ASSESSING THE FIRST AND LAST MILE PROBLEM IN INTERCITY PASSENGER RAIL: EFFECTS ON MODE CHOICE AND TRIP FREQUENCY

by

Lisa Lorena Losada Rojas

A Thesis

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Master of Science in Civil Engineering

Lyles School of Civil Engineering

West Lafayette, Indiana

December 2017

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Konstantina Gkritza, Chair

Lyles School of Civil Engineering

Dr. Jon D. Fricker

Lyles School of Civil Engineering

Dr. Jane R. Frankenberger

School of Agricultural and Biological Engineering

Dr. Dimitra Pyrialakou

Department of Civil and Environmental Engineering. West Virginia University.

Approved by:

Dr. Dulcy Abraham

Head of the Graduate Program

ACKNOWLEDGMENTS

I am deeply grateful to my major professor Dr. Konstantina Gkritza for her advice, guidance, and support to finalize this master's thesis. Thank you for giving me the opportunity to continue my graduate studies under your supervision and for your constant encouragement to be better every day. I would also like to thank my committee members Dr. Jon Fricker, Dr. Jane Frankenberger, and Dr. Dimitra Pyrialakou for their time and feedback that allow me to complete this research.

I take this opportunity to thank the Indiana Department of Transportation (INDOT) that financially supported the SPR 4404 project "Evaluating Opportunities to Enhance Hoosier State Train Ridership through a Survey of Riders' Opinions and an Assessment of Access to the Line" where the majority of the data used in this thesis was collected.

I am also grateful to the Sustainable Transportation Research Group (STSRG) members who have been a strong and positive influence on me this past year. I wish to express my sincere thanks to my friends Luisa Nuñez, Valeria Vargas, Laura Cardenas, Felipe Castro, Paula Navarro, and Raul Pineda. They have been part of some of the most important moments in both my life and career.

Finally, but first in my heart, my parents, siblings, and grandma are due my deep gratitude for their continued moral support throughout my studies. The broad education that I was able to enjoy while growing up has proven invaluable. Hector, Cecilia, Gina, Daniel, and Nelly; this is all for you.

TABLE OF CONTENTS

LIST OF TABLES viii
LIST OF FIGURES ix
ABSTRACT x
1. INTRODUCTION 1
1.1 Research Motivation 1
1.2 Research Background 3
1.2.1 Intercity Passenger Rail 3
1.2.2 First and Last Mile (FMLM) in Intercity Passenger Rail 4
1.3 Research Objectives and Questions 4
1.4 Anticipated Benefits 7
1.5 Thesis Organization 8
2. LITERATURE REVIEW 9
2.1 Research Background 9
2.1.1 Connectivity and Access to Rail Stations 9
2.1.2 Connectivity and Access to Transportation Services in Rural Areas 11
2.1.3 First and Last Mile Problem 13
2.1.3.1 First and Last Mile in Freight 13
2.1.3.2 First and Last Mile in Transit 16
2.2 Methodological Approaches 22
2.2.1 Measuring the Factor that Affect Commuter Choice 22
2.2.2 Methodologies for Measuring Access to Train Stations 24
3. EMPIRICAL SETTING AND DATA 27
3.1 Geographic Selection 27
3.2 Data Description 28
3.2.1 Primary Data 28
3.2.1.1 On-board Questionnaire 2016 29
3.2.2 Secondary Data 38
3.2.2.1 TIGER Files 38
3.2.2.2 GTFS Data 38

3.2.2.3 Smart Location Data 38
4. THE ROLE OF THE FIRST AND LAST MILE ON THE MODE CHOICE OF INTERCITY PASSENGER RAIL RIDERS 39
4.1 Introduction 39
4.2 Multi-attribute Attitude Model 39
4.2.1 Range 1 41
4.2.2 Range 2 43
4.2.3 Range 3 44
4.2.4 Range 4 46
4.2.5 Comparison of Findings 48
4.2.6 Sensitivity Analysis 50
4.3 Concluding Remarks 51
5. FACTORS AFFECTING the FREQUENCY OF TRAVEL BY INTERCITY RAIL 53
5.1 Introduction 53
5.2 Ordered Probit Model 53
5.2.1 Results 56
5.2.1.1 Marginal Effects 58
5.3 Concluding Remarks 60
6. ASSESSING THE FIRST AND LAST MILE PROBLEM IN INTERCITY PASSENGER RAIL SERVICE 61
6.1 Introduction 61
6.2 Methodology 62
6.2.1 Data and Preliminary Steps 64
6.2.2 Average Travel Cost Estimation to the Nearest Station for "driving alone” and “Someone dropped me off/picked me up” options 66
6.2.3 Average Travel Cost Estimation to the Nearest Station by Walking 68
6.2.4 Average Travel Cost Estimation to the Nearest Station by Transit 69
6.2.5 Average Travel Cost Estimation to the Nearest Station Using Ridesharing Services 70
6.3 Travel Cost Analyses Results 72
6.3.1 Scenario-based Analysis 81
6.4 Concluding Remarks 85

7. CONCLUSIONS, LIMITATIONS AND RECOMMENDATIONS 87
7.1 Summary and Contributions of this Thesis 87
7.2 Limitations 90
7.3 Future Research 91
REFERENCES 93
APPENDIX A. SURVEY 102
APPENDIX B. DISTANCE TO REACH THE STATION 109
APPENDIX C. SENSITIVITY ANALYSIS 112
APPENDIX D. ORDERED PROBIT MODEL OUTPUTS 123

LIST OF TABLES

Table 2.1 Summary of First and Last mile in freight. 14
Table 2.2 Strategies to address the first and last mile problem. (Source: TCRP Report and other contributions as stated) 17
Table 2.3 Summary of First and Last Mile in Transit Literature. 20
Table 3.1 Data Collection Schedule 29
Table 3.2 Main characteristics of the rail passenger surveyed. 33
Table 4.1 Attitude Model scores – Range 1 Quartiles (Base case) Scenario 41
Table 4.2 Average Important Rating of Attributes – Range 1 Quartiles Scenario. 42
Table 4.3 Multi-attribute Attitude Model scores – Range 2 Quartiles Scenario 43
Table 4.4 Average Important Rating of Attributes – Range 2 Quartiles Scenario 44
Table 4.5 Multi-attribute Attitude Model scores – Range 3 Quartiles Scenario 45
Table 4.6 Average Important Rating of Attributes – Range 3 Quartiles Scenario 46
Table 4.7 Multi-attribute Attitude Model scores – Range 4 Quartiles Scenario 47
Table 4.8 Average Important Rating of Attributes – Range 4 (Base case) 48
Table 5.1 Ordered Probit Estimation Results 56
Table 5.2 Marginal Effects 58
Table 5.3 Cross tabulation of predictions for the ordered probit model 59
Table 6.1 Data Sources 65
Table 6.2 Road Type Classifications (U.S. Census Bureau, 2015) 66
Table 6.3 Lyft charges according to the area (Source: (Inc, 2017)) 71
Table 6.4 Summary Statistics for Cost Analysis. 81

LIST OF FIGURES

Figure 1.1 Research Framework 6
Figure 3.1 Area traveled by respondents to reach the station. 35
Figure 3.2 Area traveled by respondents to reach the station (Origin only). 36
Figure 3.3 Area traveled by respondents to reach the station. 37
Figure 4.1 Average Score per Attribute – Range 1 Quartiles Scenario 42
Figure 4.2 Average Score per Attribute – Range 2 Quartiles Scenario 44
Figure 4.3 Average Score per Attribute – Range 3 Quartiles Scenario 46
Figure 4.4 Average Score per Attribute – Range 4 Quartiles Scenario 47
Figure 4.5 Average scores per attribute due to different ranges for Intercity Train and Driving Alone (scores are displayed as a percent of maximum possible score). 49
Figure 5.1 Histogram trip frequency in the previous year. 55
Figure 5.2 Histogram new categories for trip frequency in the prior year. 55
Figure 6.1 Methodology for Estimation the Cost to Access the HST Stations. 63
Figure 6.2 Average Travel Cost to the Nearest Station by Driving Alone 74
Figure 6.3 Average Travel Cost to the Nearest Station by Driving Alone – Parking included. ... 75
Figure 6.3 Average Travel Cost to the Nearest Station by “Someone dropped me off” option ... 76
Figure 6.4 Average Travel Cost to the Nearest Station by Walking 77
Figure 6.5 Average Travel Cost to the Nearest Station by Transit (Tippecanoe County) 78
Figure 6.6 Average Travel Cost to the Nearest Station by Using Ridesharing Services 79
Figure 6.7 Average Travel Cost to the Nearest Station by Transit (Tippecanoe and Marion Counties) 83
Figure 6.8 Average Travel Cost to the Nearest Station by Ridesharing (All Counties) 84

ABSTRACT

Author: Losada Rojas, Lisa Lorena. MSCE
Institution: Purdue University
Degree Received: December 2017
Title: Assessing the First and Last Mile Problem in Intercity Passenger Rail: Effects on Mode Choice and Trip Frequency
Major Professor: Dr. Konstantina Gkritza

Passenger rail service is an integral part of intercity transportation networks, especially in areas where residents do not have access to cars or other intercity travel options. Some municipalities in the U.S. have experienced a decline in passenger rail service in recent years, which has prompted schedule reductions and entire abandonment of service in some cases. To improve the current intercity passenger rail service predicament, two alternatives can be considered: (1) improve the rail service itself (frequency, infrastructure, etc.) and (2) improve accessibility to the rail stations, which might be cheaper and more cost-effective overall. Improvements in accessibility can impact a wider area and play a key role in passengers choosing rail service as their travel alternative. To address the above issues, the main objective of this thesis was to explore the possibilities for enhancing access to medium distance travel which is, according to the U.S. Bureau of Transportation Services (BTS), between three to five hours or more than 50 miles of travel from home to the nearest intercity passenger rail station. The approach of this thesis was to identify the factors that affect mode choice and level of usage in order to subsequently evaluate different strategies for passengers to reach a station.

