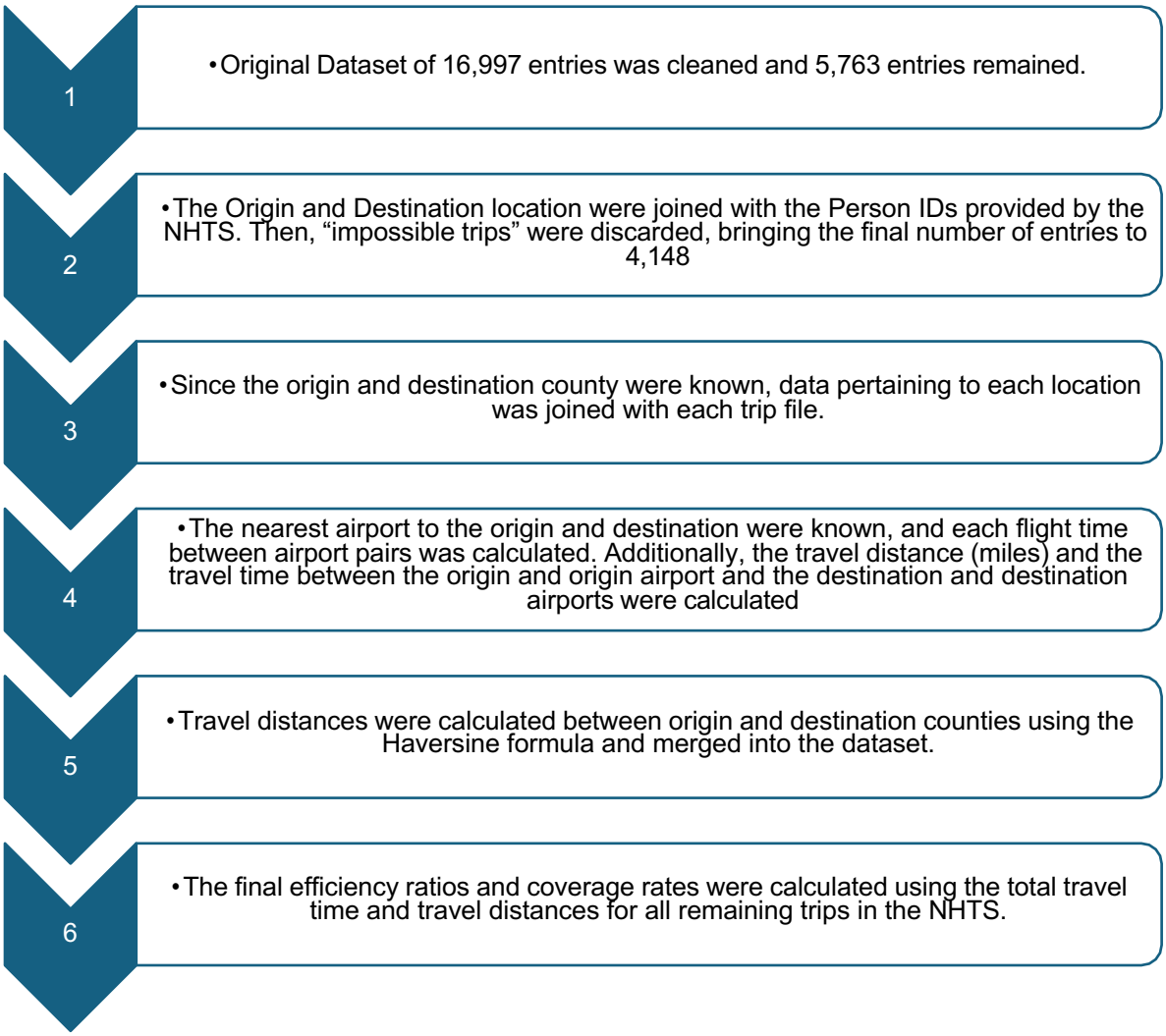


Table 2: Data Set Characteristics (Unweighted)

Respondent Characteristics						
Variable and Variable Range	Count	% of Responses	Min	Max	Mean	Std. Deviation
Respondent age	-	-	5	92	47.07	21.14
Sex of Respondent						
…Male	1971	47.5%	-	-	-	-
…Female	2177	52.5%	-	-	-	-
Respondent is Employed	1298	-	-	-	-	-
Respondent Race						
…Hispanic, No Other Races	335	8.1%	-	-	-	-
…White, Non-Hispanic	3329	80.3%	-	-	-	-
…Black/African American, Non-Hispanic	218	5.3%	-	-	-	-
…Asian, Non-Hispanic	210	5.1%	-	-	-	-
…Other, Non-Hispanic	56	1.4%	-	-	-	-
Education						
…Less than High School	630	15.2%	-	-	-	-
…Some High School or Trade School	1352	32.6%	-	-	-	-
…Undergraduate Degree	1212	29.2%	-	-	-	-
…Graduate Degree	954	23.0%	-	-	-	-
Income						
…Less than $24,000	222	5.30%	-	-	-	-
…Between $25,000 and $49,999	496	11.96%	-	-	-	-
…Between $50,000 and $99,999	1251	30.00%	-	-	-	-
…Between $100,000 and $149,999	1048	25.27%	-	-	-	-
…Greater than $150,000	1131	27.27%	-	-	-	-
Home Ownership Status						
…Owned	3409	82.2%	-	-	-	-
…Rented or Occupied without Payment	739	17.8%	-	-	-	-

Table 2 Continued: Data Set Characteristics (Unweighted)

Variable and Variable Range	Count	% of Responses	Min	Max	Mean	Std. Deviation
Type of Home Respondent Lives In						
…Single Family, Detached	3232	77.9%	-	-	-	-
…Single Family, Townhome or Condo	341	8.2%	-	-	-	-
…Apartment Building	486	11.7%	-	-	-	-
…Mobile Home	85	2.1%	-	-	-	-
…RV, Van, Boat, etc.	4	0.1%	-	-	-	-
Travel Is Difficult due to Medical Condition						
…Yes	174	4.2%	-	-	-	-
…No	3974	95.8%	-	-	-	-
Respondent is in School	99	-	-	-	-	-
Household Characteristics						
Total number of people in household	-	-	1	10	2.8	1.4
Count of household members under 5 years old	-	-	0	5	0.6	1.0
Count of household members 5-17 years old	-	-	0	4	0.4	0.4
Count of adult household members at least 18 years old	-	-	1	7	2.1	0.7
Total number of vehicles in household	-	-	0	8	2.2	1.1
Number of drivers in the household	-	-	0	7	2.0	0.7
Count of workers in household	-	-				
Number of times goods delivered in past 30 days	-	-	0	99	4.8	6.6
Origin/Home Location Characteristics						
Total Population (100,000s of People)	-	-	1328	10019635	889583.3	1511810.1
Average Yearly Temperature (°F)	-	-	37.62	77.04	58.7	7.6
Average Yearly Precipitation (inches)	-	-	0.36	8.97	3.7	1.5
GDP of County (in Millions of $)	-	-	32735	711874201	64900945.7	120329313.2
Mean Home Rent (in Thousands of $)	-	-	664	3539	1584.2	482.6
Mean Home Value (in Thousands of $1,000)	-	-	54230.65	1573548	331247.7	200915.8

Table 2 Continued: Data Set Characteristics (Unweighted)

Variable and Variable Range	Count	% of Responses	Min	Max	Mean	Std. Deviation
Median Household Income ($1,000)	-	-	31091	153716	71726.5	18289.9
Number of Establishments (1,000s of Establishments)						
...All Sectors	-	-	120	3629082	374084.3	612559.8
...Accomodation	-	-	3	23781	2102.0	3596.7
...Administration	-	-	3	11787	1232.9	1906.6
...Agriculture	-	-	3	201	22.0	30.8
...Construction	-	-	3	16915	1872.5	2818.0
...Education	-	-	3	4548	390.0	686.3
...Entertainment	-	-	3	20972	637.6	2482.5
...Finance	-	-	3	13463	1361.7	2260.1
...Health	-	-	3	36149	2832.6	5132.3
...Information	-	-	3	13241	583.6	1628.7
...Management	-	-	3	1376	194.4	284.2
...Manufacturing	-	-	3	10620	760.2	1473.3
...Mining	-	-	3	879	39.2	119.5
...Professional	-	-	3	42070	3285.7	6259.1
...Real Estate	-	-	3	18713	1476.1	2771.3
...Retail	-	-	6	28975	2649.4	4316.2
...Transportation	-	-	3	10754	827.0	1786.6
...Utility	-	-	3	343	35.0	51.3
...Whosesale	-	-	3	21470	1285.6	2849.8
Urban/Rural Indicator						
…Rural	916	22.1%	-	-	-	-
…Small Town	1023	24.6%	-	-	-	-
…Second City	691	16.7%	-	-	-	-
…Suburban	1038	25.0%	-	-	-	-
…Urban	480	11.6%	-	-	-	-
Distance From Origin to Origin Airport (Miles)	-	-	1	373	79	51

Table 2 Continued: Data Set Characteristics (Unweighted)

Trip/Tour Characteristics						
Variable and Variable Range	Count	% of Responses	Min	Max	Mean	Std. Deviation
Mode Choice for Trip						
…Personal Vehicle/Car	3718	89.6%	-	-	-	-
…Airplane	430	10.4%	-	-	-	-
Trip Purpose						
…Not Provided	4	0.0%	-	-	-	-
…Personal Vacation	1463	35.3%	-	-	-	-
…Visiting Family or Friends	1849	44.6%	-	-	-	-
…Other	832	20.1%	-	-	-	-
Number of people with respondent on long distance trip	-	-	0	35	2.5856	1.55555
Nights away on long distance trip	-	-	0	31	2.75	3.709
Travel Time for Flight (Hours)	-	-	0.9	373.1	79.3	48.5
Distance From Destination to Destination Airport (Miles)	-	-	0.9	373.1	76.8	50.3
Month the Trip was Taken In						
…January	284	6.8%	-	-	-	-
…February	266	6.4%	-	-	-	-
…March	288	6.9%	-	-	-	-
…April	380	9.2%	-	-	-	-
…May	342	8.2%	-	-	-	-
…June	229	5.5%	-	-	-	-
…July	361	8.7%	-	-	-	-
…August	277	6.7%	-	-	-	-
…September	252	6.1%	-	-	-	-
…October	360	8.7%	-	-	-	-
…November	633	15.3%	-	-	-	-
…December	476	11.5%	-	-	-	-

Table 2 Continued: Data Set Characteristics (Unweighted)

Variable and Variable Range	Count	% of Responses	Min	Max	Mean	Std. Deviation
Trip Took Place Over Weekend						
…Yes	2371	57.2%	-	-	-	-
…No	1777	42.8%	-	-	-	-
Distance Between Counties	-	-				
Total Time Spent Flying and Commuting Between Airports (Hours)	-	-	0.09	16.1	5.5	2.3
Total Time Spent Driving (Hours)	-	-	0	41.5	12.7	9.2
Hours Spent on Trip (Does not Include Travel Time)	-	-	24	768.0	89.9	89.0
Days away on Trip	-	-	1	32.0	3.7	3.7
Trip Destination Characteristics						
GDP of County (in Millions of $)	-	-	48534	711874201	63019901	124554680
Mean Home Rent (in Thousands of $)	-	-	663.97	3539.28	1640.79	523.47
Mean Home Value (in Thousands of $1,000)	-	-	54230.65	2097025.01	351820.39	228299.62
Median Household Income (in Thousands of $)	-	-	29143.00	153716.00	68615.28	16416.99
Number of Establishments (1,000s of Establishments)						
...All Sectors	-	-	337	3629082	347334.62	590622.785
...Accommodation	-	-	5	23781	1957.04	3403.727
...Entertainment	-	-	3	20972	616.22	2342.99
...Health	-	-	3	36149	2515.38	4815.088
...Retail	-	-	7	28975	2403.29	4041.649
Average Yearly Precipitation (inches)	-	-	0.2	9.0	3.6	1.5
Average Yearly Temperature (°F)	-	-	37.3	77.0	58.7	7.9
Total Population (100,000s of People)	-	-	2052.0	10019635.0	778880.5	1403161.0

Table 3: Data Set Characteristics (Weighted)

Respondent Characteristics						
Variable and Variable Range	Count	% of Responses	Min	Max	Mean	Std. Deviation
Respondent age	-	-	5	92	47.07	21.14
Sex of Respondent						
…Male	33,191,759	47.5%	-	-	-	-
…Female	36,687,353	52.5%	-	-	-	-
Respondent is Employed	1,565,155	-				
Respondent Race						
…Hispanic, No Other Races	11,121,691	15.9%	-	-	-	-
…White, Non-Hispanic	45,526,464	65.2%	-	-	-	-
…Black/African American, Non-Hispanic	6,392,456	9.1%	-	-	-	-
…Asian, Non-Hispanic	3,599,056	5.2%	-	-	-	-
…Other, Non-Hispanic	3,239,445	4.6%	-	-	-	-
Education						
…Less than High School	14,019,929	20.1%	-	-	-	-
…Some High School or Trade School	29,758,930	42.6%	-	-	-	-
…Undergraduate Degree	13,498,340	19.3%	-	-	-	-
…Graduate Degree	12,601,913	18.0%	-	-	-	-
Income						
…Less than $24,000	4,340,365	6.2%	-	-	-	-
…Between $25,000 and $49,999	8,912,474	12.8%	-	-	-	-
…Between $50,000 and $99,999	20,096,975	28.8%	-	-	-	-
…Between $100,000 and $149,999	15,952,221	22.8%	-	-	-	-
…Greater than $150,000	20,577,077	29.4%	-	-	-	-
Home Ownership Status						
…Owned	54,320,914	77.7%	-	-	-	-
…Rented or Occupied without Payment	15,558,198	22.3%	-	-	-	-

Table 3 Continued: Data Set Characteristics (Weighted)

Variable and Variable Range	Count	% of Responses	Min	Max	Mean	Std. Deviation
Type of Home Respondent Lives In						
…Single Family, Detached	52,153,842	74.6%	-	-	-	-
…Single Family, Townhome or Condo	6,116,144	8.7%	-	-	-	-
…Apartment Building	9,694,107	13.9%	-	-	-	-
…Mobile Home	1,856,168	2.7%	-	-	-	-
…RV, Van, Boat, etc.	58,851	0.1%	-	-	-	-
Travel Is Difficult due to Medical Condition						
…Yes	2,616,349	3.7%	-	-	-	-
…No	67,262,763	96.3%	-	-	-	-
Respondent is in School	1,565,155	-	-	-	-	-
Household Characteristics						
Total number of people in household	-	-	1	10	2.8	1.4
Count of household members under 5 years old	-	-	0	5	0.6	1.0
Count of household members 5-17 years old	-	-	0	4	0.4	0.4
Count of adult household members at least 18 years old	-	-	1	7	2.1	0.7
Total number of vehicles in household	-	-	0	8	2.2	1.1
Number of drivers in the household	-	-	0	7	2.0	0.7
Count of workers in household	-	-				
Number of times goods delivered in past 30 days	-	-	0	99	4.8	6.6
Count of household trips on travel day	-	-	1	10	2.8	1.4

Table 3 Continued: Data Set Characteristics (Weighted)

Origin/Home Location Characteristics						
Variable and Variable Range	Count	% of Responses	Min	Max	Mean	Std. Deviation
Total Population	-	-	1328	10019635	889583.3	1511810.1
Average Yearly Temperature (°F)	-	-	37.62	77.04	58.7	7.6
Average Yearly Precipitation (inches)	-	-	0.36	8.97	3.7	1.5
GDP Origin	-	-	32735	711874201	64900945.7	120329313.2
Mean Home Rent	-	-	664	3539	1584.2	482.6
Mean Home Value	-	-	54230.65	1573548	331247.7	200915.8
Median Household Income ($1,000)	-	-	1328	10019635	889583.3	1511810.1
Number of Establishments (1,000s of Establishments)						
...All Sectors	-	-	31091	153716	71726.5	18289.9
...Accomodation	-	-				
...Administration	-	-	120	3629082	374084.3	612559.8
...Agriculture	-	-	3	23781	2102.0	3596.7
...Construction	-	-	3	11787	1232.9	1906.6
...Education	-	-	3	201	22.0	30.8
...Entertainment	-	-	3	16915	1872.5	2818.0
...Finance	-	-	3	4548	390.0	686.3
...Health	-	-	3	20972	637.6	2482.5
...Information	-	-	3	13463	1361.7	2260.1
...Management	-	-	3	36149	2832.6	5132.3
...Manufacturing	-	-	3	13241	583.6	1628.7
...Mining	-	-	3	1376	194.4	284.2
...Professional	-	-	3	10620	760.2	1473.3
...Real Estate	-	-	3	879	39.2	119.5
...Retail	-	-	3	42070	3285.7	6259.1
...Transportation	-	-	3	18713	1476.1	2771.3
...Utility	-	-	6	28975	2649.4	4316.2
...Whosesale	-	-	3	10754	827.0	1786.6