The Hoosier State Train (HST), a short-distance intercity passenger rail system that travels between Chicago and Indianapolis four days a week, was chosen as a case study. HST has four intermediate stops in Indiana. For some of those intermediate stops, HST is the only intercity public transit service offered to reach either Chicago or Indianapolis. An HST on-board survey that explored opportunities to increase the HST ridership was conducted in November and December of 2016. The survey findings indicated that there are passengers who travel from counties farther away from a county with a station to take the train. Moreover, it was found that most of the respondents drove a personal vehicle, rented a car, or were dropped off to reach a train station in Indiana. The first and last mile (FMLM) of a trip is commonly used to describe passenger travel as far as getting to/from transit stops/stations. The findings of this thesis suggest that there

is a gap in the FMLM for intercity rail passengers. Solving the FMLM problem would extend access to transportation systems and could increase the number of passengers from remote communities, such as rural areas. The FMLM problem has been addressed in different public transit contexts, mainly within urban areas; however, limited research efforts have been undertaken to examine the FMLM problem of intercity passenger rail. This thesis intends to fill this gap by exploring the best strategies to address the FMLM problem of short distance intercity passenger rail (i.e., corridors that are less than 750 miles long according to the Passenger Rail Improvement and Investment Act, 2008).

Using the data collected on board the HST in Indiana, this thesis estimated a multi-attribute attitude model (MAM) to assess how transportation mode preferences for intercity travel are made and how the factors considered in mode choice decisions vary among individuals with different levels of access to an intercity passenger rail line. An ordered probit model was estimated to further investigate how passenger characteristics, as well as the factors associated with both access to a rail station and mode choice decisions, relate to the frequency of travel by intercity rail. This thesis also presents the results of an accessibility analysis conducted for the state of Indiana in order to identify the areas in need of FMLM service where no public transportation services exist and the cost of reaching a station from a desired origin is expensive. To that end, a cost survey for the different modes available was conducted to determine the average travel cost to the nearest station. The analysis was carried out in ArcGIS using origin-destination information from the on-board survey, transportation network information from the U.S. BTS, and general transit feed specification data.

The results of this thesis can assist Amtrak and state transportation agencies identify which aspects of rail service potentially can be enhanced to attract more passengers as well as promote the use of intercity passenger rail service in the U.S. Additionally, the findings could have extensive implications for planning strategies to provide access to passenger rail stations. While the inferences in this thesis are case-study specific for Indiana, the proposed methodology could be used to identify areas where accessibility can be improved in other U.S. states or countries with similar characteristics.

1. INTRODUCTION

1.1 Research Motivation

Transportation plays an important role in different aspects of society. Transportation systems can influence economic, social, and industrial changes in a particular area. The relationship between transportation and the economy is bidirectional: increased economic output leads to an increased amount of travel and increased travel leads to higher economic output (Sinha & Labi, 2007). On the other hand, transportation systems can have a significant influence in fostering social changes, including changes related to social exclusion that occur when segments of the population are prevented from participating in activities that affect their quality of life (McCray & Brais, 2007).

An important element of social exclusion related to transportation pertains to accessibility. Widely understood as the ability, potential, or ease of reaching desired opportunities, accessibility is an extensively applied concept in different disciplines (Martens, Golub, & Robinson, 2012; Foth, Manaugh, & El-Geneidy, 2013). There are methods to measure how accessible an area is in terms of primary care services, jobs, supermarkets, etc. and accessibility has been studied with regards to specific demographic groups such as minority and low-income households and households in rural areas. Accessibility measures also have been used to assess the transportation disadvantages of an area, which are the disadvantages of a population, group, or area due to lack of mobility and/or accessibility (Pyrialakou et al., 2016). While the literature on assessing transportation accessibility presents several methods, there is not a unique methodology to measure accessibility (Handy & Niemeier, 1997).

From a transportation system perspective, multimodal system connectivity and access to public transportation are part of the key variables that contribute to the measure of accessibility (Governors' Institute, 2017). Access to public transportation always has been a concern; but most recently, interest has increased in streamlining the journey "chain," which includes the journey to and from the designated stations or stops by different modes of transport (Givoni & Rietveld, 2007). That added link in the transportation chain becomes a problem when there are no options available. This problem is commonly referred to as the first and last mile (FMLM) problem of a trip and is generally used to describe passenger travel with regard to getting to/from stops/stations. FMLM transportation connectivity to/from a major transit line extends the access opportunities to more places of interest to commuters living in remote communities (Chandra, Bari, Devarasetty, & Vadali, 2013). While the FMLM problem is known to impact the accessibility of distant commuters to their surrounding environment, this problem is currently not well understood and a unique way to address it has yet to be defined. A few cities (New Jersey, Dallas, and Atlanta) primarily are using ride-share services like Uber, Lyft, or Juno, which are not completely subsidized and therefore make commuting more expensive (King, 2016). Other solutions are infrastructure investments around the stop/stations, park and ride facilities, and feeder systems, such as fixed routes to connect the surrounding areas, especially intercity transportation stations (Metro, 2014b).

Additionally, it has been acknowledged that the access journey to a passenger rail line can be a factor in determining if rail service is the chosen travel alternative (Rietveld, 2000). Access to transportation facilities is a factor that influences the level of usage of services (Moniruzzaman & Páez, 2012). Since railway stations, even in major cities, usually are located somewhat far away from each other, getting to or from them becomes an important part of a rail journey and therefore must be accounted for in any efforts to increase rail use. Improving access to stations might be less expensive and more cost-effective overall rather than improvements to the actual rail journey (Givoni & Rietveld, 2007).

By improving access to railway stations or bus stops, using those services could increase, which is of particular importance to intercity bus and passenger rail ridership, which has experienced a 4.4% decrease from 2005 to 2010 in the U.S. (BTS, 2011). During that same period, it was estimated that approximately 3.4 million people living in rural areas lost access to intercity passenger transportation. This loss is primarily due to the discontinuation of intercity passenger options such as bus and passenger rail services (Pyrialakou, Gkritza, & Fricker, 2016). In 2010, 40% of the rural population in the U.S. had access to intercity transportation, which refers to the rural population living within 25 miles of a bus, rail station, or airport. However, the percentage of rural population being able to use these services might be much lower because of the FMLM problem to access these public services.

This thesis proposes a methodology to explore how to enhance access to the nearest intercity passenger rail stations that are medium distance from home, which according to the BTS is between three to five hours or more than 50 miles of travel. The approach of this thesis was to identify the factors that affect mode choice and level of usage in order to evaluate different strategies for passengers to reach a station.

1.2 Research Background

1.2.1 Intercity Passenger Rail

For nearly 100 years, the U.S. was the worldwide leader in passenger rail transportation. Trains were the primary mode of transportation available for medium and long distance travel. However, according to the 2010 National Transportation Statistics, the dominant mode of transportation today is the highway system and the percentage of trips via intercity trains or Amtrak is nearly zero. The U.S. invests a small amount in passenger rail comparative to the size of the population and landmass (ACE, 2017).

The Passenger Rail Investment and Improvement Act of 2008, PRIIA Section 209, declared that short distance Amtrak corridor services (less than 750 miles) must be state-funded (PRIIA, 2008). This decision has caused intercity passenger rail to face many difficulties due to the lack of additional investment opportunities and less government funding to make the service more successful in the future.

Despite Amtrak’s growing ridership, with 2016 being the sixth year in which ridership exceeded 30 million, Amtrak must cover 94% of its operating cost with ticket sales and other operating revenue (ACE, 2017). Amtrak strongly relies on government funding for capital investment and additional projects. Even though Amtrak received a $2.45 billion loan from the U.S. Department of Transportation (DOT) in the fall of 2016, those funds were primarily used to invest in new high speed trains and has not solved the large and growing backlog of capital needs, especially in short distance corridors like the HST. It has been argued that, at the state and regional level, rail should become part of multimodal strategic policies and capital investment programs that support the role of passengers (ACE, 2017).

On the other hand, considerable attention has been given high speed rail in the U.S. This investment would potentially moderate automobile and air traffic congestion throughout the U.S and lead to significant economic, environmental, and quality of life benefits. However, the low density, automobile-oriented development that has dominated U.S. cities is not appropriate for the kind of access provided by high speed rail (Lane, 2012). Due to the lack of a centralized, dense, and highly-accessible location to place a station, and the lack of accessibility provided by non-automotive modes; the FMLM problem appears in rail service, which means that part of the population will not have access to the high speed rail system because of lack of access to stations. Consequently, development of high speed rail will not provide the needed benefits.