Table 3 Continued: Data Set Characteristics (Weighted)

Variable and Variable Range	Count	% of Responses	Min	Max	Mean	Std. Deviation
Urban/Rural Indicator						
…Rural	15,145,135	21.7%	-	-	-	-
…Small Town	16,462,387	23.6%	-	-	-	-
…Second City	12,175,418	17.4%	-	-	-	-
…Suburban	17,493,701	25.0%	-	-	-	-
…Urban	8,602,471	12.3%	-	-	-	-
Distance From Origin to Origin Airport (Miles)	-	-	1	373	79	51
Trip/Tour Characteristics						
Mode Choice for Trip						
…Personal Vehicle/Car	64,166,629	91.8%	-	-	-	-
…Airplane	5,712,482	8.2%	-	-	-	-
Trip Purpose						
…Not Provided	57,890	0.1%	-	-	-	-
…Personal Vacation	25,077,649	35.9%	-	-	-	-
…Visiting Family or Friends	31,020,816	44.4%	-	-	-	-
…Other	13,722,757	19.6%	-	-	-	-
Number of people with respondent on long distance trip	-	-	0	35	2.5856	1.55555
Nights away on long distance trip	-	-	0	31	2.75	3.709
Travel Time for Flight (Hours)	-	-	0.9	373.1	79.3	48.5
Month the Trip was Taken In	-	-	0.9	373.1	76.8	50.3
…January	5,684,132	8.1%				
…February	3,629,196	5.2%	-	-	-	-
…March	4,567,521	6.5%	-	-	-	-
…April	5,582,731	8.0%	-	-	-	-
…May	6,856,351	9.8%	-	-	-	-
…June	5,820,237	8.3%	-	-	-	-
…July	9,136,653	13.0%	-	-	-	-

Table 3 Continued: Data Set Characteristics (Weighted)

Variable and Variable Range	Count	% of Responses	Min	Max	Mean	Std. Deviation
…September	5,793,383	8.3%	-	-	-	-
…October	6,430,846	9.2%	-	-	-	-
…November	6,168,373	8.8%	-	-	-	-
…December	5,315,931	7.6%	-	-	-	-
Trip Took Place Over Weekend						
…Yes	42,249,893	60.5%	-	-	-	-
…No	27,629,218	39.5%	-	-	-	-
Distance Between Counties	-	-				
Total Time Spent Flying and Commuting Between Airports (Hours)	-	-	0.09	16.1	5.5	2.3
Total Time Spent Driving (Hours)	-	-	0	41.5	12.7	9.2
Hours Spent on Trip (Does not Include Travel Time)	-	-	24	768.0	89.9	89.0
Days away on Trip	-	-	1	32.0	3.7	3.7
Trip Destination Characteristics						
GDP of County (in Millions of $)	-	-	48534	711874201	63019901	124554680
Mean Home Rent (in Thousands of $)	-	-	663.97	3539.28	1640.79	523.47
Mean Home Value (in Thousands of $1,000)	-	-	54230.65	2097025.01	351820.39	228299.62
Median Household Income (in Thousands of $)	-	-	29143.00	153716.00	68615.28	16416.99
Number of Establishments (1,000s of Establishments)						
...All Sectors	-	-	337	3629082	347334.62	590622.785
...Accomodation	-	-	5	23781	1957.04	3403.727
...Entertainment	-	-	3	20972	616.22	2342.99
...Health	-	-	3	36149	2515.38	4815.088
...Retail	-	-	7	28975	2403.29	4041.649
Average Yearly Precipitation (inches)	-	-	0.2	9.0	3.6	1.5
Average Yearly Temperature (°F)	-	-	37.3	77.0	58.7	7.9
Total Population (100,000s of People)	-	-	2052.0	10019635.0	778880.5	1403161.0

The values unweighted data set represent the sample used for this study. In total, 4,148 trips were used. Each trip had its own unique characteristics, along with the respondents who filled out the surveys. The weighted sample represents the given sample in the NHTS if it was scaled up for a much larger population, which was around 70 million people. These distributions were computed and displayed for several reasons, the first being to compare the data to national trends. First, most of the population of the United States identifies as “White or Caucasian”, with the next largest sample size being those who were “Black or African American”. These two trends (in the weighted sample), closely mirror. Additionally, both the weighted and unweighted sample observe many levels of income earners. Overall, the main conclusions that can be drawn from this data set is that it is representative of the population of the United States, since it surveyed a variety of Americans with different socioeconomic and demographic backgrounds.

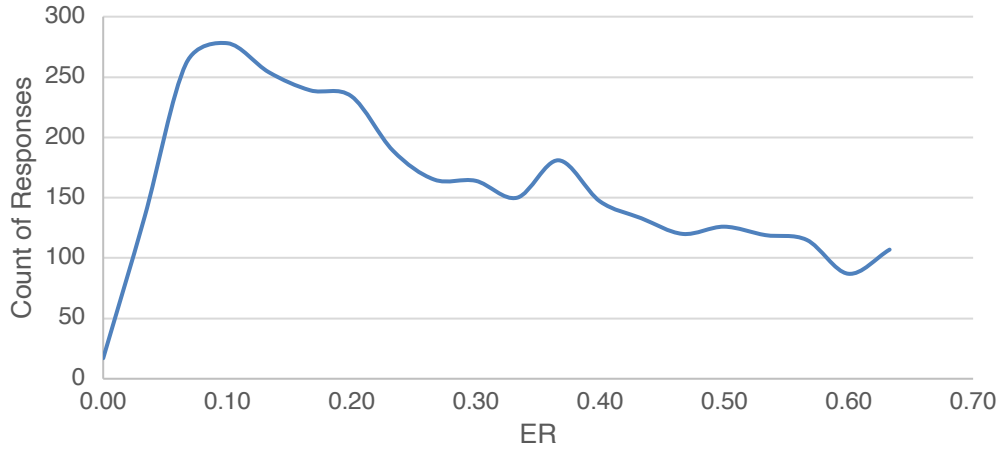
Chapter 5: Analysis Methods

This study aimed to understand the relationship between travel times and times at destination for various modes (ER) and the relationship between travel distance (or radius) and time at destination for various modes (CR). Based on the literature and available data, there will be significant differences in ER and CR for varying household types, demographics, socioeconomic traits, and origins of the traveler.

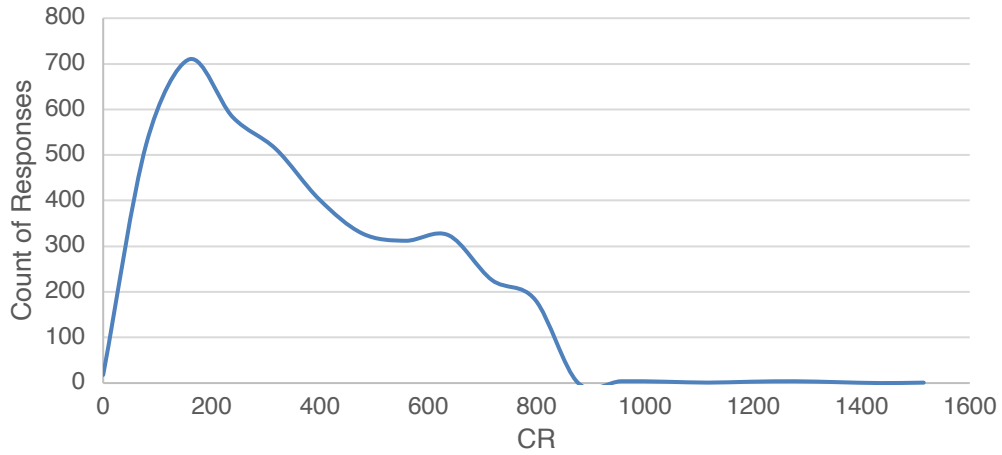
This study utilized gamma regressions. A gamma regression is a generalized type of linear regression used for datasets with positive skew. These types of regressions are also used when a dependent variable cannot be zero or negative. Gamma regressions also have the capability to observe how a variable will increase or decrease by a certain percentage. This differs from a normal linear regression, which quantifies impact by increases or decreases in a value by a fixed amount every single time. This regression type was chosen for these data sets due to the positive skew of all three datasets, thus leading to a gamma regression for ER, CR, and SR. This skew was determined by examining the Probability Distribution Function (PDF) and Cumulative Distribution Function (CDF) of all three variables. A PDF measures the distribution of variable values in a dataset while a CDF measures the likelihood of a variable being less than or equal to a certain value. Additionally, the PDF and CDFs were calculated again with the weighted samples to observe how representative the ER, CR and SR samples were for the population of the United States.

Figure 4: Probability Distribution Functions

(a) ER



(b) CR



(c) SR

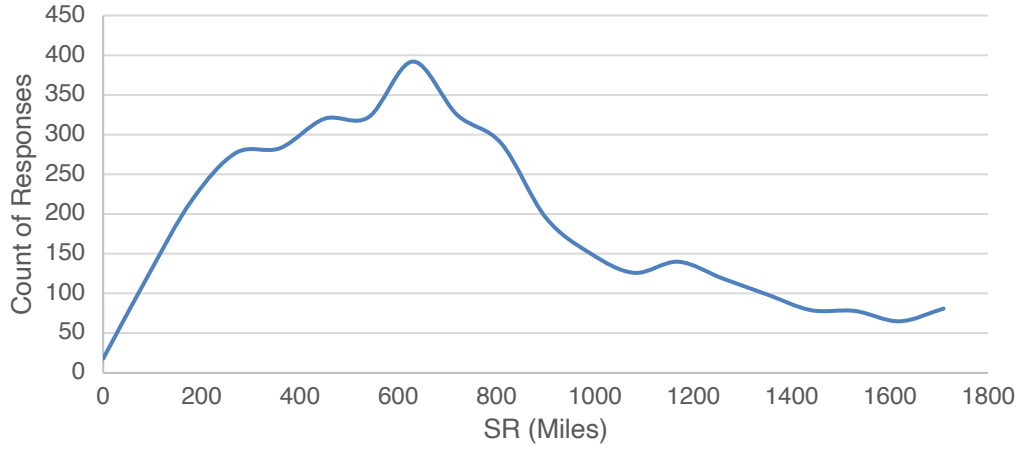
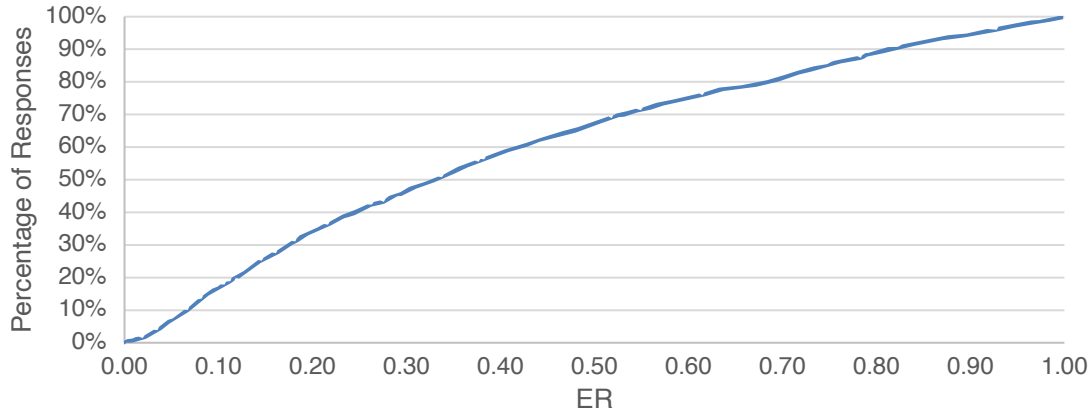
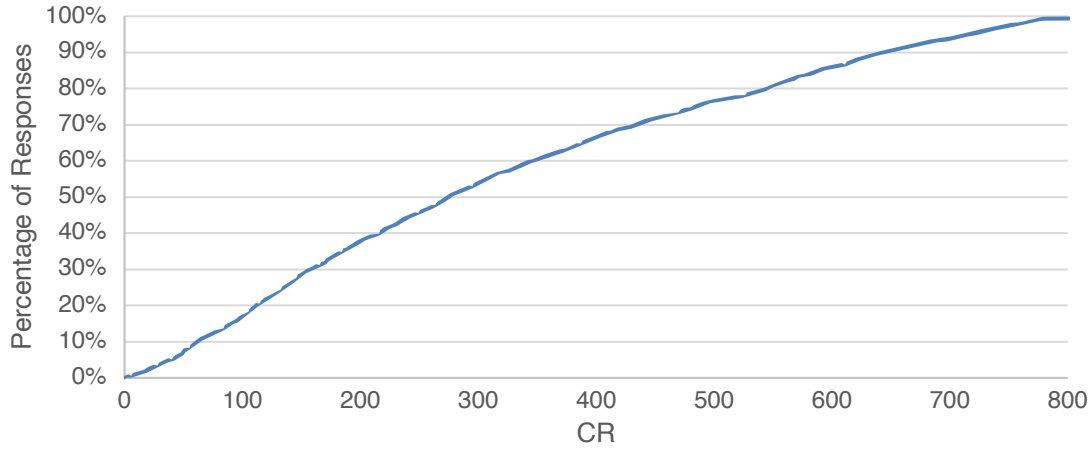


Figure 5: Cumulative Distribution Functions

(a) ER



(b) CR



(c) SR

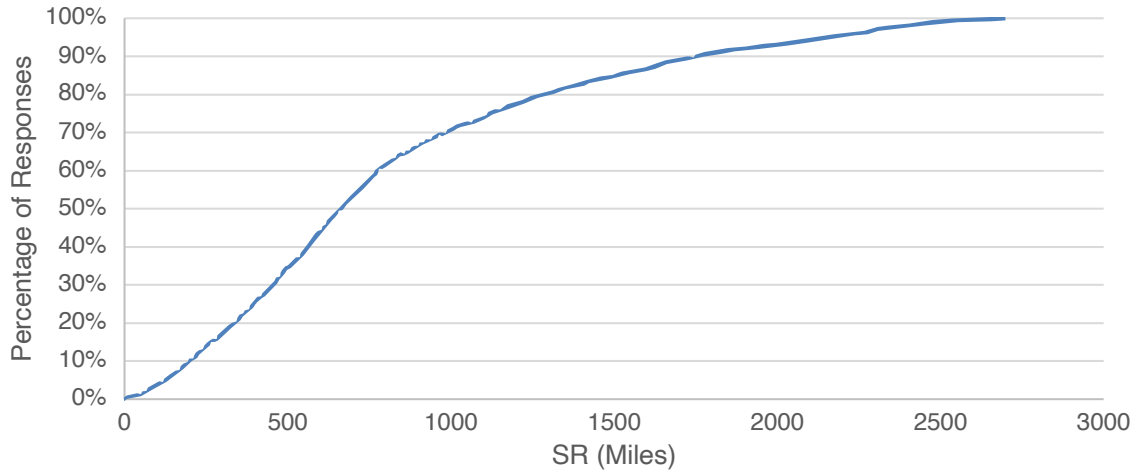
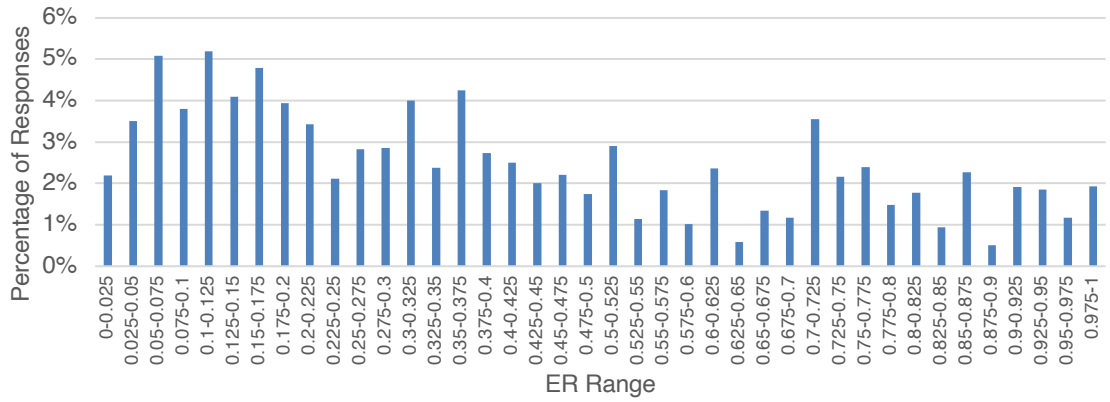
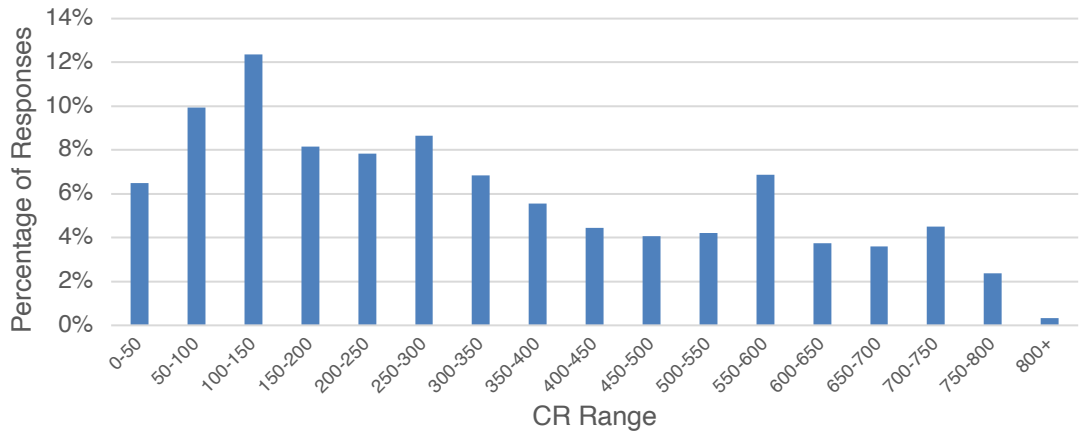