1.2.2 First and Last Mile (FMLM) in Intercity Passenger Rail

The FMLM of a trip has been used to describe passenger travel with regard to getting to and from transit stops. This problem was identified first in freight transportation with failed attempts to deliver a product the first time as well as the congestion that this procedure created in the road system. The FMLM problem also has been addressed in different public transit contexts, mainly in urban areas. However, it is also an important part of the commute journey in an intercity trip.

In 2006, an "America on the Go" report presented the percentage of trips by access mode for long distance trips by public transportation mode (Patterns & Choice, 2006), which indicated that for the train mode, 54.4% of people were reaching the stations using personal vehicles, followed by 20.4% of people who were reaching the station by multiple modes. The remaining 25.6% of people were either walking, bicycling, using public transportation, or availing other modes to reach the train station (BTS, 2007).

Efforts to increase rail use have focused mainly on the rail service itself while expanding access to the rail network has received a reduced amount of consideration. This alternative could increase rail use by making rail services more accessible to more potential passengers from a wider geographical coverage of access services. Brons et al. (2009) noted that an important way to improve access to railway stations is through the implementation of public transport services around the station. From a policy perspective, the authors emphasized the importance of integrated transport to accomplish a shift from personal cars to public transport modes and, specifically, for long distance trips a shift to rail.

1.3 Research Objectives and Questions

This thesis had two general objectives. The first objective aimed to identify the different factors that influence the mode choice of passengers for medium distance travel as well as the level of usage of intercity passenger rail service. The second objective of this thesis aimed to explore strategies to attract a wider number of passengers by addressing the FMLM problem in an intercity passenger rail system. To achieve these objectives, the following research questions were posed:

1. What role does the FMLM play in the mode choice of medium distance passengers?
2. What is the relationship between frequency of travel by intercity rail and (i) mode choice-related factors, (ii) factors associated with access to a rail line, and (iii) passenger characteristics?
3. Which strategies are the most helpful for accessing an intercity passenger rail service?

The first two research questions correspond to the first objective of this thesis and the third question addresses the second objective. To answer those questions, a research framework was developed, which is shown in Figure 1.1. The research framework is composed of three parts that correspond to the three questions addressed. The HST in Indiana served as a case study. The survey data were collected on board the HST in the fall of 2016; and the 908 responses served as the primary source of data for this thesis.

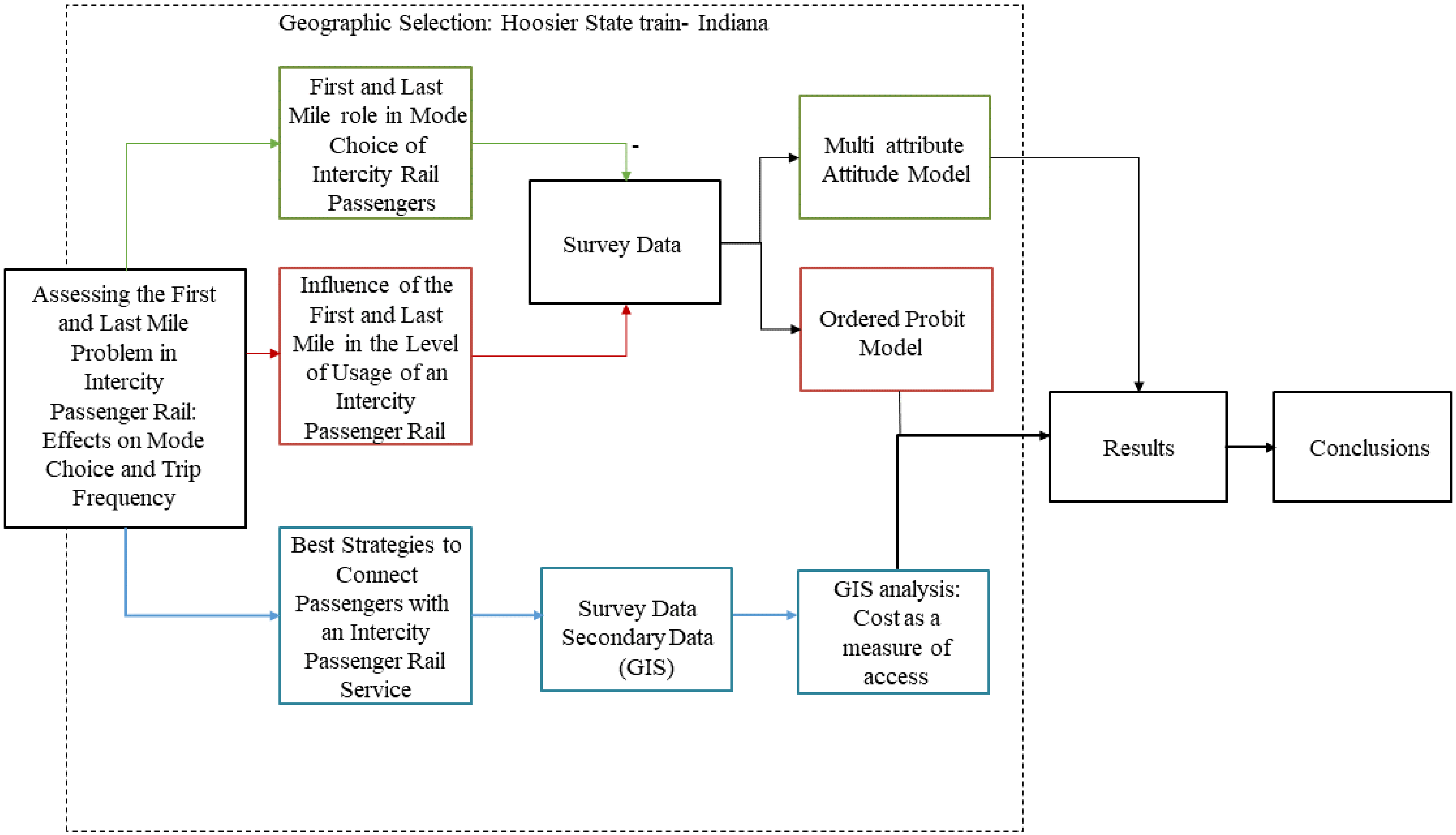


Figure 1.1 Research Framework

To assess the FMLM problem in short distance intercity passenger rail systems, it was necessary to identify which factors are more significant when choosing between different options for medium-distance travel. To address the first question, data from Section 3 of the on-board survey conducted in the fall of 2016 on the HST (see Appendix A) were used to estimate Fishbein's theory models as to the range of distances traveled by the passengers to reach their stations. Such distances served as an indicator of access to the station and helped to comprehend the FMLM journey. On the other hand, Fishbein's theory model, also called MAM, has been used to measure a consumer's attitude toward a service (Wilkie & Weinreich, 1972) and therefore was used to measure passengers' attitudes towards different medium-distance transportation modes available in the area. This analysis enabled identifying whether the factors that affected the mode choice decision varied across different distances traveled to reach a station and which ones were more important to target in the different types of areas that these ranges covered.

The second research question aimed to identify the relationship between the frequency of travel by intercity rail and other factors associated with the passenger characteristics, the mode choice-related factor, and factors associated with access to a rail station. For this question, the same survey data were used. The dependent variable in this analysis was intended to be the frequency of passengers; and because that variable was grouped into ranges, an ordered probit model was proposed. The ordered probit model was chosen because of the nature of the dependent variable.

Lastly, an accessibility analysis was carried out in ArcGIS to address the third question. This last question aimed to identify the best strategies to connect nearby passengers with intercity passenger rail stations. The most advantageous strategy was evaluated in terms of the cost to reach the station. To do so, cost rasters were developed to measure accessibility in terms of the cost to reach a desired destination. The methodology was performed in ArcGIS Pro in order to identify the areas where there is a need of FMLM service and then propose strategies that have been used in other transportation systems to address the problem.

1.4 Anticipated Benefits

There has been increasing interest in the FMLM problem in public transportation in urban areas, but it has not been widely extended to rural applications. There has not been a lot of research that measures how the distance to and from the stations affects commuters' decisions in modal choice. Although, the survey questionnaire used in this thesis was not designed to address this specific problem, some of the questions on it were useful to analyze the FMLM problem. The results may have far-reaching implications for planning the strategies to provide accessibility and connectivity to rail stations. The methodology used in this thesis could be used to identify areas where accessibility could be improved, not just in Indiana but elsewhere. Also, the proposed strategies to address the FMLM problem could provide useful suggestions to intercity passenger rail services that would serve to attract a wider number of passengers. Specifically, the case study and empirical results not only could provide the Indiana Department of Transportation (INDOT) Multimodal Planning and Programs Division a vision for expanding the area where the HST serves as well as provide access to more people in the future, but it also identifies some factors and strategies for making passenger rail service a viable and attractive option.

1.5 Thesis Organization

This thesis includes seven chapters. This first chapter provides a contextualization of the topic and discusses the anticipated benefits. Chapter 2 presents a synthesis of previous studies on the different topics and methodologies referred to in this thesis. Chapter 3 defines the geographical area and the data used for each of the steps. Chapter 4 describes the methodology and the results of the first research question proposed in this thesis. Chapter 5 describes the methodology and results of the second research question proposed in this thesis. Chapter 6 defines the methodology and results of the third research question proposed in this thesis. Finally, Chapter 7 discusses the conclusions and limitations of this work and offers recommendations for future research.

2. LITERATURE REVIEW

To address the thesis research questions, the author develops the background of this research. A literature review related to connectivity to public transportation stops/stations, rural connectivity to public transport modes, first and last mile problems and common solutions was conducted. A literature review related to methodological approaches to conduct the analysis was also undertaken. The key findings and research gaps identified are presented below.

2.1 Research Background

2.1.1 Connectivity and Access to Rail Stations

The term ‘connectivity’ refers to the availability of a transportation system to facilitate travel between services (Pyrialakou, 2016). The improvement of transportation connectivity is one of the essential tasks for transportation operation planning (Ceder, Net, & Coriat, 2009). On the other hand, the term ‘access’ refers to the opportunity for potential riders to get from where they are to the transit service (Murray, Davis, Stimson, & Ferreira, 1998). Access coverage reflects the fact that riders only use a service that they can get to and it is a key topic in public transportation planning (Wu & Murray, 2005). Additionally, Murray et al. (1998) discusses access and accessibility in the context of public transportation. Accessibility relates to the suitability of the transit system to move people from where they board to where they exit in realistic amount of time, whereas access typically has to do with proximity to service and its cost.

Connectivity and access to rail stations is an important factor when the expansion of a public transportation system is being considered. For light rail stations, Kuby, Barranda, & Upchurch (2004) carried out research on the characteristics of the access and egress journeys to and from stations mainly with respect to distance, time and other variables, such as park and ride spaces, and number of bus connections. With that in mind, they used multiple regression to determine factors that contribute to higher light-rail ridership. The significant variables for explaining ridership were grouped in five main categories, such as land use traffic generator, network structure, city wide, socioeconomic variables, and intermodal category. That analysis was performed at the station level by using network based buffers generated by GIS. Although that study was conducted around urban and suburban areas, the identified factor gives insight to which aspects could increase transit use and the importance of them in the accessibility of surrounding zones (Kuby, Barranda, & Upchurch, 2004).

Improving public-transit connectivity is one of the most important tasks in transit-operations planning (Hadas & Ranjitkar, 2012). To this end, Hadas et al. (2012) estimated a multi-objective model to integrate spatial and non-spatial data for assessing public-transit connectivity and offer a decision-support tool for the identification of inefficiencies. Auckland, New Zealand served as a case of study for this analysis. Geographical information systems and non-spatial data (General Transit Feed Specification from Google) was used to provide a tool for a public transit network analysis. As a contribution, that study constructed an analysis on how to measure transit-network connectivity in terms of the value of time and quality of transfer, which were calculated within a GIS package. Since the connectivity indicators were calculated within GIS, the author stated that it was possible to examine large public transport networks, as was confirmed in Auckland, Vancouver and Portland (Hadas, 2013).

Welch et al. (2013) studied a measure of equity for public transit connectivity (how quality transit was distributed between households in the area studied). In that study, the connectivity index, built with an assessment of service quality, incorporated features of each transit line and stop, such as distance, activity density, capacity, speed, required transfers, and frequency of the underlying land use served by a transit mode (Welch & Mishra, 2013). This index is a measure of transit-service quality at each stop, along every transit line. The Washington-Baltimore region was used for case study. The tool proposed in that study could be used by transit agencies in measuring the distribution of transit service between specific groups to offer better access to captive riders, who are unable to afford automobiles.

Transit connectivity is also influenced by factors such as frequency, in-vehicle travel time, service reliability, access/egress times, waiting time, and transfers along multimodal paths. A study in the Greater Copenhagen Area was conducted to assess the equity for the multimodal transit system. The methodology used for measuring equity in transit provision involved three strategies: (i) measuring transit connectivity, (ii) calculating location-based and potential accessibility measures, and (iii) computing Gini coefficients per area that provides an equity measurement. It was found that Transit Oriented Development (TOD) in that specific case study was useful from a spatial equity perspective. Although that study was conducted in a metropolitan area, the proposed methodology could be applied to other study areas (Kaplan, Popoks, Prato, & Ceder, 2014).

An evaluation of accessibility of two different feeder transit services, fixed route transit (FRT) and demand responsive transit (DRT), was conducted in Denver, Colorado (Chandra et al., 2013). The authors stated that transportation network-related improvements access to destinations and impact mobility. A gravity-based model was used to estimate the potential accessibility. The authors concluded that the method defined in that research provides guidance for using available decay factors in assessing how far or close transit agencies are from reaching the best possible accessibility through feeder services. The authors also concluded that feeder transit services are a cost-effective, safe, and reliable transportation mode for first/last mile connectivity, because they are specifically planned to cater to the requirement of a door-to-door type of service. Similar to previous studies, that study was carried out in a metropolitan area, which is not directly comparable to the connectivity of an intercity passenger rail, but the methodology could be used to study how these two services will perform in such as areas.

Turning to intercity rail, a methodology to assess the connectivity of high speed rail (HSR) in Extremadura, Spain is presented in Gallego et al. (2015). That study was based on the use of tools for network design and GIS to explore accessibility produced by the HSR. The first variable used to weight the degree of importance to access the services station was the least access time of population to the train stations with a cut-off buffer of 600 km. The second variable was the improved index of absolute accessibility (IAA) which measures the interconnection of a population with its surrounding region. It was concluded that a high speed rail favors the Extremadura region to change from a peripheral region to become a more integrated territory (Gutiérrez Gallego, Naranjo Gómez, Jaraíz-Cabanillas, Ruiz Labrador, & Jeong, 2015).

2.1.2 Connectivity and Access to Transportation Services in Rural Areas

Lack of transportation services for people is often identified as an important constraint on rural development. Rural communities face multiple challenges associated with accessibility and connectivity that relate to both a physical and virtual sphere (Velaga, Beecroft, Nelson, Corsar, & Edwards, 2012). The landscape in the United States has been transformed over the past century. The rural population represented 54.4% of the population in 1910 census and, in 2010, the total population that lived in rural areas was merely 19.3% (United States Census Bureau, 2010). Departments of Transportation (DOTs) have placed less planning emphasis and resources on rural transportation concerns, particularly in areas of multimodal planning. The interaction of multimodal planning outside of urban areas will likely continue to slowly progress (Dixon, Sarasua, Daniel, & Mazur, 2001). In fact, some studies have considered the analysis of rural road network planning (Rao & Jayasree, 2003) where the connectivity is evaluated in terms of infrastructure (transportation network such as roads). In Rao's and Jayasree's analysis, the need for spatial planning of network configuration was confirmed. That can achieve the desired results of social interaction, economic, and broad development in a region. Other studies have considered the travel pattern and accessibility in areas with little, or no, public transport service. For instance, a study carried out in Australia, which is akin to the USA due to low rural densities, searched for an indicator of transport-related problems using Census Data. Socio-economic variables were selected for potential explanatory connections such as population density, age groups and employment, Native American populations, and low income households (Nutley, 2003). Also, a 'travel needs index' was proposed that combines both potential mobility and service center accessibility. Although the authors could not predict a 'problem' in rural areas in terms of trends, location, or relationship with socioeconomic factors, they encouraged further investigation due to their findings of lower vehicle ownership in the remote areas.

In Scotland, the context for accessibility and connectivity in rural communities was examined, highlighting key transportation and technology changes. The study identified that the lack of transportation infrastructure, fewer passengers, sparse population, and mobile communication systems are the problems suffered by provision of a passenger information system in rural areas. The authors concluded that innovations oriented towards advanced technologies and transportation telematics can make a substantial contribution to address the accessibility problem (Velaga et al., 2012).

Another study identified that the current lack of alternative transport modes means that car ownership has become necessary and not a choice for the rural commuter (Cheyne & Imran, 2010). This study, carried out in New Zealand, presented data from research on shared (or flexible) transport in metropolitan areas. Cheyne and Imran concluded that many groups in this type of population do not have the option of private vehicles and alternatives in the form of shared or public transport, which is key for economic and social health. The methodology of the study was the analysis of census data, a primary survey conducted along the line, as well as interview data. After analyzing the data obtained, the authors concluded that the current lack of alternative transport modes means that car ownership has become necessary and not a choice in rural areas, but the increasing availability of information and communications technologies means shared transport services can enhance transport choices for low density communities.

In regards to the exposed problems, there is still a need for research about public transportation in rural areas, specially addressing how connectivity can be improved and providing an enhancement on accessibility to those areas. The next subsection is dedicated to the first and last mile problem in different transportation modes.

2.1.3 First and Last Mile Problem

The FMLM problem was drawn originally from telecommunications. In that realm, the FMLM is the final leg (or first leg) to the consumer. In the 1970's and 1980's, as cable TV was being deployed across the US, cable companies had to individually wire each and every household, at a tremendous but necessary cost (King, 2016). This problem has been faced also in transportation systems. The following subsections presents the literature review relative to first and last mile in freight and transit.