Figure 6: Weighted Probability Density Functions

(a) ER



(b) CR



(c) SR

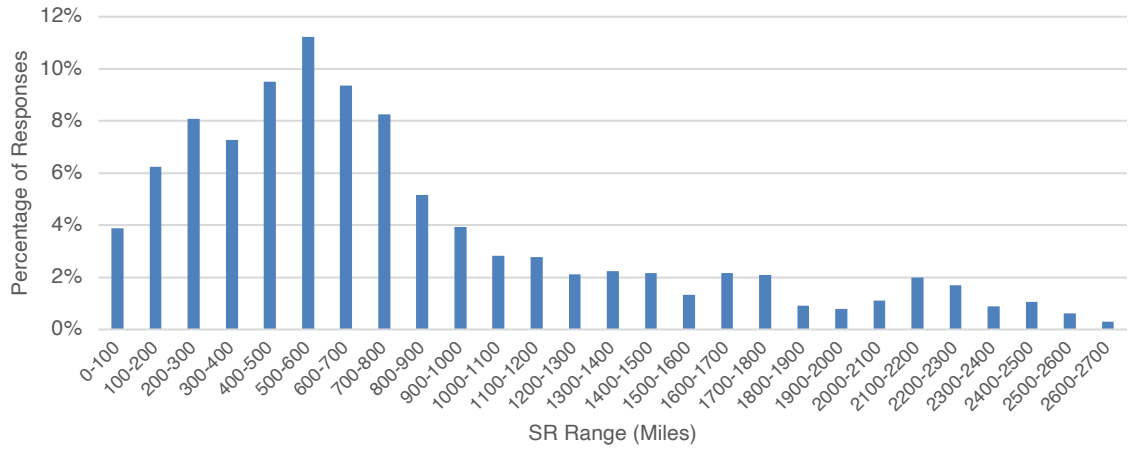
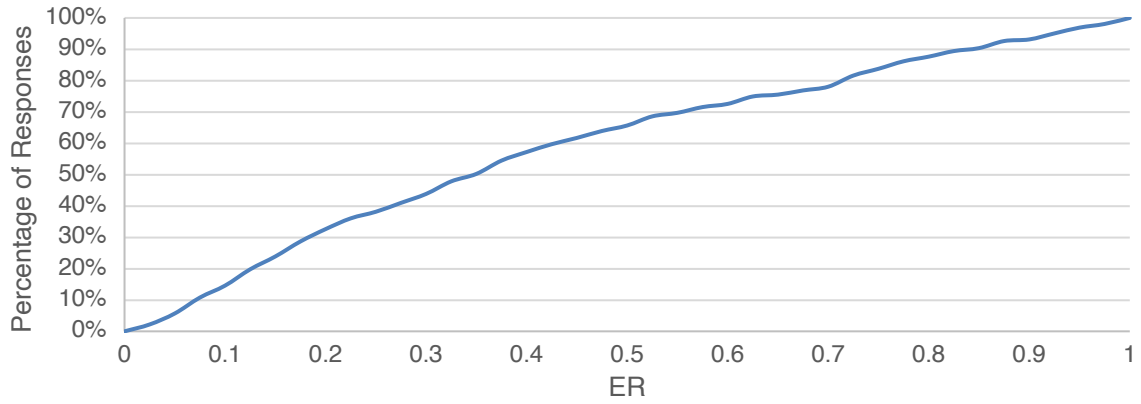
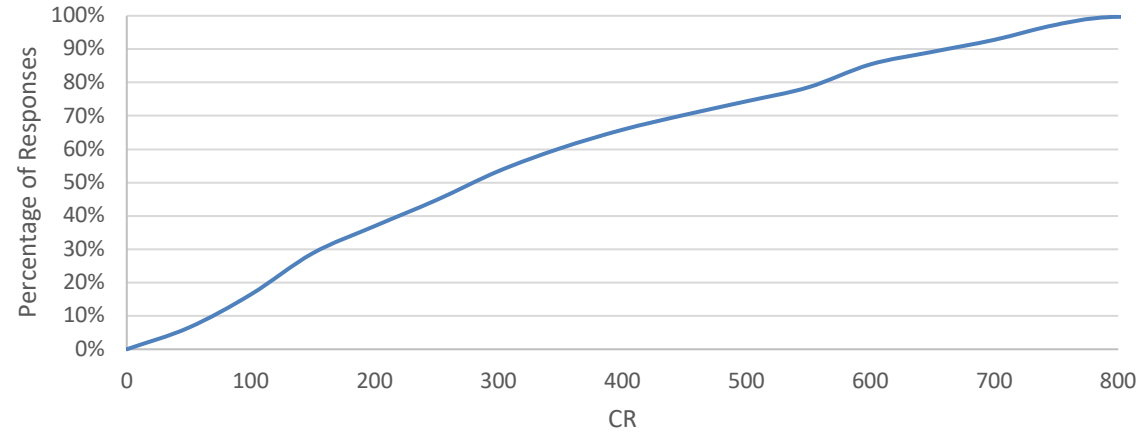


Figure 7: Weighted Cumulative Distribution Functions

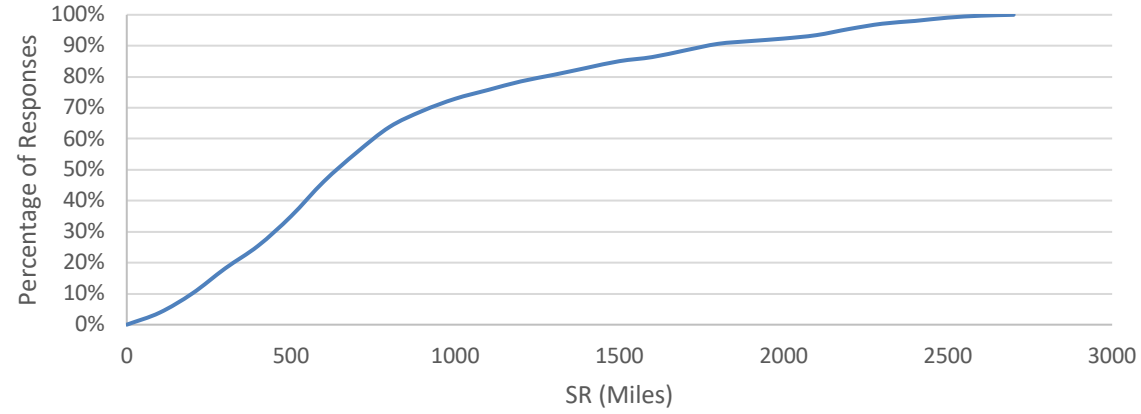
(a) ER



(b) CR



(c) SR



These sets of data provide insight into the distribution of the data. Each PDF has “positive skew”, or data that is concentrated on the left side of the graph. This indicates that the data is concentrated in the lower ranges the data can be in. Therefore, a gamma distribution was the best fit for this analysis. Additionally, both the PDF and CDFs for the weighted samples are very similar, including the positive skew. This trend indicates that the data is representative for the population of the United States. Furthermore, these findings indicated that a gamma regression would be the best fit for the unweighted data since the shape of the weighted data PDFs and CDFs were positively skewed.

Chapter 6: Results and Discussion

6.1 Summary Statistics for Models and Key Variable Comparisons

For this study, three dependent variables were used for analysis: ER, CR, and SR. All three were calculated using the steps described in the previous section. The first variables that were compared were the ER of personal vehicle and air trips and CR, which can be found in Figures 8 and 9. It should be noted that this section was used to compare ER and Cr, or the travel times and distance travelled. Therefore, SR was not included in this section.

Figure 8: Distribution of ER Values for Personal Vehicle Trips Compared to CR

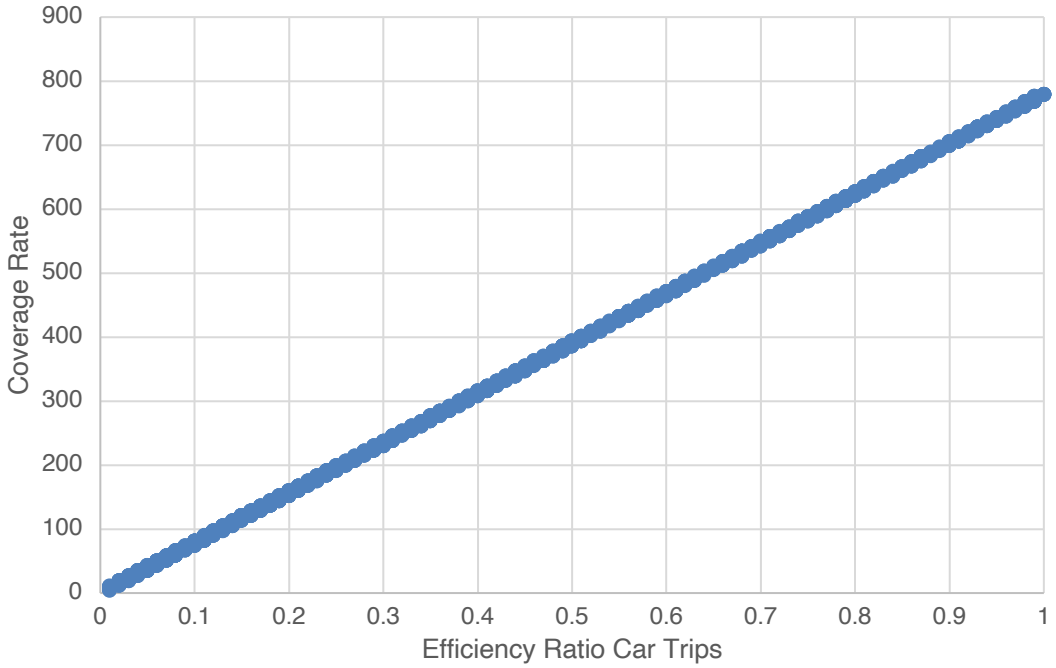
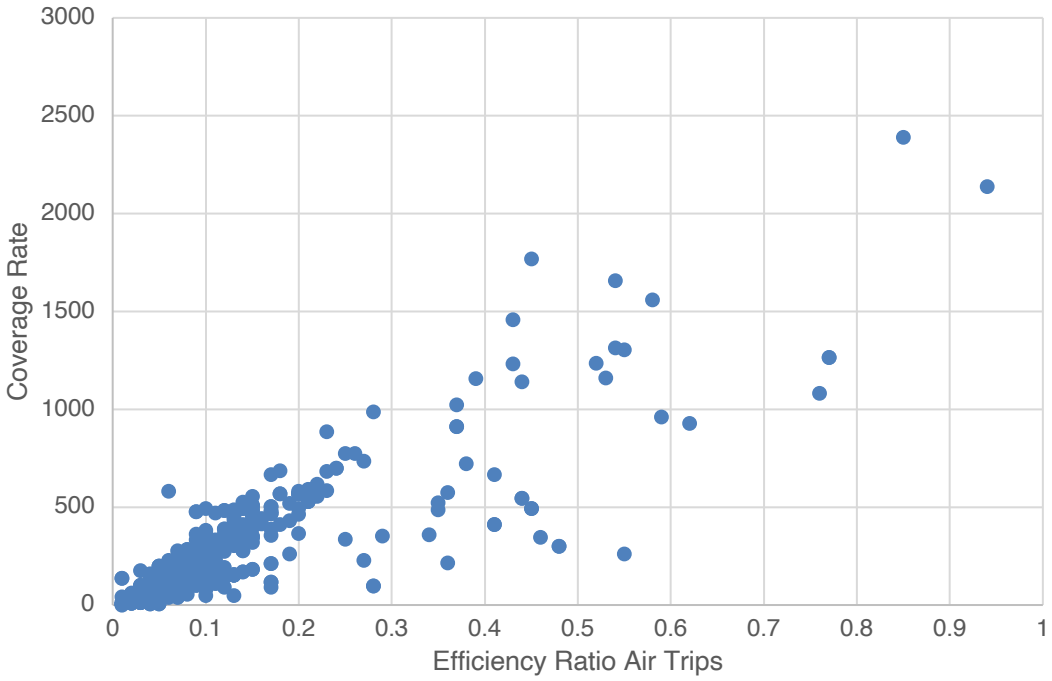
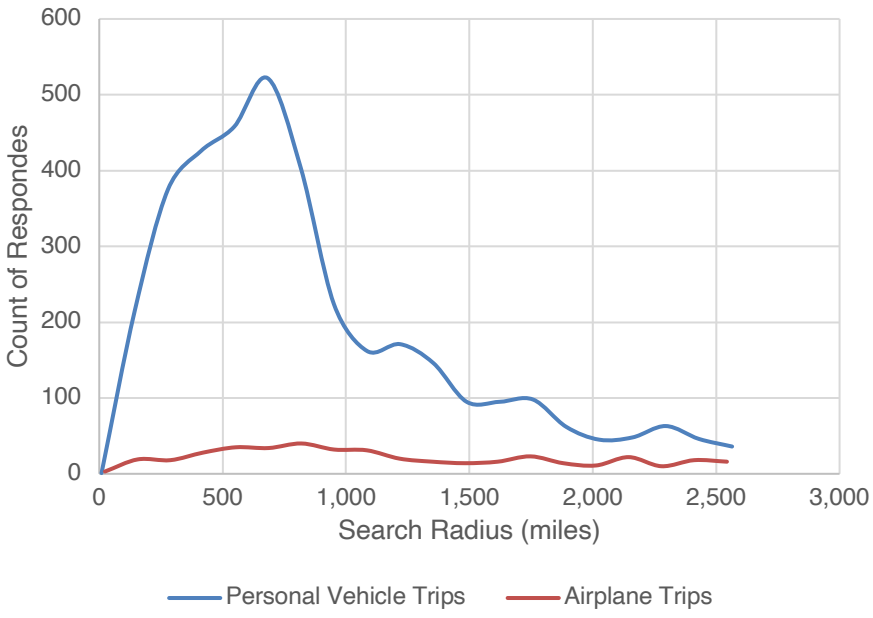


Figure 9: Distribution of ER Values for Airplane Trips Compared to CR



While the trends for personal vehicle travel may seem strange, they are explainable based on the way each were calculated. As mentioned previously, travel times for personal vehicle trips were calculated using travel distances. For this study, a speed 65 miles per hour was used. This means that, for car trips, as travel times increase, travel distance increase at the exact same rate, thus creating the correlation seen in Figure 8. The trends in the distribution of ER and CR values for airplane trips are less straightforward. For these trips, a large cluster of the trips can be seen in the lower ranges of both the ER and CR data, which explains the shapes of the PDFs and CDFs of the ER and CR models. Next, the distribution of values for each trip's search radius based on the mode of the trip were calculated. This distribution aimed to compare the observed average miles travelled by car versus average miles travelled by air. The results of this distribution can be found below in Figure 10.