2.1.3.1 First and Last Mile in Freight

The growth of e-commerce and the congestion caused by freight transportation in urban areas has motivated the study of the last mile in delivery activities. The growth in home delivery activity has increased concerns over freight traffic in often not suitable residential areas. Also, freight traffic leads to a CO_2 emissions increase due to road transport and some environmental cost for added vehicle trips for all parties. (Iwan, Kijewska, & Lemke, 2016; Song, Cherrett, McLeod, & Guan, 2009; Wygonik & Goodchild, 2016). The problem of FMLM in freight is also referred as "city logistics". It has been studied in developing economies, where it represents up to 28% of the total freight cost (Muñoz-Villamizar, Montoya-Torres, & Vega-Mejía, 2015).

The FMLM problem in freight transportation has been addressed in different ways. Song et al., (2009) proposed theoretical collection and delivery points (CDPs) in the network of existing business establishments. On the other hand, Dell'Amico & Hadjidimitriou, (2012) presented the City Log project where the combination of two types of vehicles and the Modular Bentox-Box (M-BB) were introduced as an innovative logistics model for urban delivery. A collaborative scenario between at least two actors demonstrated to be noteworthy, especially for cities that have several delivery locations (stores) to supply and an inadequate infrastructure to deal with the increased traffic (Muñoz-Villamizar et al., 2015). Additionally, as a solution of the last mile problem in freight transportation, a study on an effective crowd-tasking model with scalable solutions was conducted by Wang et al. (2016). That study concluded that the concept of crowd delivery in city last mile problem has a substantial impact on urban logistics development.

In regards to the previous discussion, most of the last mile solutions in freight transportation are focused on the proximity of a certain place where the delivery needs to be made to the final destination. Most of the results on the previous studies talked about the possibility to serve the last mile by a third party. Those solutions concentrated the goods in one place where they will be collected and sent to their final destinations. However, the idea of gathering passengers in a hub to be transported to another terminal or station will add an additional link to the existing transport "chain" that passengers have to face. The literature review in terms of freight transportation did not give a clear insight about how to approach the problem of FMLM for passenger transportation, although the problem is also related to proximity as it is in transit. Nevertheless, the review of this literature helped to understand how this problem has been faced in other modes and which factors are affected by this problem, such as transportation externalities. A summary of the studies is presented below (Table 2.1).

Table 2.1 Summary of First and Last mile in freight.

Study	**Area and Data**	**Objectives**	**Summary of results**
(Song et al., 2009)	● West Sussex, United Kingdom. ● Household database (Home delivery questionnaire). Networks of existing business establishments as theoretical collection and delivery points (CDPs).	● Assess the impacts of failed first-time home deliveries on extra carrier journeys and consumer trips. ● Quantify the transport benefits if they were to be implemented across the county	● Benefits might increase from using networks of Local Collect post offices, railway stations, and supermarkets as CDPs, when compared to the traditional delivery method.

(Dell'Amico & Hadjidimitriou, 2012)	● Lyon, France. ● Districts. Network. Environmental calculator (EcoTransiIT) Location of depot.	● Show the savings that the new customer distribution allows to obtain in terms of kilometer traveled, pollutant emissions and cost for transportation.	The successful of urban consolidation center (UCC) depends on: ● The nature and volume of traffic. ● The possibility to introduce financial support to operation and to enhance the service offering to attract greater throughput.
(Muñoz-Villamizar et al., 2015)	● Bogota, Colombia ● Real-data from convenience stores operating in the case study. OD matrix and driving distance.	● Address the problem of the last mile urban freight transport under collaborative systems. ● Aim at proposing an analytical approach, to assess the benefits of collaborative freight delivery in urban areas. ● Compare the allocation-routing decisions in both non-collaborative and collaborative scenarios.	● The collaborative scenario was shown to be significant especially for cities whose have several delivery points to supply and deficient infrastructure to manage the increased traffic.
(Wygonik & Goodchild, 2016)	● Seattle and King County, Washington ● Household database. VMT, road type, speed, vehicle type and emissions for three different goods movement schemes	● Examine the relationship between good movements and development patter characteristics including density and distance from warehousing. The work questioned if the impact from last mile goods movements strategies differ with	Last mile goods movement relying on: ● Delivery services result in the lowest generation of CO_2 per customer, except in road-dense locations. ● Passenger vehicles always result in the lowest generation of NO_x and PM_{10}.

		urban form characteristic.	
(Iwan, Kijewska, & Lemke, 2016)	● Szczecin, Poland. ● Location and data from the parcel lockers operation. Population around the parcel. Usability survey.	● Analysis of usability and efficiency of parcel lockers system as a solution of the last mile problem.	● The growth of e-commerce has an influence on the growing demand for last mail delivering. ● The most important factor of efficiency of solutions as locker parcel is the proper location of the machines used for deliveries.
(Wang et al., 2016)	● Singapore and Beijing. ● Bus database. Taxi database. Travel records. Pop-station location.	● Investigate how to use the power of crowd-workers to improve the last-mile delivery.	● The solution presented can support real-time delivery optimization in the large-scale mobile crowd-sourcing problem. ● The crowd delivery in city last-mile application has substantial effect on the urban logistics development.

2.1.3.2 First and Last Mile in Transit

The FMLM problem has also been faced in transit. The lack of adequate connectivity between transits stop and trip origin or end points has limited transit in playing a bigger role as a transportation mode in cities and urban areas. There is a need for transportation alternatives to make transit more competitive and appealing. The ride itself is a vital part of making transit appealing: the quality of the waiting environment, fare level, service frequency, and in-vehicle amenities (Tilahun, Thakuriah, Li, & Keita, 2016). In addition to proximity, as it was discussed in the freight transportation, access conditions also depends significantly on various features of the built and social environment in which the last-mile trips make place. The conditions also depend on the physical connectivity, place-based barriers and lack of specific solutions such as connecting transport or lack of information. Tilahum et al. (2016) also presented that the last mile problem is a multipart issue that needs to be addressed with a multi-pronged approach that not only consist of transportation, but also urban design solutions and more comprehensive social policies. Chicago, Illinois was used as a case of study. For a last mile standpoint, the paper findings suggested that enhancements to accessibility and related built environment structures, such as job density and diversity, at the terminating end of the trip may be much more significant in influencing choice. Also, the paper concluded that, by involving the use of information technology-based solutions that affect mobile technologies, it would be easier to find, in real-time, “walking or traveling buddies” from bus stops and train stations in insecure areas. Social media can be an important factor in implementing these types self-organizing strategies.

The FMLM problem has been solved in different ways according the mode of transportation used as a feeder (defined as a peripheral route or branch in a system, which connects minor or more remote nodes with a route carrying heavier traffic). Shared-use vehicle service is a term including booth car sharing and station car programs as solutions to the first mile and last mile problem (Shaheen, Meyn, & Wipyewski, 2003). However, the difference between these two concepts is that car sharing enables an individual to obtain the benefits of private-vehicle use at a lesser cost relative to vehicle ownership, taxis, or conventional rental. On the other hand, station car programs primarily facilitate transit access. Nevertheless, both are now used as a solution of their last or first mile problem. Some other solutions were summarized in the TCRP Research Report 188 (Murphy, Transit Cooperative Research Program, Transportation Research Board, & National Academies of Sciences, Engineering, and Medicine, 2016), and elaborated by the author in the following table (Table 2.2). It is worth to mention that these options could be also combined to produce multimodal solutions to the first and last mile problem.

Table 2.2 Strategies to address the first and last mile problem. (Source: TCRP Report and other contributions as stated)

Term	**Description**	**Source**
Crossing and Connections	A set of strategies focused on pedestrian mode that include enhance crosswalks to protect	(Metro, 2014)

	pedestrian and active transportation users when crossing vehicular traffic, cut-troughs and shortcuts to provide more direct routes to and from the station, raised crossings, among other.	
Bikesharing	It is a short-term bike rental, usually for short periods of an hour or less that typically requires a membership. Information technology (IT)-enabled public bikesharing provides real-time information about the position and availability of bikes at stations in an area.	(Murphy et al., 2016)
Carsharing	Automobile rental for intervals of less than a day. Major car sharing business models include traditional or round-trip, which has users borrow and return vehicles to their original location; one-way or free-floating, which permits users to pick up a vehicle at one location and drop it off at a different one; and peer-to-peer (p2p), which allows car owners to rent out their vehicle, when they are not using it, to other carsharing members. Neighborhood Electric Vehicles (NEVs) proposed in the First Last Mile Strategic Plan is an example of it.	(Murphy et al., 2016)
Micro transit	IT-enabled private multi-passenger transportation services that attend passengers by dynamically generated routes, and may expect passengers to make their way to and from common pick-up or drop-off points. Because they provide transit-like service but on a reduced and more flexible scale, these new types of services have been referred to as “micro transit.”	(Murphy et al., 2016)
Private shuttles	Traditional private shuttle services include corporate, regional, and local shuttles that make fewer stops, often only picking up designated riders.	(Murphy et al., 2016)
Ridesharing	Ridesharing implicates adding passengers to a private trip in which driver and passengers share a destination. Such an organization provides additional transportation options for riders. Traditional forms of ridesharing include carpooling and vanpooling.	(Murphy et al., 2016)

Kiss and Ride	As part of the Plug in Components in the FLM Strategic Plan, this is a designated pick-up/drop-off area in a convenient location.	(Metro, 2014)
Park-and-Ride	These facilities are parking lots with public transport connections that allow passengers and other people heading to city centers to leave their vehicles and transfer to a bus, rail system, or carpool for the remainder of the journey.	(Metro, 2014a)
AV's	Autonomous vehicles (AV's) have been studied as a potential solution for the last mile trips between a train station and the traveler's final end.	(Yap, Correia, & van Arem, 2016)

Furthermore, to address the FMLM problem in the rail journey, the Integration Between and Access-to-rail-stations Modes (IBRAM) has gained attention as a research topic. This term refers to the integration of the journey component that is critical to achieve continuous travel, door to door, in order to make the rail attractive alternative to the car (Givoni & Rietveld, 2007). As an initial point to the IBRAM research, a study examined the modes that passenger used to get to or from railway station in the Netherlands. Using a regression model, the authors measure the significance of the access and egress journey in passenger's overall satisfaction. They concluded that, in general, passenger would accept extensive journey time and distances for the access journey than for the egress journey.