Figure 10: Distribution of SR for Airplane vs. Personal Vehicle Trips



The results in Figure 10 reveal a couple of trends regarding the data. First, the average search radius for personal vehicle trips was 794 miles, while the average distance for airplane trips was 1,143 miles. Furthermore, this data set contains an even distribution of trips that were taken by air for varying search radii. Since only around 400 trips were studied in this work were airplane trips, future NHTS versions should aim to collect more information regarding airplane trips. In this specific data set, however, many airplane trips occurred between 500 and 1000 miles, which correlates with findings in the literature that describe the typical ranges of airplane travelers. These findings also follow general trends pertaining to average distance taken by personal vehicle. The literature specifies that most long-distance trips taken by car are within 250 and 500 miles of the origin, and this data set is representative of those findings. In this data set, around 40% of the trips taken were within 500 miles. These same trends were observed for the ER of personal vehicle and airplane trips. The results can be found below in Figure 11.

Figure 11: ER for Personal Vehicle and Airplane Trips

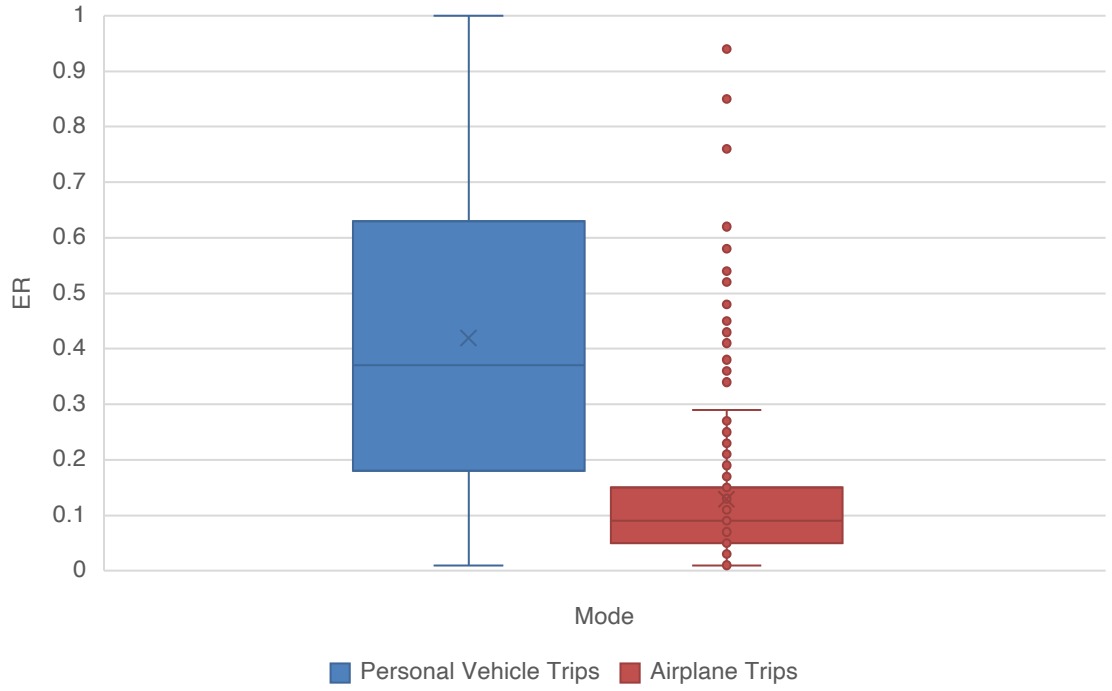


Figure 11 reveals a few key trends in ER for both personal vehicle and airplane trips, including the difference in average ER. For personal vehicle trips, the ER was 0.42 on average, while airplane trips had an average ER of 0.13. This difference is significant and further confirms findings from the literature, which indicate that airplane trips may cover longer distances but have lower travel times than personal vehicle trips. Additionally, airplane trips do not differ much in travel times in relation to time at destination. Furthermore, personal vehicle trips have varied ER values and a larger range of ER values. This indicates that personal vehicle trips are more likely to have much higher and much lower travel time ranges compared to airplane trips.

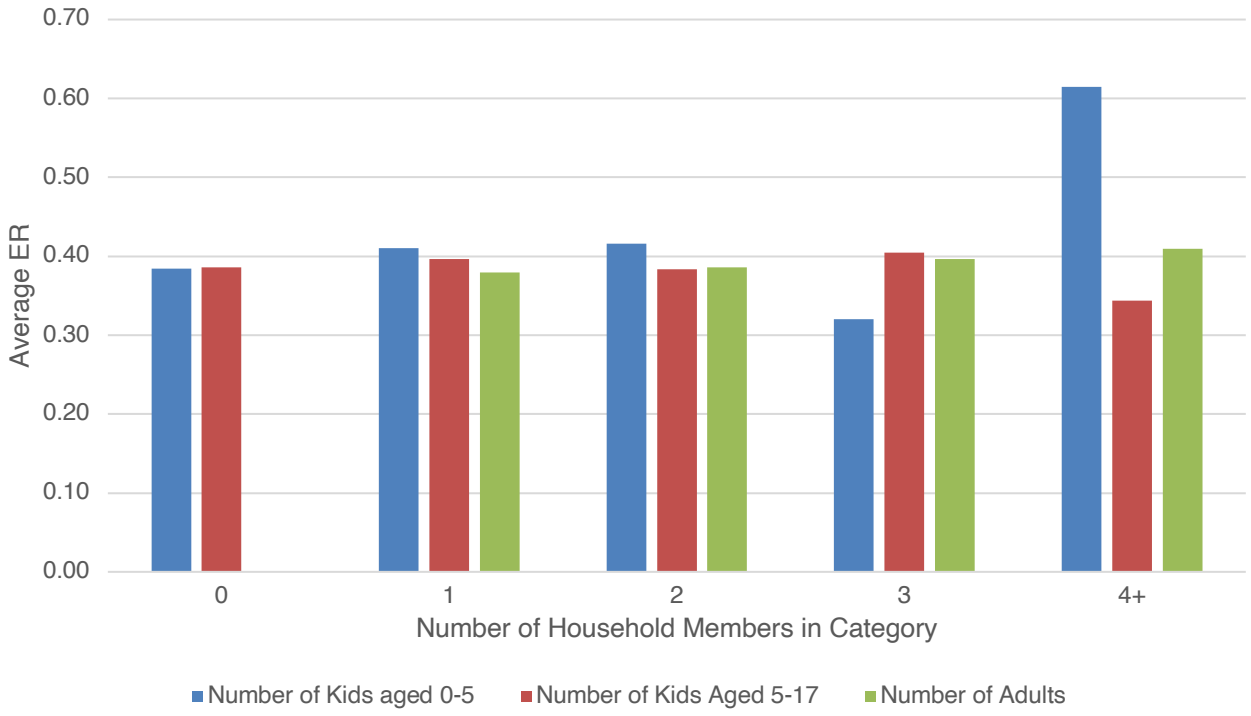
Once these variables were summarized, they were compared to a few variables collected in the National Household Travel Survey. While over 100 variables were used for analysis in this study, a few were observed to see how ER and CR change as certain variables change. Future work should compare ER and CR to all independent variables used in this study to observe trends in ER and CR changes. For this work however, only a few variables were analyzed. Those were number of

household members within age ranges (0-5 years old, 5–17 years old, and 18 years and older), income of respondent, reason respondent and travel party are taking the trip, and mode the trip was taken with. For this model, the ERs used for both personal vehicle and airplane choices were based on the actual mode taken by the traveler and not the predicted mode. This was not done for figures that compare ER and CR, which combine all ER values into the single data set.

As mentioned, the first variable analyzed was the number of household members within certain age ranges. In the data, there were between zero and four children under the age of five, zero and five children between 5 and 17 years old, and between one and seven adults in each household. Future work should also be done to observe ER and CR changes as the number of family members in each age range increases. The results of both ER and CR trends can be found in Figure 12 below.

Figure 12: Various Counts of Household Member Ages Comparisons

(a) ER



(b) CR

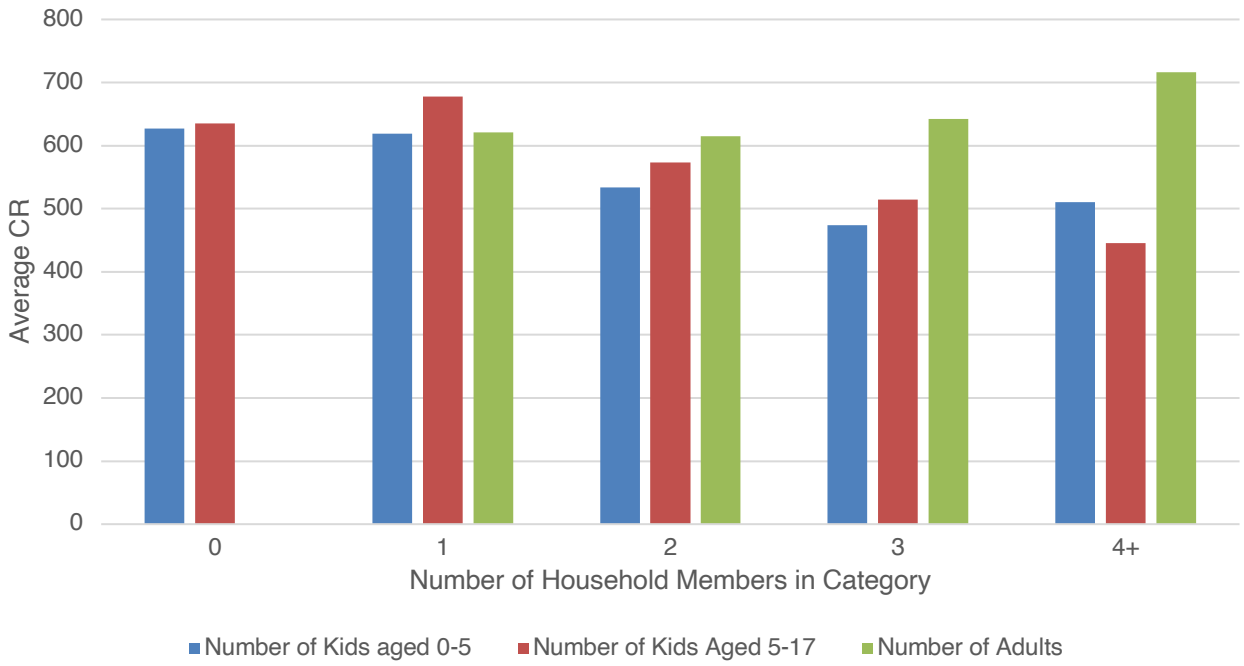
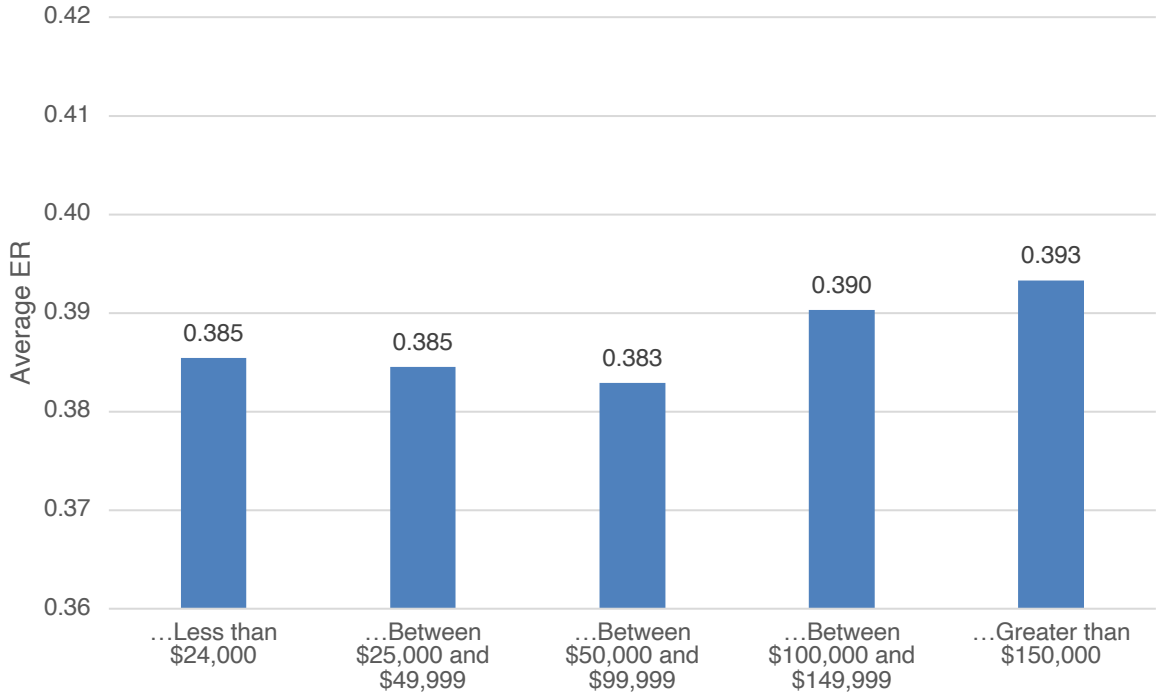


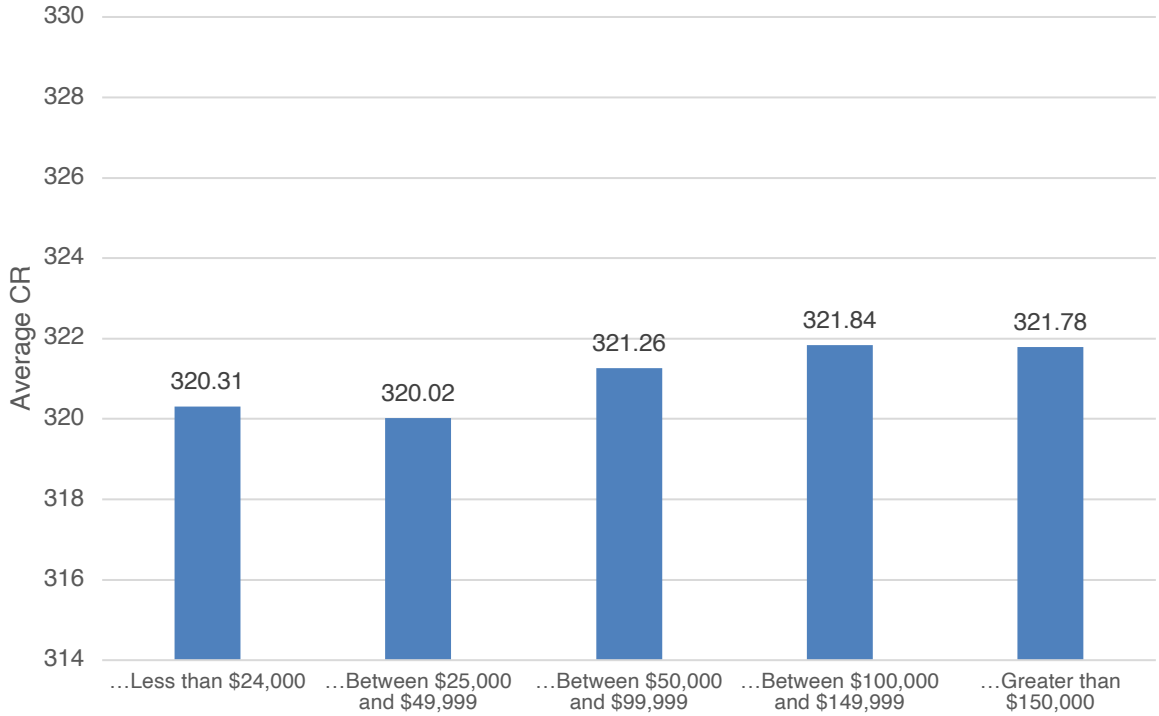
Figure 12 reveal a few trends in how the number of household members within certain ranges influences ER and CR. Based on these results, there is no noticeable trend in the increase and decrease of ER and CR as number of members in each category changes. However, ER and CR change together as number of members in each category changes. While ER and CR do not increase the more members of a certain age range are in each household, they increase and decrease at the same rates. However, ER and CR greatly increase the number of children aged 0-5 is at least 4. Also, CR increases as the number of adults increases.

Figure 13: ER and Income Ranges Comparisons

(a) ER



(b) CR

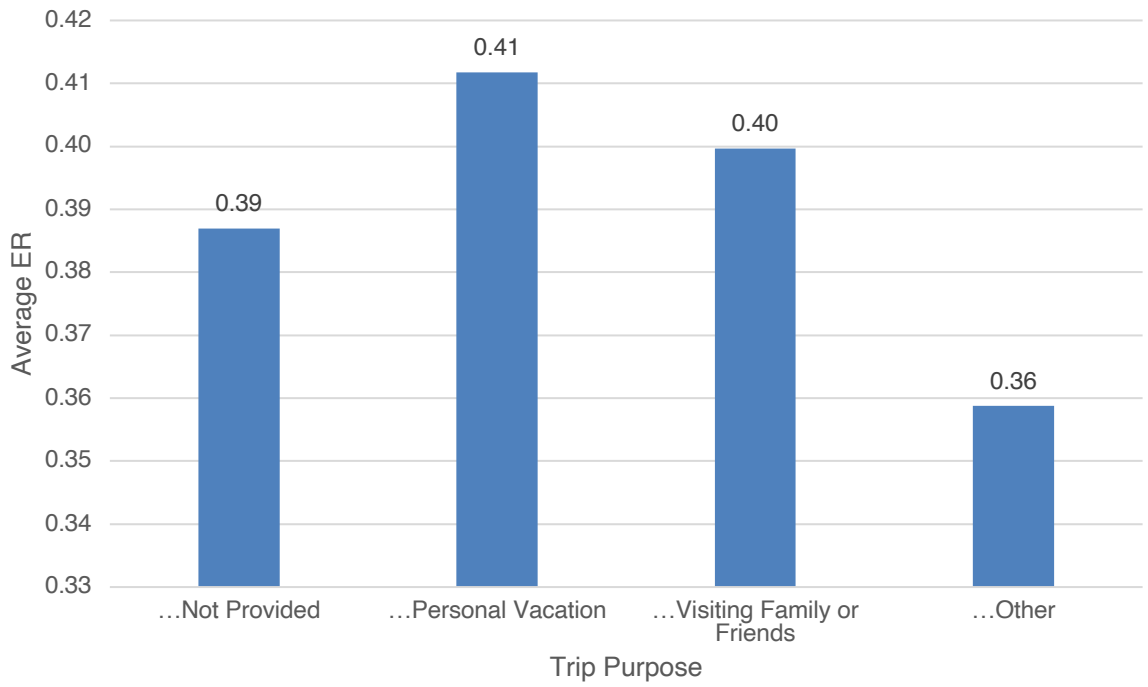


Based on the results found in Figure 13, a few conclusions can be drawn. First, both ER and CR increase as income increases. This increase, however, is not a significant increase for either variable. This means that, while higher income earners are more likely to travel further and endure longer travel times, those distances and times are not significantly longer than those with lower incomes.

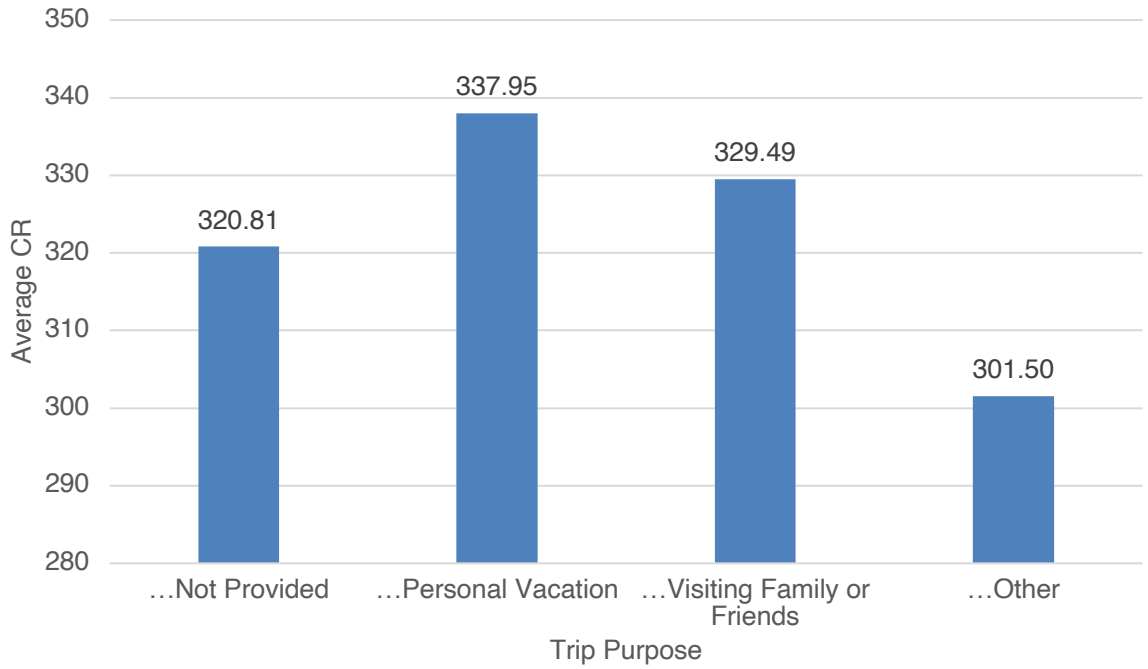
Next, this study measured the average ER and CR for varying trip purposes. For this study, business trips were not considered and aggregated out of the dataset. However, leisure trips and trips where the respondent visited family or friends were kept. Also, trips that did not have a specific purpose were kept as well, despite not fitting either leisure trip category. This was done since the specific trip reason is not known and therefore cannot be eliminated from consideration. The results of these trends can be found in Figure 14.

Figure 14: Varying Trip Purposes Comparisons

(a) ER



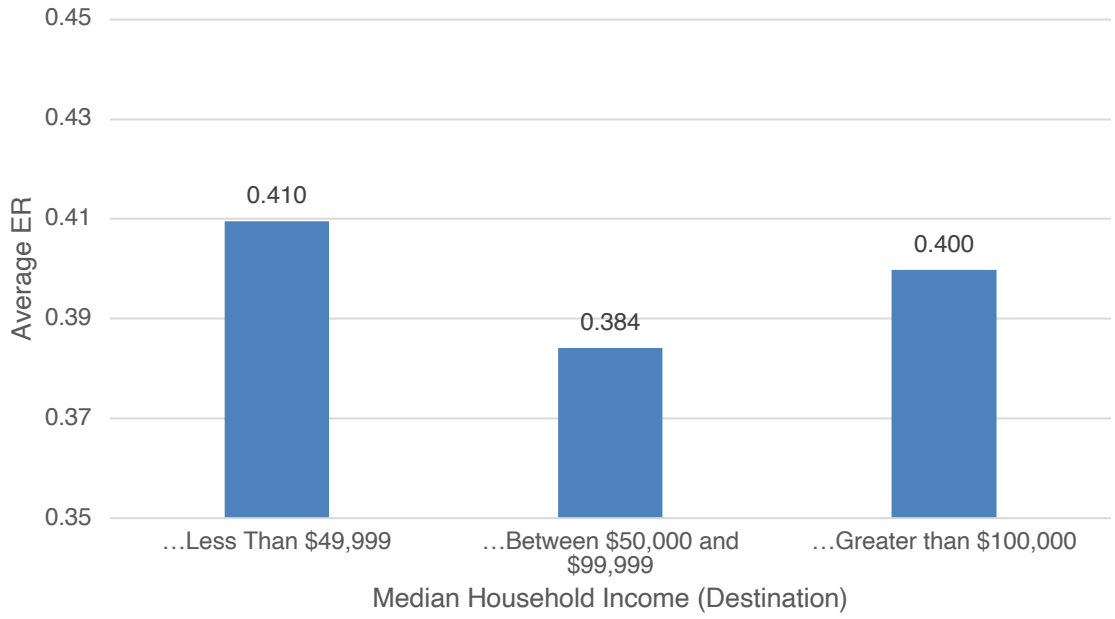
(b) CR



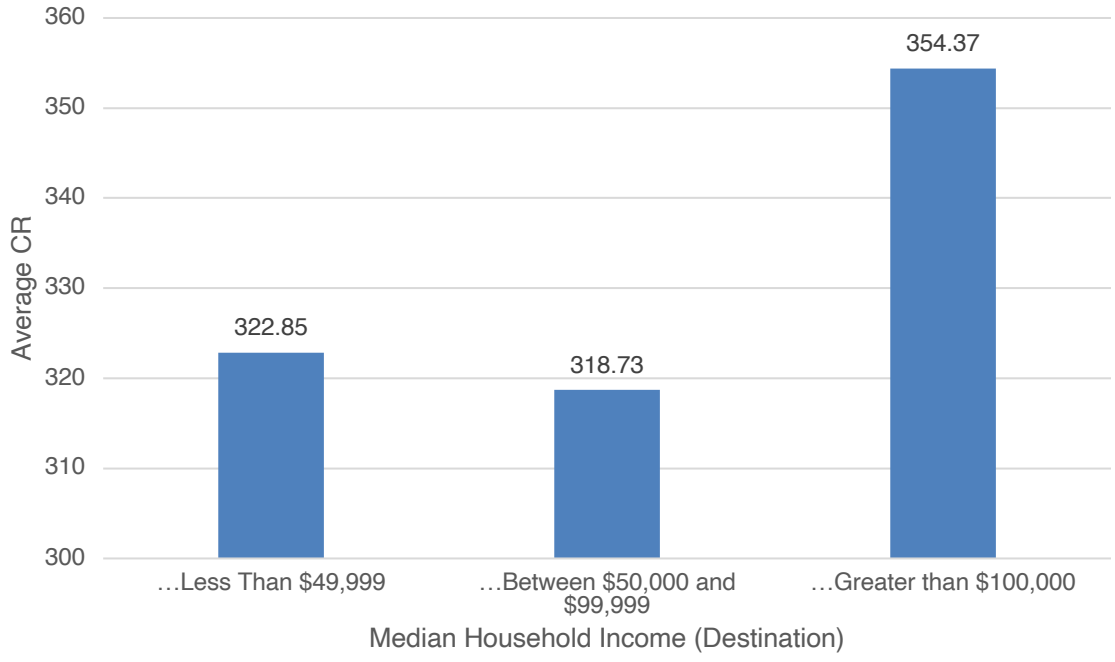
These results reflect different trends than income increase and decrease. Based on these results, trips that were classified as “personal vacations” had the highest ER and CR, meaning that travelers taking a personal vacation are more likely to travel further and endure longer travel times. The next highest ER and CR were trips aimed to visit family and friends. While these trips were lower than personal vacations, they were still higher than trips where the purpose was not provided. Future versions of the NHTS and this work should explore ER for business travel and provide more trip purposes for study. Next, ER and CR were compared to the Median Household Income of the traveler’s destination. This variable was chosen since, like income and trip purpose, was significant in some capacity in all three regressions. The results of this comparison can be found in Figure 15.

Figure 15: Median Household Income at Destination Comparisons

(a) ER



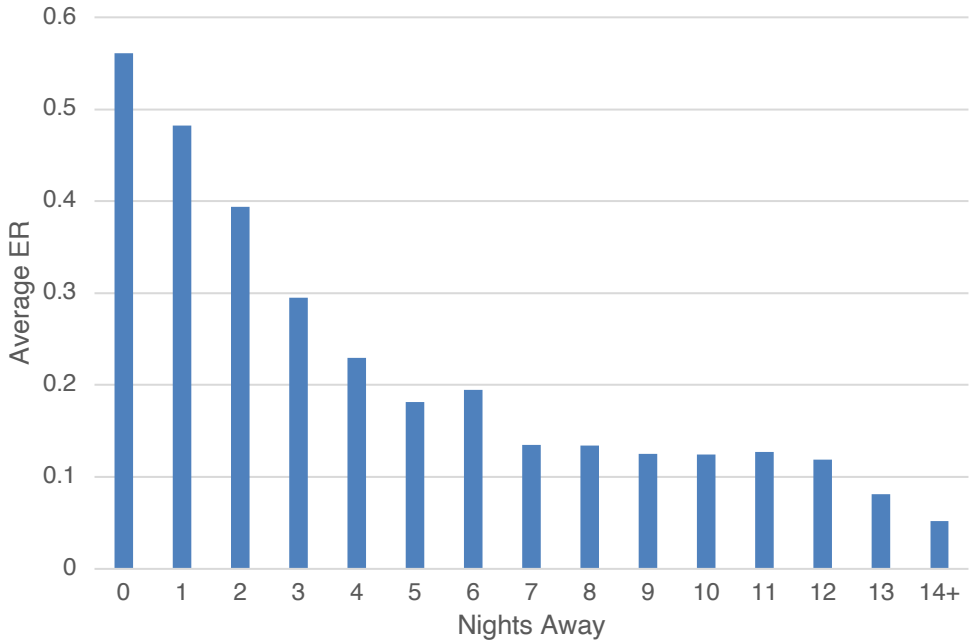
(b) CR



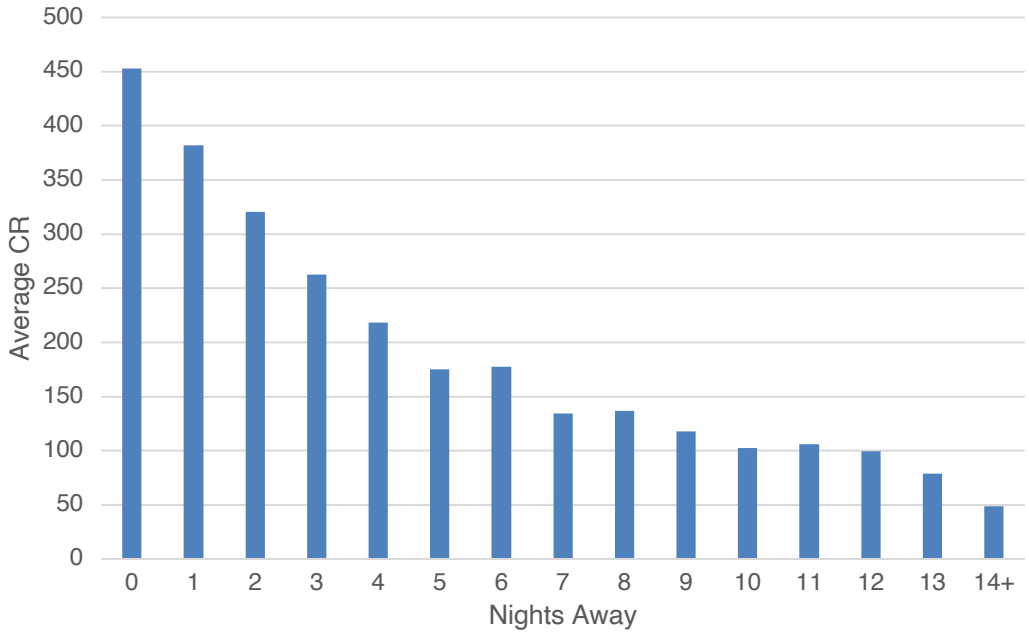
The results of Figure 15 show the differences in how ER and CR change as the median household income of the destination grows. First, travelers are willing to travel much further given the same time constraints to areas with higher median household incomes on average. Conversely, the CR for locations with lower median household incomes is far lower than for higher range median household income locations, meaning higher local incomes produce much higher CR values for those travelers visiting there. The same trends are not quite the same for ER values, however. Median household income is less of a factor in significantly increasing the ER of a trip due to the smaller range of average ER values for varying median household income ranges. As mentioned before, this variable was selected for this analysis since it was significant in all three models. However, it was also selected since it encapsulates another trend later observed in the regressions, which is that the economic well-being of the county being travelled to is an important factor in determining ER, CR and SR. The most prominent variable in this was the median household income at the destination location. Also, it was the only variable that was significant in all three models for This is the main reason why this variable was selected out of all economic characteristics. The final variable selected for this portion of analysis was the number of nights away, which was another significant variable in all three models. The results of this comparison can be found below in Figure 16.

Figure 16: Nights Away Comparisons

(a) ER



(b) CR



These results indicate that as the number of Nights Away increases, the ER and CR decrease. This does not mean, however, that distance or travel time increase or decrease as nights away increases. These findings make sense, since the number of nights away was a partial component in calculating ER and CR. To calculate ER, the number of nights away was converted to number of hours away. To calculate the CR, the number of nights away was converted to number of days away.