In addition, there is a study focused on the preferences of travelers for using automated vehicles as last-mile public transport of multimodal train trips (Yap et al., 2016). This study aims to evaluate automated vehicles as a last mile solution when traveling by train. That study conducted a stated preference survey in a large national online panel. Using a discrete choice model to explore inclinations of travelers for using automated vehicles, it was concluded that the usage of AV's as last mile transport between the train station and the final end for first class train travelers is preferred versus the use of non-motorized modes or public transit.

Another study (Liang, Correia, & van Arem, 2016), investigated the potential of using automated taxis (ATs) as a last mile solution of train trips. They defined that the use of entirely automated electric vehicles to feed this system could be a worthy alternative to bring more people to public transport and increase sustainability. Two integer programming, formulations were made; these formulations depended on how trips were selected from the total number of reservations made in one typical day where passengers needed to book in advance. The mathematical models tested the effect of service zone location and trip selection on the profitability of the AT system.

The Los Angeles County Metropolitan Transportation Authority (Metro) released the first-last mile strategic plan and planning guidelines in March 2014. These guidelines addressed 3 main goals: i) expand the reach of transit through infrastructure improvements, ii) maximize multimodal benefits and efficiencies and iii) build on the Regional Transportation Plan/Sustainable Communities Strategy and Countywide Sustainable Planning Policy (multimodal, green, equitable, and smart). This plan was developed to be around the rail system in Los Angeles County, which will be a short distance from Los Angeles County residents.

A summary of the previous studies and their most notable conclusions can be found in Table 2.3.

Table 2.3 Summary of First and Last Mile in Transit Literature.

Study	Area and Data	Objective	Methodology	Summary of results
(Tilahun et al., 2016)	● Chicago, Illinois ● Eight Counties Census track level. Socio-economic and built environment related data.	● Study the role that public transport last mile problem plays in mode choice decision of travelers. ● Analyze the problem considering a wider range of area factors including but not limited to transit availability, and social characteristics such as street-level crime.	It undertakes that mode decisions are made based on utility maximization. The analysis, was applied to home-based work, work related school and school-related trips.	● The incidence of non-domestic violent crimes decreases the probability of using non-motorized alternatives. ● Improved destination accessibility significantly boosts transit use more when compared to increases in origin-level accessibility. ● The results dispute for enhanced accessibility and related job densities at job locations.

(Yap et al., 2016)	● Netherlands. ● State preference experiment distributed as online survey with 9 different mode alternatives, four of which involve the use of an AV as egress.	● Place automated vehicles in the public transport market. ● Understand the sympathy of travelers toward instrumental travel attributes.	A Bayesian efficient design, which aims to maximize the expected D-error, was estimated. To explore the preferences of travelers for using automated vehicles, a discrete choice model was estimated and the utility maximization framework was used.	● Travelers' attitudes play an important role in the attractiveness of using AVs as last mile transport. ● Travelers associate more disutility to the in vehicle time in an automatically driven AV than in a manually driven vehicle. ● Potential for AVs as new mode of transportation between the train station and the final destination.
(Liang et al., 2016)	● Delft, Netherlands. ● 48 potential zones (average size of 500 m X 500 m). The data needed was taken from a survey conducted at the train station. Request during an average day, driving distance and travel trim and cost of	● Examine the likely of using automated taxis (ATs) as a last mile connection of train trips. ● Present a way to optimize the service area of an AT system which fulfills passenger's request to access or egress a train station.	Two integer programming (IP) models were estimated which aim to define the optimal service area and trips to be completed by the AT system. The models were for a scheme called free service and the scheme called full service.	● Zone location and trip selections are likely to reduce the negative effect of lack of taxis and contribute to the maximization of the profit. ● Fleet size is a key factor of the productivity of the AT system.

	running the system.

The literature reviewed in this section accounts for different strategies studied to resolve the first and last mile problem. However, most of the studies are related to giving access to transit stations located in urban areas. Although some of those studies such as the one developed in Netherlands Yap et al., (2016), addressed the FMLM problem for train mode, the characteristics of those systems are largely different from the ones found in the United States concerning intercity passenger rail. The characteristics of the territory and its density allowed to address the FMLM problem with the approaches exposed previously, which could not be exactly replicated in the United States. In New Zealand, some strategies of demand responsive passenger transport (DRT) in reduced demand situations were analyzed. An operational component of DRT is the origin-destination pairs where the alternative many origins to one destination was studied. This alternative aims to serve areas with more densify demand and can be useful to any service including rail stations (Scott & NZ Transport Agency, 2010). Other services, such as the Flexible Transport Service (STF) experienced in Scotland, seem like a promising solution for widespread public transport in rural areas (Velaga et al., 2012). Even though those studies serve as first and last mile solutions, they are being developed in countries with different characteristics than in the United States. The lack of first and last mile research for intercity passenger rail has not been a significant matter of concern for researchers until now. To the author's knowledge, there lacks studies that address the FMLM problem and effects of rural commuters in the United States.

2.2 Methodological Approaches

In order to address the proposed goals, a revision of the past methods to measure the factors that affects commuter choice, the level of usage of intercity passenger train, and access to train stations are identified herein.

2.2.1 Measuring the Factor that Affect Commuter Choice

The selection a specific good or service is a difficult choice process when a consumer is making a decision (Lindgren Jr & Konopa, 1980). Researches have attempted to model process of information evaluation in order to better understand consumer behavior choice. Those models of attitude formation and change have been suggested and used in different fields such as economics, psychology, and marketing with the intent to measure a consumer's attitude toward a service (i.e. mode choice), and to identify the specific attributes related with those objects. These models are called Multi-attribute Attitude Models.

To predict the factors that affect commuter choice, a Multi-attribute Attitude Model (MAM) is considered. The MAM was originally proposed by Fishbein and Rosenberg in 1967. It was based on the notion that an individual's attitude towards an objective is a function of his/her beliefs about the object that are significant to the evaluation and the implicit evaluative responses pertaining to those beliefs (Wilkie & Weinreich, 1972). With respect to marketing, this method has been extended to suggest that attitudes toward brands are driven by a consumer belief regarding the ability of different brands to satisfy the specifically desired product attribute intensities. This can be seen in the following equation:

$$A_j = \sum_{i=1}^{n} b_{ij} a_i \qquad (2.1)$$

where, for each individual, A_j represents the attitude toward brand j, in this case brand will take the form of a transportation mode, b_{ij} represents the rating of mode of transportation j on attribute i, a_i represents the importance of attribute i in forming an overall attitude toward the transportation mode, and n represents the number of attributes that a person will look at.

Pyrialakou (2016) estimated a Multi-attribute Attitude Model to better explain the mode choice decisions by passengers who perform medium distance trips. According to Pyrialakou, that analysis could support a prioritization of the policy and planning choices promoting a mode shift towards an intercity passenger rail service. The attributes considered by that study were defined as qualities or features that characterized a transportation mode. The following attributes were considered:

- Cost
- Travel Time
- Comfort
- Safety
- Amenities
- Flexibility of travel
- Convenient/flexible schedule
- Reliability

- Ease of traveling

According to Pyrialakou (2016), those attributes were also chosen from the pilot survey results. In addition, the competing modes of transportation selected for that analysis included personal vehicles (driving alone and carpooling), intercity bus, and airplane. For the application of the model, passengers were asked to rate the current attribute i in the transportation mode (b_{ij}) on a scale of 1 (poor) to 5 (excellent) and the evaluation of attribute i (a_i) was rated on a 5 points scale, from (1) not important at all to (5) extremely important (Pyrialakou, 2016).

The use of a MAM to predict the factors affecting mode choice decision in terms of the distance needed to reach a station/stop has not been explored, however, it could be useful to identify the lack/need of improvement in a given factor according to the distance traveled to reach or leave the stations. In that way, a prioritization of policies could be made at different area levels such as blocks, counties or economic development region level.

2.2.2 Methodologies for Measuring Access to Train Stations.

Spatial analysis is a type of geographical analysis that pursues to describe patterns of human behavior and its spatial expression in terms of mathematics and geometry (Mayhew, 2009). Correspondingly, network analysis is a network-based spatial analysis tool for solving complex routing problems. These type of analyses have been used to measure access to train stations. For example, a spatial analysis of access to and accessibility surrounding train stations was conducted in Perth, Western Australia to study the accessibility for elderly people. Accessibility is measured by a composite index based on three travel modes (walk-and-ride, park-and-ride and bus-and-ride) using spatial methods. ArcGIS was used to perform the index calculations considering street blocks for the walk and ride analysis in an 800 meter area, census districts for the bus analysis in an 800 meter service area around the bus stops, and census districts in an area of 90 percent of access trips for the park and ride analysis (Lin et al., 2014).