6.2 Gamma Regressions

The next step in analysis was the gamma regressions and summary statistics for each model. The summary statistics can be found in Table 4. To determine the effectiveness of a gamma distribution model, three parameters need to be determined: Log-likelihood, likelihood ratio chi-squared and the Significance. The log-likelihood of a gamma distribution describes the overall fit of the model compared to a typical gamma distribution. The likelihood ratio chi-square test determines the improvement of the model compared to the null model. Finally, the significance (or p-value) of the model determines how much the hypothesis and null hypothesis correlate. A lower p-value indicates that there is strong evidence to reject the null hypothesis. Based on this, these three values were determined for the ER, CR, and SR models. They can be found below in Table 5.

Table 4: Summary Statistics for ER, CR, and SR

Dependent Variable Statistics				
Variable	Minimum	Maximum	Mean	Standard Deviation
ER	0.000	0.999	0.387	0.278
CR	0.000	2392.059	320.966	225.860
SR	0.000	2698.964	827.222	598.267

Table 5: Gamma Distribution Results Statistics

Statistics for Gamma Distribution Models			
Variable	Log-Likelihood	Likelihood Ratio Chi Squared	Significance
ER	305.959	692.666	< 0.001
CR	-9029.882	594.624	< 0.001
SR	-10334.588	431.196	< 0.001

These results provide insight into the fit and details of each model. The goal is to have the log-likelihood be as close to zero as possible, with the ER regression being the best fitting and the SR regression with the largest log-likelihood. All three values are reasonable and acceptable for our estimations. The second test performed on all three models is the likelihood ratio chi-squared test. These statistics evaluate if each model provides more statistical prediction than a constants-only model. All three models have a chi-squared value greater than a critical value at any confidence level, indicating that they are all provide strong predictions.

As mentioned before, gamma regressions are used when the distribution of the dependent variable fits a gamma distribution. Based on the log-likelihood, likelihood ratio chi-squared and p-value of each test, a gamma regression was the best fit. For all three regressions, a confidence interval of 95%, and an alpha value of 0.05 were used for this study. The first regression analysis conducted was for ER.

6.3 Efficiency Ratio Gamma Regression

The first of three regressions performed was for the Efficiency Ratio of the NHTS trips. In total, 86 variables were used for this regression. Each variable used in the regression was from one of the five main categories of data combined in the dataset: personal, household, origin/home location, trip/tour, and destination characteristics. Also, variables pertaining to distance to and from airports, hours and days away, and time spent travelling were not included in the regression models since they were used to calculate ER. The results of the regression can be found in Table 6.

Based on the respondent section of the ER regression, a few conclusions can be drawn. First, the age of the respondent does not influence ER, despite the literature determining that trip length and time at destination were both influenced by age. Additionally, the literature argues that time constraints for the respondent such as school or employment do not influence ER, trip length or travel time. This is true despite findings in the literature that indicate that time constraints such as commitments back home influence frequency of trips taken and how long the respondent can be gone for. However, one trend that is consistent with the literature is the influence of higher incomes on ER. These findings indicate that having a higher income (specifically those who earn at least $100,000 per year) is more influential to ER changes than lower income levels, a common finding in the literature.

The literature also points out that variables such as demographics may affect the travel times and time at destination some, which was not found to be true in this work. These insignificant variables include household size, the number of children between

the ages of zero and five and five and 17, along with other variables such as number of drivers and number of vehicles in the household. Based on these findings, future work should be done to determine the significant variables pertaining to the respondent's household that influence ER.

These results reveal a few trends that influence and do not influence ER at the origin and home location level, including entertainment at home, industry specific characteristics of the home location, among others. In this model, the only variables that pertained to the home location and were significant related to the types of industries present in that location. The most unsurprising was the presence of entertainment options at the origin location. The model further indicates that home locations with more entertainment and novelty locations tend to lower the ER of the traveler. This means that travelers who live at these types of locations spend less time travelling to their destinations or stay longer on their trip when they take one. While these two cannot be true at the same time, the presence of entertainment at the home destination is a crucial variable for determining ER. Another finding in this data is that the type of area (urban or rural) that the respondent lives in does not influence ER. This is a finding that counters a finding in the literature, which indicates that those living in urban areas are more inclined to travel further due to access to the nearest airport.

There were many variables pertaining to trip characteristics that were significant in influencing ER. Specifically, trips that had a defined purpose were the most influential in determining ER. Furthermore, these types of trips had lower ERs. Since these two trip types are both considered leisure trips, future trips should (a) compare the ER of these trips to trips taken for business and (b) explore more specific types of leisure

trips (or more specific activities done while on the trip that may increase or decrease ER). Another variable that decreases ER is nights away. While this variable was not used directly in the ER calculation, this influence was to be expected since hours spent away, and nights away somewhat align. Also, trips that took place over a weekend had lower ER values than trips that did not. This is an interesting finding, since variables discussed earlier such as the respondent being in school or being employed not being influential. Therefore, other time constraints should be explored in the future. Finally, as echoed in the household characteristics, party size for the trip does not influence ER, which is a finding consistent in the literature. The final subset of data analyzed in the ER regression was the destination characteristics. These variables indicated that a few cost-of-living variables including average home rent and the median household income of the location were the most likely to influence ER. Furthermore, destinations with higher costs of living were the most likely to increase ER. This means that individuals are likely to travel further to places with higher costs of living. Despite this finding, places with more entertainment are likely to increase ER, but are not significant enough to influence it greatly. These findings collectively illustrated the significant and insignificant factors in ER calculation.

Table 6: Efficiency Ratio Regression Results

Variable and Variable Range	β	p-value	Significant?
Respondent Characteristics			
Intercept	0.124	0.784	
Respondent age	0.002	0.154	NO
Sex of Respondent (Base: Female)			
…Male	-0.016	0.663	NO
Respondent Race (Base: Other, Non- Hispanic)			
…Hispanic, No Other Races	-0.028	0.898	NO
…White, Non-Hispanic	-0.191	0.365	NO
…Black/African American, Non-Hispanic	-0.206	0.363	NO
…Asian, Non-Hispanic	-0.103	0.645	NO
Education (Base: Less than High School)			
…Some High School or Trade School	-0.061	0.487	NO
…Undergraduate Degree	-0.048	0.591	NO
…Graduate Degree	-0.122	0.205	NO
Income (Base: Less than $24,000)			
…Between $25,000 and $49,999	-0.215	0.066	NO
…Between $50,000 and $99,999	-0.209	0.057	NO
…Between $100,000 and $149,999	-0.271	0.019	YES
…Greater than $150,000	-0.276	0.020	YES
Home Ownership Status (Base: Rented or Occupied without Payment)			
…Owned	-0.002	0.974	NO
Type of Home Respondent Lives In (Base: Other)			
…Single Family	-0.254	0.059	NO
…Apartment Building	-0.125	0.407	NO
Travel Is Difficult due to Medical Condition	-0.036	0.736	NO
Respondent is Employed (Base: Yes)			
…No	-0.001	0.980	NO
Respondent is in School	-0.022	0.857	NO
Number of Long Distance Trips in last 30 Days	0.001	0.571	NO

Table 6 Continued: Efficiency Ratio Regression Results

Variable and Variable Range	β	p-value	Significant?
Household Characteristics			
Total number of people in household	-0.027	0.611	NO
Count of household members under 5 years old	0.039	0.612	NO
Count of household members 5-17 years old	0.039	0.470	NO
Total number of vehicles in household	-0.005	0.813	NO
Number of drivers in the household	0.083	0.152	NO
Number of times goods delivered in past 30 days	0.002	0.799	NO
Number of Deliveries to Residence in Last 30 Days	-0.002	0.774	NO
Origin/Home Location Characteristics			
Total Population (100,000s of People)	0.016	0.466	NO
Average Yearly Temperature (°F)	-0.006	0.140	NO
Average Yearly Precipitation (inches)	-0.003	0.860	NO
GDP of County (in Millions of $)	-0.002	0.231	NO
Mean Home Rent (in Thousands of $)	-0.201	0.129	NO
Mean Home Value (in Thousands of $1,000)	0.042	0.251	NO
Median Household Income ($1,000)	0.001	0.585	NO
Number of Establishments (1,000s of Establishments)			
...All Sectors	0.000	0.891	NO
...Accomodation	-0.010	0.933	NO
...Administration	0.037	0.814	NO
...Agriculture	1.811	0.121	NO
...Construction	0.031	0.642	NO
...Education	-0.910	0.063	NO
...Entertainment	-0.414	0.015	YES
...Finance	0.074	0.585	NO
...Health	-0.077	0.310	NO
...Information	0.885	0.011	YES
...Management	-0.634	0.105	NO

Table 6 Continued: Efficiency Ratio Regression Results

Variable and Variable Range	β	p-value	Significant?
...Manufacturing	-0.021	0.863	NO
...Mining	0.470	0.174	NO
...Professional	0.017	0.764	NO
...Real Estate	0.189	0.034	YES
...Retail	-0.051	0.530	NO
...Transportation	0.073	0.208	NO
...Utility	0.117	0.909	NO
...Whosesale	-0.053	0.483	NO
Urban/Rural Indicator (Base: Urban)			
…Rural	-0.010	0.896	NO
…Small Town	0.002	0.973	NO
…Second City	-0.086	0.262	NO
…Suburban	-0.002	0.978	NO
Trip/Tour Characteristics			
Trip Purpose (Base: Not Provided/Other)			
…Personal Vacation	-0.196	0.001	YES
…Visiting Family or Friends	-0.151	0.004	YES
Number of people with respondent on long distance trip	0.006	0.724	NO
Nights away on long distance trip	-0.130	0.000	YES
Month the Trip was Taken In (Base: December)			
…January	-0.003	0.973	NO
…February	-0.057	0.554	NO
…March	0.047	0.651	NO
…April	0.107	0.264	NO
…May	0.058	0.551	NO
…June	0.103	0.336	NO
…July	-0.029	0.757	NO
…August	0.013	0.904	NO
…September	0.095	0.336	NO
…October	-0.072	0.436	NO
…November	-0.039	0.601	NO
Trip Took Place Over Weekend (Base: No)			
…Yes	-0.081	0.049	YES

Table 6 Continued: Efficiency Ratio Regression Results

Trip Destination Characteristics			
Variable and Variable Range	β	p-value	Significant?
GDP of County ($1,000,000)	0.000	0.721	NO
Mean Home Rent ($1,000)	-0.308	0.000	YES
Mean Home Value ($1,000)	0.040	0.120	NO
Median Household Income ($1,000)	0.004	0.013	YES
Number of Establishments (1,000s of Establishments)			
…All Sectors	0.000	0.544	NO
…Accommodation	0.133	0.085	NO
…Entertainment	0.018	0.519	NO
…Health	-0.157	0.001	YES
…Retail	0.039	0.439	NO
Average Yearly Precipitation (inches)	-0.008	0.598	NO
Average Yearly Temperature (°F)	0.005	0.135	NO
Total Population (100,000s of People)	0.013	0.343	NO
Scale	0.434		

6.4 Coverage Rate Gamma Regression

The next regression performed was for Coverage Rates. This regression used the same variable set as the Efficiency Rate regression. The regression for CR yielded similar results to the ER regression, but a few noticeable differences can be observed. First, there were no variables in the respondent or household characteristic subsets that were significant in determining CR. The most influential variable in both datasets was the time commitment to school or academic programs (even though this variable was not considered significant). While the ER model had very few variables that were insignificant in both subsets, income was significant in ER model, more specifically higher income. Despite this, income was not a significant variable in the CR model. Also, the presence of mining industries was a significant variable in both models. However, despite the information and entertainment sectors being significant in the ER model, they were not in the CR model. Instead, the presence of retail and transportation industries were both positive influences on CR. This indicates that individuals who live in areas with a higher presence of these two industries are more likely to travel further given a certain number of days to be away from home. Finally, both ER and CR were influenced by the same variables pertaining to the trip itself, including trip purpose, number of nights away, and if the trip occurred on a weekend.

The final subset from the CR regression pertained to the destination county. Unlike the ER model, many variables were significant in this area. One common finding in the literature was that tropical destinations (or warmer) were the most popular destination types for American long-distance trips. While this study does not contain a variable pertaining to if the destination was “tropical” or not, it does contain information pertaining

to the yearly temperatures of the destination. Based on these results, trips that have destinations that have higher yearly temperatures are also more likely to have higher CRs. This also indicates that trips to states such as Florida, Arizona, and California have visitors from much further away compared to destinations that are colder. This trend, however, is not true for destinations with higher average rainfall. These destinations, unlike tropical destinations, tend to deter visitors. Furthermore, visitors will also travel shorter SRs to reach these locations. This model also explores other tourism infrastructure such as presence of entertainment and accommodations. Much like the ER model, presence of entertainment does not significantly change the CR of an individual. However, trips to areas with more accommodations (or higher presence of hotels) are more likely to have higher CRs, a trend deemed insignificant in the ER model. A trend observed in both models, however, is the cost-of-living in the destination counties. Just like the ER model, destinations with higher costs-of-living have higher CRs. Therefore, it can be concluded that travelers are willing to travel longer and further to areas with higher costs of living.

These results closely mirror those in the ER model besides how each variable was influenced by the destination of the trip. Additionally, the economic status and cost of living variables are not noteworthy in the CR model besides the median household income in the county. Future work for both ER and CR should incorporate more economic variables such as cost of living and average price of gasoline, among other variables, in the models. These numbers were not available at the time this work was done, so they were not incorporated into this study. These findings concluded the CR study. A final regression was performed, this time for travel SR. While SR is a component of CR, a

regression was performed to observe if correlation existed between ER/CR and SR, or if these trends were isolated.

Table 7: Coverage Rate Regression Results

Variable and Variable Range	β	p-value	Significant?
Respondent Characteristics			
Intercept	6.335	0.000	
Respondent age	0.002	0.284	NO
Sex of Respondent (Base: Female)			
…Male	-0.028	0.440	NO
Respondent Race (Base: Other, Non- Hispanic)			
…Hispanic, No Other Races	0.055	0.796	NO
…White, Non-Hispanic	-0.022	0.913	NO
…Black/African American, Non-Hispanic	-0.016	0.941	NO
…Asian, Non-Hispanic	-0.076	0.725	NO
Education (Base: Less than High School)			
…Some High School or Trade School	-0.071	0.407	NO
…Undergraduate Degree	-0.047	0.591	NO
…Graduate Degree	-0.063	0.500	NO
Income (Base: Less than $24,000)			
…Between $25,000 and $49,999	-0.103	0.363	NO
…Between $50,000 and $99,999	-0.122	0.250	NO
…Between $100,000 and $149,999	-0.135	0.222	NO
…Greater than $150,000	-0.117	0.308	NO
Home Ownership Status (Base: Rented or Occupied without Payment)			
…Owned	-0.003	0.961	NO
Type of Home Respondent Lives In (Base: Other)			
…Single Family	-0.083	0.528	NO
…Apartment Building	0.000	1.000	NO
Travel Is Difficult due to Medical Condition	0.018	0.864	NO
Respondent is Employed (Base: Yes)			
…No	0.012	0.799	NO
Respondent is in School	0.181	0.123	NO
Number of Long Distance Trips in Last 30 Days	0.001	0.354	NO

Table 7 (Continued): Coverage Rate Regression Results

Variable and Variable Range	β	p-value	Significant?
Household Characteristics			
Total number of people in household	-0.007	0.891	NO
Count of household members under 5 years old	0.048	0.524	NO
Count of household members 5-17 years old	0.034	0.517	NO
Total number of vehicles in household	-0.009	0.692	NO
Number of drivers in the household	0.028	0.623	NO
Number of times goods delivered in past 30 days	0.002	0.774	NO
Number of Deliveries to Residence in Last 30 Days	-0.003	0.711	NO
Origin/Home Location Characteristics			
Total Population (100,000s of People)	0.002	0.918	NO
Average Yearly Temperature (°F)	-0.007	0.087	NO
Average Yearly Precipitation (inches)	0.024	0.126	NO
GDP of County (in Millions of $)	-0.002	0.195	NO
Mean Home Rent (in Thousands of $)	-0.156	0.236	NO
Mean Home Value (in Thousands of $1,000)	0.049	0.180	NO
Median Household Income ($1,000)	0.002	0.362	NO
Number of Establishments (1,000s of Establishments)			
...All Sectors	0.000	0.526	NO
...Accomodation	-0.124	0.274	NO
...Administration	0.144	0.347	NO
...Agriculture	1.080	0.344	NO
...Construction	-0.004	0.952	NO
...Education	-0.537	0.263	NO
...Entertainment	-0.310	0.062	NO
...Finance	-0.073	0.583	NO
...Health	-0.050	0.495	NO
...Information	0.659	0.051	NO
...Management	-0.531	0.172	NO

Table 7 (Continued): Coverage Rate Regression Results

Variable and Variable Range	β	p-value	Significant?
...Manufacturing	0.105	0.384	NO
...Mining	0.900	0.008	YES
...Professional	-0.006	0.910	NO
...Real Estate	0.285	0.001	YES
...Retail	-0.055	0.487	NO
...Transportation	0.147	0.009	YES
...Utility	-0.379	0.704	NO
...Whosesale	-0.070	0.340	NO
Urban/Rural Indicator (Base: Urban)			
…Rural	-0.090	0.250	NO
…Small Town	-0.072	0.302	NO
…Second City	-0.129	0.084	NO
…Suburban	-0.059	0.392	NO
Trip/Tour Characteristics			
Trip Purpose (Base: Not Provided/Other)			
…Personal Vacation	-0.156	0.005	YES
…Visiting Family or Friends	-0.113	0.026	YES
Number of people with respondent on long distance trip	-0.017	0.311	NO
Nights away on long distance trip	-0.120	0.000	YES
Month the Trip was Taken In (Base: December)			
…January	0.009	0.924	NO
…February	-0.091	0.328	NO
…March	0.084	0.404	NO
…April	0.058	0.530	NO
…May	-0.021	0.823	NO
…June	0.079	0.452	NO
…July	-0.059	0.510	NO
…August	0.002	0.982	NO
…September	0.042	0.661	NO
…October	-0.048	0.596	NO
…November	-0.002	0.982	NO
Trip Took Place Over Weekend (Base: No)			
…Yes	-0.102	0.012	YES

Table 7 (Continued): Coverage Rate Regression Results

Trip Destination Characteristics			
Variable and Variable Range	β	p-value	Significant?
GDP of County ($1,000,000)	-0.002	0.158	NO
Mean Home Rent ($1,000)	-0.302	0.000	YES
Mean Home Value ($1,000)	0.065	0.008	YES
Median Household Income ($1,000)	0.003	0.055	NO
Number of Establishments (1,000s of Establishments)			
…All Sectors	0.000	0.637	NO
…Accomodation	0.171	0.024	YES
…Entertainment	-0.020	0.448	NO
…Health	-0.117	0.013	YES
…Retail	0.075	0.118	NO
Average Yearly Precipitation (inches)	-0.036	0.022	YES
Average Yearly Temperature (°F)	0.007	0.025	YES
Total Population (100,000s of People)	-0.005	0.696	NO
Scale	0.415		

6.5 Search Radius Gamma Regression

The final regression analysis conducted was for the search radii. The same data set was used for this analysis as was used in the ER and CR analyses. However, the SR variable itself was used in the ER and CR analyses as an independent variable but was relegated to being the dependent variable for this regression. Based on these results, a few conclusions can be made. First, these results mirror many results found in the ER and CR models. For instance, income was a significant predictor of travel SR. However, the SR model shows that all income ranges were significant in predicting travel SR. Like the results of the ER and CR models, there were no variables pertaining to household characteristics that were significant in determining ER. This means that, for future work, new variables pertaining to the characteristics of the household should be used in this study. Like household characteristics, the characteristics of the origin location, specifically the presence of certain industries. For that model however, the presence of the construction and education sector was significant unlike ER and CR. Finally, the same trip characteristics such as nights away, weekend trips and trip purpose were significant in the SR model. However, trips that were taken to visit family or friends had less of an effect on the model than they did for ER and CR. If the confidence interval was lower for this model, such as 90%, then it would have been significant.

The final subsection of variables used in the analysis for SR pertained to the destination of the trip. The results of this subsection mirror the results of the ER and CR models, indicating that destinations that have higher costs-of-living are visited by people who live further SRs away. These results mirror those of the ER and CR model, indicating that future research should study the socioeconomic, county level economic and built

environment variables to compare to ER, CR and travel SR. Finally, the presence of employment options in the destination county also yields higher travel SR, indicating that economic opportunity and business opportunity positively influence travel SR, furthering the emphasis for study into economic factors at destinations and how they influence ER, CR and travel SR. Additionally, the presence of accommodations positively influences the SR traveled for a trip. Also, the SR model echoes the findings in the literature that say that travelers are willing to travel further to destinations that are warmer year-round and have lower levels of precipitation.

Table 8: Search Radius Regression Results

Variable and Variable Range	β	p-value	Significant?
Respondent Characteristics			
Intercept	6.067	0.000	
Respondent age	0.002	0.151	NO
Sex of Respondent (Base: Female)			
…Male	-0.033	0.372	NO
Respondent Race (Base: Other, Non- Hispanic)			
…Hispanic, No Other Races	-0.013	0.954	NO
…White, Non-Hispanic	-0.045	0.830	NO
…Black/African American, Non-Hispanic	0.000	0.998	NO
…Asian, Non-Hispanic	-0.117	0.596	NO
Education (Base: Less than High School)			
…Some High School or Trade School	-0.040	0.641	NO
…Undergraduate Degree	0.001	0.987	NO
…Graduate Degree	-0.030	0.755	NO
Income (Base: Less than $24,000)			
…Between $25,000 and $49,999	-0.243	0.038	YES
…Between $50,000 and $99,999	-0.295	0.008	YES
…Between $100,000 and $149,999	-0.246	0.032	YES
…Greater than $150,000	-0.239	0.044	YES
Home Ownership Status (Base: Rented or Occupied without Payment)			
…Owned	0.099	0.123	NO
Type of Home Respondent Lives In (Base: Other)			
…Single Family	-0.063	0.637	NO
…Apartment Building	0.047	0.752	NO
Travel Is Difficult due to Medical Condition	-0.027	0.799	NO
Respondent is Employed (Base: Yes)			
…No	-0.034	0.472	NO
Respondent is in School	0.105	0.382	NO
Number of Long-Distance Trips in Last 30 Days	-0.001	0.501	NO

Table 8 (Contined): Search Radius Regression Results

Variable and Variable Range	β	p-value	Significant?
Household Characteristics			
Total number of people in household	-0.040	0.445	NO
Count of household members under 5 years old	0.043	0.571	NO
Count of household members 5-17 years old	0.052	0.323	NO
Total number of vehicles in household	0.005	0.843	NO
Number of drivers in the household	0.011	0.844	NO
Number of times goods delivered in past 30 days	-0.002	0.791	NO
Number of Deliveries to Residence in Last 30 Days	0.001	0.933	NO
Origin/Home Location Characteristics			
Total Population (100,000s of People)	-0.023	0.304	NO
Average Yearly Temperature (°F)	-0.004	0.329	NO
Average Yearly Precipitation (inches)	0.003	0.869	NO
GDP of County (in Millions of $)	-0.001	0.428	NO
Mean Home Rent (in Thousands of $)	-0.140	0.300	NO
Mean Home Value (in Thousands of $1,000)	0.073	0.052	NO
Median Household Income ($1,000)	0.000	0.913	NO
Number of Establishments (1,000s of Establishments)			
...All Sectors	0.001	0.168	NO
...Accomodation	-0.168	0.153	NO
...Administration	-0.163	0.299	NO
...Agriculture	0.730	0.533	NO
...Construction	0.171	0.009	YES
...Education	-1.353	0.006	YES
...Entertainment	-0.181	0.297	NO
...Finance	-0.185	0.172	NO
...Health	0.093	0.221	NO
...Information	0.464	0.187	NO
...Management	-0.186	0.652	NO

Table 8 (Continued): Search Radius Regression Results

Variable and Variable Range	β	p-value	Significant?
...Manufacturing	0.047	0.701	NO
...Mining	0.710	0.044	YES
...Professional	0.037	0.528	NO
...Real Estate	0.287	0.001	YES
...Retail	0.006	0.941	NO
...Transportation	0.156	0.008	YES
...Utility	0.388	0.703	NO
...Whosesale	-0.184	0.017	YES
Urban/Rural Indicator (Base: Urban)			
…Rural	-0.118	0.140	NO
…Small Town	-0.100	0.163	NO
…Second City	-0.079	0.303	NO
…Suburban	-0.036	0.609	NO
Trip/Tour Characteristics			
Trip Purpose (Base: Not Provided/Other)			
…Personal Vacation	0.117	0.043	YES
…Visiting Family or Friends	0.101	0.056	NO
Number of people with respondent on long distance trip	0.007	0.670	NO
Nights away on long distance trip	0.054	0.000	YES
Month the Trip was Taken In (Base: December)			
…January	0.017	0.863	NO
…February	-0.134	0.159	NO
…March	0.134	0.193	NO
…April	0.096	0.311	NO
…May	-0.043	0.658	NO
…June	0.155	0.151	NO
…July	-0.077	0.408	NO
…August	0.114	0.303	NO
…September	-0.017	0.864	NO
…October	-0.021	0.817	NO
…November	0.000	1.000	NO
Trip Took Place Over Weekend (Base: No)			
…Yes	0.299	0.000	YES

Table 8 (Continued): Search Radius Regression Results

Trip Destination Characteristics			
Variable and Variable Range	β	p-value	Significant?
GDP of County ($1,000,000)	-0.003	0.006	YES
Mean Home Rent ($1,000)	-0.365	0.000	YES
Mean Home Value ($1,000)	0.099	0.000	YES
Median Household Income ($1,000)	0.005	0.009	YES
Number of Establishments (1,000s of Establishments)			
…All Sectors	0.000	0.937	NO
…Accomodation	0.238	0.002	YES
…Entertainment	-0.033	0.232	NO
…Health	-0.065	0.181	NO
…Retail	0.058	0.241	NO
Average Yearly Precipitation (inches)	-0.045	0.005	YES
Average Yearly Temperature (°F)	0.012	0.000	YES
Total Population (100,000s of People)	-0.024	0.085	NO
Scale	0.433		

Chapter 7: Conclusions

This study aimed to determine if a reasonable Efficiency Ratio, Coverage Rate, and Search Radius could be determined for trip entries in the 2024 NHTS. This study also aimed to determine the most influential factors associated with ER, CR, and SR.

To compute ER, CR, and Travel Distance, several pieces of data needed to be obtained. First, data from the National Household Travel Survey was collected using the person, household, and long-distance subsections of the dataset. These three datasets contain data regarding the characteristics of the household (such as age ranges of household members, type of home, etc.), respondent (such as income, employment status, race, etc.) and long-distance trip taken by the travel party. Along with this data, all 4,148 trips in the dataset had an origin and destination county. This county was used for several purposes, including calculating the distance between the origin and destination in miles, determining the nearest airport to the origin/destination and the distance to it, and assigning each trip with the economic and labor characteristics of each county. This, coupled with the NHTS data, was a collection of data that covered a variety of topics related to the traveler, trip, and regions visited on the trip.

The results of this study proved that ER and CR values can be determined, and that measurable impacts of socioeconomic, labor, trip, and home and destination variables can be observed. . On average, ER was 0.13 for car trips and 0.42 for airplane trips. CR was around 283 on average for car trips and around 325 for airplane trips, meaning both values, along with travel distance, are much higher for airplane trips than for car trips. These findings are consistent with the findings in the literature, which indicate that trips that

cover longer distances (given the same time at destination regardless of mode) are more likely to be taken by airplane. Furthermore, certain variables were discovered and classified as accurate predictors of ER, CR and SR. The specific variables proven to be the most impactful were the economic status of the destination county, a defined trip purpose, a presence of accommodations (or places to stay) at the destination, higher income levels, and if the trip was taken over a weekend. These variables will be useful in future work pertaining to ER, CR and SR.

Understanding the relationship between travel time/time away from home, travel distance/time away from home and the search radii of NHTS travelers was performed for a few reasons. Currently, no research has been conducted that observe these variables together and how they increase/decrease with each other. Currently, planning agencies, local governments and other tourism related industries plan for tourists based on distance, nights away, and a few demographic variables such as age, race, sex, etc. These are used to approximate tourism volumes at a particular location and the types of places those travelers are from. Using the findings in this research, these agencies and businesses can:

- Have a more accurate representation of the distances travelers who visit certain destinations travel, the time they spend travelling there, and the mode they chose.
- Quantify how many nights these travelers are likely to stay based on the trip/travel characteristics.
- Have a better plan for marketing the destination to the travelers and types of travel parties most likely visit that destination.

- Prepare the destination and the surrounding areas for the types of travelers including access to amenities.

- Evaluate the existing infrastructure and economic state of the destination to quantify its appeal to a traveler.

Chapter 8: Future Work

This work can be used for many things in the future. First, this work quantifies the relationship between travel time/travel distances and time spent at the destination by the traveler. Since these trends have not been studied much in the literature, they can be used to further explain various travel trends such as destination choice, trip length, among other factors. This study faced a few limitations. First, the specific airports for air trips were not known, so the closest airports to the home and destination locations were used instead. These airports may or may not have been the actual chosen airports for the trip, but since this information was not provided in the NHTS, this was the closest information that could be used for trip time calculation and trip distance calculation. Furthermore, while the specific origin and destination counties were known and used heavily for this analysis, many assumptions regarding travel distance needed to be made, since county level data was the most exact data available. Future work in this field should examine the relationship between ER/CR and the specific variables used in this study. While this was done for a portion of the dataset's variables, it was not done for all of them, nor was it done for significant variables in the dataset. Additionally, future work should examine the types of novelties and activates that influence ER and CR the most. Since all three models underscore the impact that this has on ER and CR, future work should explore the types of entertainment that influence ER and CR, such as amusement parks, natural and historical sites, food related events and others. Finally, future work should observe the changes in ER and CR for different regions of the United States, such as states and

census regions, to determine which regions of the United States have higher ERs and CRs.

References

1. "U.S. Travel Economic Impact." *Ustravel.org*, 2023, impact.ustravel.org/national.

2. Hill, William. "Understanding the Typical Vacations of U.S. Southern Travelers." *Journal of Applied Business Research (JABR)*, vol. 32, no. 4, 30 June 2016, p. 1169, https://doi.org/10.19030/jabr.v32i4.9729. Accessed 3 May 2019.

3. LaMondia, Jeffrey, et al. "An Annual Time Use Model for Domestic Vacation Travel." *Journal of Choice Modelling*, vol. 3, no. 1, 29 Sept. 2007.

4. Kitamura, Ryuichi, et al. "Traveler Destination Choice Behavior: Effects of Time of Day, Activity Duration, and Home Location." *Transportation Research Record Journal of the Transportation Research Board*, vol. 1645, no. 1, 1 Jan. 1998, pp. 76–81, https://doi.org/10.3141/1645-10. Accessed 10 May 2025.

5. Bernini, Cristina, and Maria Francesca Cracolici. "Demographic Change, Tourism Expenditure and Life Cycle Behavior." *Tourism Management*, vol. 47, Apr. 2015, pp. 191–205, https://doi.org/10.1016/j.tourman.2014.09.016. Accessed 15 Apr. 2020.

6. Yang, Yang, et al. "A Shrinking World for Tourists? Examining the Changing Role of Distance Factors in Understanding Destination Choices." *Journal of Business Research*, vol. 92, Nov. 2018, pp. 350–359, https://doi.org/10.1016/j.jbusres.2018.08.001. Accessed 23 June 2021.

7. Rossi, Sebastian Dario, et al. "The Role of Distance in Peri-Urban National Park Use: Who Visits Them and How Far Do They Travel?" *Applied Geography*, vol. 63, 1 Sept. 2015, pp. 77–88, www.sciencedirect.com/science/article/pii/S0143622815001526, https://doi.org/10.1016/j.apgeog.2015.06.008. Accessed 22 Oct. 2021.

8. Wynen, Jan. "Explaining Travel Distance during Same-Day Visits." *Tourism Management*, vol. 36, June 2013, pp. 133–140, https://doi.org/10.1016/j.tourman.2012.11.007. Accessed 2 Sept. 2020.

9. Van Nostrand, Caleb, et al. “Analysis of Long-Distance Vacation Travel Demand in the United States: A Multiple Discrete–Continuous Choice Framework.” *Transportation*, vol. 40, no. 1, 5 Apr. 2012, pp. 151–171, https://doi.org/10.1007/s11116-012-9397-6. Accessed 19 Jan. 2021.

10. Varga, Levente, et al. “Further We Travel the Faster We Go.” *PLOS ONE*, vol. 11, no. 2, 10 Feb. 2016, p. e0148913, https://doi.org/10.1371/journal.pone.0148913. Accessed 16 Sept. 2019.