Additionally, access has been measured in different way considering cost, and travel time, among other factors to reach core services such as stops/stations, health centers, and healthy food. (Murray, 2003) stated that essential characteristics of well utilized transportation system are, among others, being accessible and efficient. Also, physical distances are important measures of core services utilization and are affected by the transportation cost that includes time spent in travel, monetary cost of travel, and discomfort related to travel.

Transportation adds complexity to the measurement of geographical access (McLafferty, 2003). To overcome this complexity, GIS have been used to further explore the measurement of access. The possibility to consider various transportation systems (i.e. automobile, public transit, and walking) have become part of the analysis and it has been useful to understand the access situation of a stop/station. The use of GIS has enabled integrating different tools to calculate travel time, the visualization of the network, and results. Murray, (2003) discusses that one access consideration in potential use of public transportation is the travel distance or time from the origin to a stop/station. Another considerations are also cost of service, safety in getting from origin to stop, and barriers in travel to/from stop/station where cost can be included. Using Euclidian distance in the evaluation of access, the author stated that geographic information packages could support the strategic analysis of a transit system. Most of the studies used travel time as a measure of accessibility. Fewer studies have considered cost in their accessibility analysis. Modeling travel time or cost of public transport has been recognized as complex and difficult, particularly with regards to accurately representing the travel cost for bus routes (Liu & Zhu, 2004; Lovett, Haynes, Sünnenberg, & Gale, 2002; O'Sullivan, Morrison, & Shearer, 2000). In a research conducted in Glasgow, Sullivan et al. (2000) demonstrated that an effective set of desktop GIS tools could produce isochrones maps. Isochrones provide a simple method for determining accessibility when using public transport. However, this research stated that simplifications were required to make the work feasible due to the complexity of the task in GIS and the availability of data. The authors of that study also mentioned that an isochrones approach could be use either with time or cost as a measure of access. Comparatively, Lovett (2002) used GIS to calculate measures of accessibility to surgery centers by public and private transport. To that end, the travel times to nodes on the road network to a health center were calculated using a Triangulated Irregular Network (TIN). The bus analysis was performed in an 800-meter buffer around the centers. A circular buffer zone was defined and the routes that served the area were selected. However, the location of the stops was ignored. However, the time at which passenger were reaching the surgery location was not considered and this added a limitation to the study. Furthermore, Liu et al. (2004) used ACCESS, an integrated GIS tool designed to support the integrated GIS approach to accessibility analysis, as an instrument to support the accessibility analysis process. The ACCESS tool uses mainly spatial analysis, network analysis, 3D analyst and patch analyst operations. This tool also allowed the authors to calculate different accessibility measures such as constrained potential model and modified potential model, which could have either cost or time as parameter to calculate the accessibility. The authors concluded that this GIS tool could be used for transportation analysis and land use planning in a larger scale.

Other authors (Burns & Inglis, 2007; Hallett & McDermott, 2011) have used transportation cost methods to analyze access to core services by threated cost as a fixed value represented by the type of road, the cost to operate a vehicle and the U.S. federal minimum wage. Burns et al. (2007) examined access to healthy and unhealthy food in Melbourne by creating a cost surface to determinate the travel cost to either supermarket or fast food outlets. In that study, travel cost depends on the limit speed when personal vehicle is analyzed and road type and frequency for the bus analysis. Likewise, Hallett et al, (2011) examined and refined the discussion of food deserts by using GIS. That research also measure the cost of distance imposed on consumer to reach healthy food. To that end, the location of full-service grocery stores were shown in maps with respect to transportation networks and population distribution. In that study, the federal minimum wage, the speed, and the value for cost of operation a motor vehicle were considered in the cost model. A raster representation of the road network was created and each cell was given a cost that represented the transportation and opportunity cost of traversing that cell. The use of the minimum wage was justified stating that it was a conservative value of the opportunity cost, but it was recognized that it could lead to an underestimation of the cost. Similarly, Bailey (2016) explored the transportation cost by mode using the federal minimum wage and the cost of operating a motor vehicle to reach healthy food and built a methodology to identify food deserts. By using spatial analysis tools, the cost to reach a supermarket in different modes such as driving, walking, and transit was estimated. The results showed that driving was the less expensive way to reach a supermarket when comparing with transit and walking. This finding helped identify the areas where it was most expensive to reach healthy food and then compare them with the location of low-income households in order to identify food deserts. The use of the minimum wage in that study was supported by past literature but it recognized that the use of that value as a cost of time could underestimate the real cost of transportation.

3. EMPIRICAL SETTING AND DATA

Chapter 3 details the empirical setting and data to address the three research questions stated in the first chapter. This chapter is composed of two sub-sections. The first sub-section addresses the reason for the selection of the HST in Indiana as the case of study. The second sub-section includes the description of the data needed to examine the proposed research questions.

3.1 Geographic Selection

The State of Indiana, and particularly the Hoosier State Train (HST), was chosen for several reasons. First, after the Passenger Rail Investment and Improvement Act of 2008 (PRIIA), the State of Indiana took charge of the HST, because this line is classified as short distance service (196 miles). Since then, the Indiana Department of Transportation (INDOT) has tried to maintain the service using local communities' funding, agreements with Amtrak, and private partnerships. There has been a growth in passengers and revenue, but there is still not enough to pay the full operating cost and forthcoming capital investments. Second, the HST has stops in five counties in Indiana, each with different characteristics. Lake, Jasper, Tippecanoe, Montgomery and Marion are the counties served by the HST. Based on the analysis performed by Pyrialakou (2016), who used three different classifications of urban-rural schemes (OMB metropolitan-micropolitan statistical areas, rural-urban continuum codes (RUCC) by ERS and urban influence codes (UIC) by ERS), three of the counties along the HST are considered large urban areas (Marion, Lake, and Jasper), one is a medium to small urbanized area (Tippecanoe), and one is a non-metropolitan area (Montgomery). Also, only three of those counties are served by intercity buses, which typically run to and from Chicago.

In addition, Pyrialakou (2016) found that a percentage of the HST riders (27%) were not residents of the counties with a station, but rather traveled to a station from other counties in Indiana such as Hamilton, Madison, and Hancock, among others (Pyrialakou, 2016). That information was obtained through on-board survey conducted in Fall 2015. A follow-up of that on-board survey was designed and funded by INDOT through the Joint Transportation Research Program. The follow-up survey was conducted in Fall 2016. Based on the findings of Pyrialakou (2016), there was a need to explore how much people were willing to travel to reach a train station and which mode they were using to do that. Also, there were additional questions about household location and opinions related to different factors around the train. The data collected through that project is the main data source of this thesis and is discussed next.

3.2 Data Description

3.2.1 Primary Data

The analysis presented in this thesis is based on a follow up on-board survey carried out during nine days between November 13 and December 2, 2016 (Sunday, Wednesday, and Friday). This survey was part of a project [SPR 4044: *Evaluating opportunities to enhance the Hoosier State Train ridership through a survey of riders' opinions and an assessment of access to the line]* funded by the INDOT. This project aimed to (i) assess the potential impact on ridership if improvements were made to the services, and (ii) identify population that would be more likely to ride the train. The survey was conducted by the author and another Purdue graduate student. The survey instrument was reviewed and approved by the Institutional Review Board (IRB protocol # 1503015896A002). A total number of 908 responses were gathered with a response rate of 85%.

In order to design the follow-up survey, the questionnaire used in Fall 2015 was considered as a primary source. However, the follow-up questionnaire was modified to address different issues that were identified as part of the 2015 survey results. Those issues were related to origin-destination responses, accessibility perception, need for more information about the perceived ease of use and usefulness of the passenger rail services, and information about the future usage of the service. The 2015 questionnaire was shared with the project's Study Advisory Committee, which evaluated the number of questions that would remain identical to 2015 and which new questions. After this revision, a pilot survey was conducted on September 28, 30 and October 2nd at the Lafayette Amtrak Station, gathering 30 responses (3% of the expected sample).

The on-board data collection was scheduled for nine days over three weeks (see Table 3.1). Permission from Amtrak to conduct the survey was obtained in advance with a request for "temporary permit to enter upon Amtrak property" and the completion of a contractor safety and security awareness training session by both graduate students. On board, the questionnaires were distributed to all eligible passengers (individuals who had already completed the survey once persons younger than 18 years old were excluded), who boarded the HST after the train departed from each station.