11. Aultman-Hall, Lisa. “Incorporating Long-Distance Travel into Transportation Planning in the United States.” *UC Davis Whitepapers*, 1 Oct. 2018. *California Digital Library*.

12. Small, Kenneth A. “Valuation of Travel Time.” *Economics of Transportation*, vol. 1, no. 1-2, Dec. 2012, pp. 2–14, https://doi.org/10.1016/j.ecotra.2012.09.002. Accessed 24 June 2020.

13. Koster, Paul, et al. “Travel Time Variability and Airport Accessibility.” *Transportation Research Part B: Methodological*, vol. 45, no. 10, Dec. 2011, pp. 1545–1559, https://doi.org/10.1016/j.trb.2011.05.027. Accessed 14 Dec. 2021.

14. Adler, Thomas, et al. “Modeling Service Trade-Offs in Air Itinerary Choices.” *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1915, no. 1, Jan. 2005, pp. 20–26, https://doi.org/10.1177/0361198105191500103. Accessed 19 May 2020.

15. Vany, Arthur De. “The Revealed Value of Time in Air Travel.” *The Review of Economics and Statistics*, vol. 56, no. 1, Feb. 1974, p. 77, https://doi.org/10.2307/1927529. Accessed 9 May 2020.

16. Gössling, Stefan, and Sara Dolnicar. “A Review of Air Travel Behavior and Climate Change.” *WIREs Climate Change*, vol. 14, no. 1, 18 Aug. 2022, https://doi.org/10.1002/wcc.802.

17. Gayle, Philip G., and Jules O. Yimga. “How Much Do Consumers Really Value Air Travel On-Time Performance, and to What Extent Are Airlines Motivated to

Improve Their On-Time Performance?" *Economics of Transportation*, vol. 14, June 2018, pp. 31–41, https://doi.org/10.1016/j.ecotra.2018.01.003. Accessed 9 May 2019.

18. Abdella, Juhar Ahmed, et al. "Airline Ticket Price and Demand Prediction: A Survey." *Journal of King Saud University - Computer and Information Sciences*, vol. 33, no. 4, Feb. 2019, pp. 375–391, www.sciencedirect.com/science/article/pii/S131915781830884X, https://doi.org/10.1016/j.jksuci.2019.02.001.

19. Thrane, Christer. "Examining Tourists' Long-Distance Transportation Mode Choices Using a Multinomial Logit Regression Model." *Tourism Management Perspectives*, vol. 15, July 2015, pp. 115–121, https://doi.org/10.1016/j.tmp.2014.10.004. Accessed 25 Apr. 2021.

20. Rasmidatta, Isaradatta. "Mode Choice Models for Long Distance Travel in United States of America - ProQuest." *Proquest.com*, 2026, www.proquest.com/docview/304902688?fromopenview=true&pq-origsite=gscholar&sourcetype=Dissertations%20&%20Theses. Accessed 19 Mar. 2026.

21. O'Neill, Wende A., et al. "Evaluating Role of Distance and Location in Statewide Travel Demand Forecasting by Using American Travel Survey." *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1660, no. 1, Jan. 1999, pp. 41–47, https://doi.org/10.3141/1660-06. Accessed 18 Oct. 2021.

22. Aultman-Hall, Lisa, et al. "Understanding the Relationship between Mode and Destination Choices for Personal Out-of-Town Travel." *Travel Behavior and Society*, vol. 31, Apr. 2023, pp. 399–409, https://doi.org/10.1016/j.tbs.2023.02.004. Accessed 16 Feb. 2023.

23. Moeckel, Rolf, et al. "Mode Choice Modeling for Long-Distance Travel." *Transportation Letters*, vol. 7, no. 1, Aug. 2014, pp. 35–46, https://doi.org/10.1179/1942787514y.0000000031.

24. Reichert, Alexander, and Christian HolzRau. “Mode Use in Long-Distance Travel.” *Journal of Transport and Land Use*, vol. 8, no. 2, 2015, pp. 87–105. *JSTOR*, www.jstor.org/stable/26202714, https://doi.org/10.2307/26202714.

25. Gudzinas, Brad. *Development of Long-Distance Multimodal Passenger Travel Modal Choice Model*. 23 July 2012, www.fhwa.dot.gov/policy/modalchoice/modalchoice.pdf.

26. “America on the Go…Long Distance Transportation Patterns: Mode Choice.” *Bts.gov*, May 2006, rosap.ntl.bts.gov/view/dot/6310, https://doi.org/10.21949/1501435. Accessed 19 Mar. 2026.

27. Andersson, Angelica, et al. “Long-Distance Mode Choice Model Estimation Using Mobile Phone Network Data.” *Journal of Choice Modelling*, vol. 42, Mar. 2022, p. 100337, https://doi.org/10.1016/j.jocm.2021.100337. Accessed 1 Apr. 2022.

28. Bursa, Bartosz. “Mode Choice and Value of Time on Long-Distance Vacation Trips.” *Transportation Research Record: Journal of the Transportation Research Board*, 24 July 2024, https://doi.org/10.1177/03611981241252837.

29. Chen, Patricia L. “TTB or Not TTB, That Is the Question: A Review and Analysis of the Empirical Literature on Travel Time (and Money) Budgets.” *Transportation Research Part A: Policy and Practice*, vol. 38, no. 9-10, 2 Feb. 2004, pp. 643–675, ideas.repec.org/a/eee/transa/v38y2004i9-10p643-675.html. Accessed 19 Mar. 2026.

30. Limtanakool, Narisra, et al. “The Influence of Socioeconomic Characteristics, Land Use and Travel Time Considerations on Mode Choice for Medium- and Longer-Distance Trips.” *Journal of Transport Geography*, vol. 14, no. 5, Sept. 2006, pp. 327–341, https://doi.org/10.1016/j.jtrangeo.2005.06.004. Accessed 12 July 2021.

31. Bhat, Chandra R. “Work Travel Mode Choice and Number of Non-Work Commute Stops.” *Transportation Research Part B: Methodological*, vol. 31, no. 1, Feb. 1997, pp. 41–54, https://doi.org/10.1016/s0191-2615(96)00016-1. Accessed 25 Nov. 2019.

32. Pendyala, R M, et al. “Application of an Activity-Based Travel-Demand Model Incorporating a Rule-Based Algorithm.” *Environment and Planning B: Planning and Design*, vol. 25, no. 5, 1998, pp. 753–772, https://doi.org/10.1068/b250753. Accessed 13 Sept. 2022.

33. Newman, Jeffrey P., and Vincent L. Bernardin. “Hierarchical Ordering of Nests in a Joint Mode and Destination Choice Model.” *Transportation*, vol. 37, no. 4, 17 Apr. 2010, pp. 677–688, https://doi.org/10.1007/s11116-010-9277-x. Accessed 11 Jan. 2021.

34. Dodds, Rachel, and Mark Robert Holmes. “Beach Tourists; What Factors Satisfy Them and Drive Them to Return.” *Ocean & Coastal Management*, vol. 168, Feb. 2019, pp. 158–166, www.sciencedirect.com/science/article/abs/pii/S0964569118305830, https://doi.org/10.1016/j.ocecoaman.2018.10.034.

35. Hasan, Md. Kamrul, et al. “Determining Factors of Tourists’ Loyalty to Beach Tourism Destinations: A Structural Model.” *Asia Pacific Journal of Marketing and Logistics*, vol. 32, no. 1, 22 July 2019, pp. 169–187, https://doi.org/10.1108/apjml-08-2018-0334. Accessed 26 Oct. 2020.

36. Goodrich, Jonathan N. “The Relationship between Preferences for and Perceptions of Vacation Destinations: Application of a Choice Model.” *Journal of Travel Research*, vol. 17, no. 2, Oct. 1978, pp. 8–13, https://doi.org/10.1177/004728757801700202.

37. Jarè Struwig, and Elizabeth du Preez. “The Effect of Income on the Relationship between Travel Motives and Destination Choices.” *South African Journal of Economic and Management Sciences*, vol. 27, no. 1, 2024, p. 13, sajems.org/index.php/sajems/article/view/5286/3032.

38. Megehee, Carol. “Top 25 Most Visited Tourist Destinations in America.” *ResearchGate*, 2026, www.researchgate.net/figure/Top-25-Most-Visited-Tourist-Destinations-in-America_tbl1_290504156. Accessed 19 Mar. 2026.

39. Nicolau, Juan L. “Characterizing Tourist Sensitivity to Distance.” *Journal of Travel Research*, vol. 47, no. 1, 14 Jan. 2008, pp. 43–52, https://doi.org/10.1177/0047287507312414.

40. Petroman, Cornelia. “Typology of Tourism Destinations.” *SCIENTIFIC PAPERS ANIMAL SCIENCE and BIOTECHNOLOGIES*, vol. 48, no. 1, 2015, pp. 338–338, mail.spasb.ro/index.php/public_html/article/view/290. Accessed 16 Sept. 2025.

41. Gitelson, Richard J., and John L. Crompton. “Insights into the Repeat Vacation Phenomenon.” *Annals of Tourism Research*, vol. 11, no. 2, Jan. 1984, pp. 199–217, https://doi.org/10.1016/0160-7383(84)90070-7. Accessed 5 Nov. 2019.

42. Hong, JungHwa (Jenny), and Kalpesh Kaushik Desai. “Variety-Seeking Behavior and Information Processing in Choosing a Vacation Destination.” *Journal of Travel Research*, vol. 59, no. 5, 22 July 2019, pp. 850–863, https://doi.org/10.1177/0047287519862871.

43. Paulino, Isabel, et al. “Identifying Tourism Destinations from Tourists’ Travel Patterns.” *Journal of Destination Marketing & Management*, vol. 19, Mar. 2021, p. 100508, https://doi.org/10.1016/j.jdmm.2020.100508. Accessed 12 Dec. 2020.

44. Ángel Rodríguez-Pallas, et al. “Gender and Age in the Travel Choice by Spanish Travel Agency Consumers.” *Societies*, vol. 14, no. 6, 15 June 2024, pp. 90–90, https://doi.org/10.3390/soc14060090. Accessed 5 Aug. 2024.

45. Gómez-Déniz, Emilio, and Jorge V Pérez-Rodríguez. “Modelling Bimodality of Length of Tourist Stay.” *Annals of Tourism Research*, vol. 75, 1 Mar. 2019, pp. 131–151, https://doi.org/10.1016/j.annals.2019.01.006. Accessed 17 May 2023.

46. Jackman, Mahalia, et al. “Distance Matters: The Impact of Physical and Relative Distance on Pleasure Tourists’ Length of Stay in Barbados.” *Annals of Tourism Research*, vol. 80, Jan. 2020, p. 102794, https://doi.org/10.1016/j.annals.2019.102794. Accessed 9 Apr. 2020.

47. Archer, Brian H., and Sheila Shea. “Length of Stay Problems in Tourist Research.” *Journal of Travel Research*, vol. 13, no. 3, Jan. 1975, pp. 8–10, https://doi.org/10.1177/004728757501300302. Accessed 2 Mar. 2023.

48. Montaño, Juanjo, et al. “A New Method for Estimating Tourists’ Length of Stay.” *Tourism Management*, vol. 75, Dec. 2019, pp. 112–120, https://doi.org/10.1016/j.tourman.2019.04.009. Accessed 22 Jan. 2021.

49. Brida, Juan Gabriel, et al. “Factors Influencing Length of Stay of Cultural Tourists.” *Tourism Economics*, vol. 19, no. 6, Dec. 2013, pp. 1273–1292, https://doi.org/10.5367/te.2013.0248. Accessed 21 Mar. 2020.

50. Alegre, Joaquín, et al. “A Latent Class Approach to Tourists’ Length of Stay.” *Tourism Management*, vol. 32, no. 3, June 2011, pp. 555–563, https://doi.org/10.1016/j.tourman.2010.05.003. Accessed 24 Mar. 2020.

51. Xu, Shuangyu, et al. “Impact of Travel Distance and Experience Use History on Visitors’ Climate Friendly Behavior and Support for Climate Friendly Management Action.” *Journal of Sustainable Tourism*, vol. 29, no. 6, 15 Dec. 2020, pp. 981–999, https://doi.org/10.1080/09669582.2020.1855435. Accessed 19 May 2021.

52. Nyaupane, Gyan P., and Kathleen L. Andereck. “Understanding Travel Constraints: Application and Extension of a Leisure Constraints Model.” *Journal of Travel Research*, vol. 46, no. 4, 29 Nov. 2007, pp. 433–439, https://doi.org/10.1177/0047287507308325.

53. Dean, Matthew D. “Understanding Party Size in Ride-Hailing: Solo versus Group Travel.” *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2679, no. 2, 19 Sept. 2024, pp. 2111–2126, https://doi.org/10.1177/03611981241277754. Accessed 9 Nov. 2025.

54. Federal Highway Administration. *Understanding Long-Distance Travel Behavior*. www.fhwa.dot.gov/publications/research/ear/13095/13095.pdf. Accessed 19 Mar. 2026.

55. Andersson, Angelica, et al. “Long-Distance Mode Choice Estimation on Joint Travel Survey and Mobile Phone Network Data.” *Transportation Research Part A: Policy and Practice*, vol. 190, 30 Oct. 2024, p. 104293, www.sciencedirect.com/science/article/pii/S0965856424003410, https://doi.org/10.1016/j.tra.2024.104293.

56. Reichert, Alexander, and Christian HolzRau. “Mode Use in Long Distance Travel.” *Journal of Transport and Land Use*, vol. 8, no. 2, 2015, pp. 87–105. *JSTOR*, www.jstor.org/stable/26202714, https://doi.org/10.2307/26202714.

57. Lee, Sun Hee, and Carmen Tideswell. “Understanding Attitudes towards Leisure Travel and the Constraints Faced by Senior Koreans.” *Journal of Vacation Marketing*, vol. 11, no. 3, July 2005, pp. 249–263, https://doi.org/10.1177/1356766705055716. Accessed 9 May 2019.

58. Ellegård, Kajsa, and Uno Svedin. “Torsten Hägerstrand’s Time-Geography as the Cradle of the Activity Approach in Transport Geography.” *Journal of Transport Geography*, vol. 23, July 2012, pp. 17–25, https://doi.org/10.1016/j.jtrangeo.2012.03.023. Accessed 10 Dec. 2020.

59. Montello, Daniel R. *Handbook of Behavioral and Cognitive Geography*. Cheltenham, Uk ; Northampton, Massachusetts, Ee Edward Elgar Publishing, 2018.

60. Pred, Allan. “The Choreography of Existence: Comments on Hagerstrand’s Time-Geography and Its Usefulness.” *Economic Geography*, vol. 53, no. 2, Apr. 1977, p. 207, https://doi.org/10.2307/142726. Accessed 14 Feb. 2020.

61. Miller, Harvey. “Modelling Accessibility Using Space-Time Prism Concepts within Geographical Information Systems.” *International Journal of Geographical Information Systems*, vol. 5, no. 3, Jan. 1991, pp. 287–301, https://doi.org/10.1080/02693799108927856.

62. Wilson, Clarke. “Activity Patterns in Space and Time: Calculating Representative Hagerstrand Trajectories.” *Transportation*, vol. 35, no. 4, 8 Mar. 2008, pp. 485–499, https://doi.org/10.1007/s11116-008-9162-z. Accessed 14 Feb. 2020.

63. Miller, Harvey. “Time Geography and Space-Time Prism.” *International Encyclopedia of Geography*, 6 Mar. 2017, pp. 1–19, https://doi.org/10.1002/9781118786352.wbieg0431. Accessed 29 June 2020.

64. *NHTS NextGen Study 2022 NHTS Address-Based Sample Weighting Plan*. 2024.

65. United States Census Bureau. “Explore Census Data.” *Census.gov*, 2026, data.census.gov/table/CBP2021.CB2100CBP?t=Employment+Size&g=010XX00US. Accessed 24 Mar. 2026.

66. “Airport Rankings 2023 | Bureau of Transportation Statistics.” *Bts.gov*, 2023, www.bts.gov/topics/airlines-and-airports/airport-rankings-2023.

67. flightroutes.com. “Flight Routes and Connecting Flights.” *Www.flightroutes.com*, 2026, www.flightroutes.com/.

68. Zillow. “Housing Data - Zillow Research.” *Zillow Research*, 2025, www.zillow.com/research/data/.

69. National Center for Environmental Impact. “Climate at a Glance | National Centers for Environmental Information (NCEI).” *Noaa.gov*, 2026, www.ncei.noaa.gov/access/monitoring/climate-at-a-glance/county/mapping/110/pcp/202101/1/value. Accessed 24 Mar. 2026.