Table 3.1 Data Collection Schedule

Day	Date	Departure Station	Arrival Station
Sunday	11/13/2016	Indianapolis	Chicago-Union Station
Sunday	11/13/2016	Chicago-Union Station	Indianapolis
Wednesday	11/16/2016	Indianapolis	Chicago-Union Station
Wednesday	11/16/2016	Chicago-Union Station	Indianapolis
Friday	11/18/2016	Indianapolis	Chicago-Union Station
Friday	11/18/2016	Chicago-Union Station	Indianapolis
Sunday	11/20/2016	Indianapolis	Chicago-Union Station
Sunday	11/20/2016	Chicago-Union Station	Indianapolis
Wednesday	11/23/2016	Indianapolis	Chicago-Union Station
Wednesday	11/23/2016	Chicago-Union Station	Indianapolis
Friday	11/25/2016	Indianapolis	Chicago-Union Station
Friday	11/25/2016	Chicago-Union Station	Indianapolis
Sunday	11/27/2016	Indianapolis	Chicago-Union Station
Sunday	11/27/2016	Chicago-Union Station	Indianapolis
Wednesday	11/30/2016	Indianapolis	Chicago-Union Station
Wednesday	11/30/2016	Chicago-Union Station	Indianapolis
Friday	2/12/2016	Indianapolis	Chicago-Union Station
Friday	2/12/2016	Chicago-Union Station	Indianapolis

3.2.1.1 On-board Questionnaire 2016

The questionnaire used for the on-board survey began with a brief introduction of the HST, and the improvements that it had undergone since the joint partnership between was formed between Iowa Pacific Holdings, Indiana Department of Transportation, Amtrak, and the Cities of Crawfordsville, Lafayette, West Lafayette/Tippecanoe County, and Rensselaer in 2015. The following sections explain the content of the final survey. The questionnaire can be found in Appendix A.

- **Section 1: Trip Characteristics and Experience with the Hoosier State Train.**

The first section was composed of one sub-section 1.1 "*Trip characteristics and experience with the Hoosier State Train*". This section included questions about the characteristics of the trip and the familiarity of respondents with the service. Some of these questions were not included in the previous survey, but they were found important for this follow-up survey to gather information needed to conduct the accessibility analysis.

There were two questions related to riders' origin and destination pairs. Question 1 and 4 asked about the station where people boarded and got off, respectively. The options for these two questions were the 5 stations that HST serves. In the same way, there were two questions associated with the distance people needed to travel to reach the departure station, and also, the distance needed in order to reach their final destination. These two questions were numbered as 2 and 5, and they were open-answer questions.

Section 1 includes questions designed to identify the mode that riders used to reach and leave their departure and arrival station, respectively. These questions included modes such as driving or renting a car, riding the bus, walking, being dropped by someone, using a bicycle, taking a taxi or a ride-sharing service like Uber, Lyft or other mode. The question related to the mode used to reach the station where the riders boarded was associated with a sub-question about the location of parking in case they used their personal vehicle to access the station. That last question was intended to capture the ease of parking around the station for those who drove a car as an access mode. These questions were 3a, 3b, and 6.

Four additional questions related to the experience on the HST. The question number 7 was associated with the frequency with which riders traveled on the HST in the year previous to the survey. Question number 8 asked about the purpose of the trip. Questions 9 and 10 were related to the experience on the train as part of a big group and the possible discounts that could have been applied when purchasing tickets for the HST, respectively. Those questions were intended to measure the level of usage of the HST, as well as the level of usage of the available options to ride the train in a cheaper way.

- **Section 2: Ease of Use and Usefulness of the Hoosier State Train**

Section 2 is composed of 4 sub-sections. Overall, these sub sections tested the perceptions of the passenger about the HST service nowadays and in the future. Section 2.1 "*Ease of using*

Hoosier State train" included 11 questions about the ease of using some resources that people interact with during their experience as riders of the HST. This section included questions related to the interaction with the ticketing system and the information system (Questions 1 and 2). Moreover, this section included questions about the perception of the distance from riders' house location to the station as well as the parking availability near the HST stations (Questions 3, 4a, and 4b). Section 2.1 also included questions about access to the platform for riders with and without disabilities (Questions 5a and 5b). In addition, questions about riders' perception on the storage space of luggage or essentials goods on board (question 6 and 7). Question 8 and 9 were related to the improvements that the HST has introduced after the joint collaboration started. These questions asked about the changes for on-board amenities (e.g., Wi-Fi, hot meal services, snacks and beverages) and the feature where people can ride with a pet on the train. Question number 10 referred to the ease in finding travel brochure information related to Indiana destinations at the HST stations. Finally, question 11 asked for the overall ease in traveling with the HST. The responses provided to these questions ranged from "strongly agree" to "strongly disagree" with the statements made. Questions 1, 4a, 4b, and 9 permitted the response of "not applicable" for those who did not relate those statements with their current situation.

Section 2.2 "*Usefulness of the Hoosier State train*" consisted of 6 questions. These questions aimed to provide information about when people consider that would be more likely to travel with the HST, based on speed, safety, time, cost, and travel purposes. Question 1 asked about the possibility to reach a destination faster by traveling with the HST. Question 2 asked about the perception of a safer trip on the HST, and Question 3 asked about the perception of productivity on board. Questions 4 and 5 were related to the cost of traveling alone or with a group in the train. Lastly, Question 6 questioned whether riding the HST line fit the traveling purposes of the respondents. The responses provided from these questions ranged from "very unlikely" to "very likely" to the statements made.

Section 2.3 questions "*Your on riders' thoughts about the Hoosier State train*" were included in order to learn the opinions of the HST. This section included 6 questions. The first question asked that if more people used the HST, it would be good for the environment. Similarly, Question 2 asked if using the HST would contribute to the reduction of traffic congestion and Question 4 asked if it would enhance economic development in Indiana. Question 4 was explicit and asked if the State of Indiana should invest funding to support the HST service. The schedule convenience for riders' travel purposes was asked in Question 5. Finally, Question 6 asked about the on-time perception to reach a destination using this train service. The responses provided to these questions ranged from "strongly agree" to "strongly disagree" towards the statements made.

Section 2.4 "*Using the Hoosier State train in the future*" asked about the intention of use the HST service in different scenarios. The first question asked about the intention to travel on the train in the next month, which was aimed to gauge respondents' short term intention to travel on the HST. The second question asked about the expectation to travel on the train in the foreseeable future, which was aimed to gauge respondents' long term intentions. Question 3 examined the possibility of riding on the HST if gas prices were higher in the future. Similarly, Question 4 asked about the possibility of riding the HST if about parking costs would be higher in the future. The last question of this section (Question 5) asked about the possibility of riding the HST if one's bicycle could be brought on the HST. The responses provided to these questions ranged from "strongly agree" to "strongly disagree". Question 5 considered the option "Not applicable" for those who did not own a bike.

- **Section 3: Mode Choice**

The third section was consisted of one sub section 3.1 "Mode choice". This subsection led to tables that provided the primary information needed to conduct a Multi-attribute attitude analysis. The attributes measured in the 2015 survey by Pyrialakou (2015) were the same considered in the 2016 survey. The attributes measured were defined as qualities that characterized a transportation mode. Based on Fishbein's theory, the following attributes were considered:

- Cost
- Travel Time
- Comfort
- Safety
- Amenities (Wi-Fi, food, etc.)
- Flexibility of travel (ability to go wherever one chooses)
- Convenient/flexible schedule
- Reliability (not being late)
- Ease of traveling (minimize the effort required to travel)

The first table summarizes the level of importance for each of these attributes when the respondent was selecting a medium distance mode to commute (Medium distance is understanding as more than 50 miles from home to the furthest destination (Cho, 2013)). The evaluation of attributes was rated on a 5 point importance scale, from (1) not important at all to (5) extremely important.

The second table asked to rate each of the attributes considered in the previous table in terms of five different modes: 1. Automobile-Drive Alone, 2. Automobile-Carpool, 3. Intercity Bus, 4. Intercity Train and 5. Airplane. Respondents were asked to rate the attributes in each mode choice on a scale of 1 (poor) to 5 (excellent).

Finally, the section considered a question about daily mode choice. This question asked whether the respondent would always travel by car to go to work or go shopping.

- **Section 4: Demographic Questions**

Finally, socioeconomic and demographic questions were included in Section 4 in order to examine variations in the attitudes and behaviors towards passenger rail among different socioeconomic and demographic groups. This group of questions asked about the sex of the respondent, age range, employment situation, annual household income, and education level, number of children in the household, personal vehicles, and household state, county and city location.

The main characteristics of the passengers surveyed are summarized in Table 3.2. Most of the passengers used the HST for the first time, and did not reside in Indiana. With respect to the age, over half of the respondents were in the 18-35 age group and a minor share are over 65. The distribution of passengers by gender was 46% male and 54% female.

Table 3.2 Main characteristics of the rail passenger surveyed.

Variable	***Mean or Percentage (Standard Deviation)***
Gender Female/Male	54/46
Age 18-24/25-34/35-44/45-54/55-64/65 and over	35/20/11/9/11/14
Employment situation Full Time/Part Time/Student/Retired/Unemployed/Other	43/7/32/14/2/2
Household Income	23/20/21/14/11/11

Under $25,000/$25,001-$49,999/$50,000-$74,999/$75,000-$99,999/$100,000-$149,999/$150,000 and over	
Education level Grade School/Some High School/High School Graduated/Technical Training/Some College/College Graduate/Graduate School	1/2/12/3/30/28/24
Household Size One/Two/Three/Four/Five or more	33/29/14/13/11
Number of children in the household None/One/Two/Three/Four	80/9/7/3/1
Weakly vehicle mileage 5-99/100-299/300-499/500-1,000/more than 1,000/I do not own a vehicle	40/17/5/1/37
Household located in Indiana Yes/No	59/41
Origin Stations Chicago/Dyer/Rensselaer/Lafayette/Crawford /Indianapolis	48/3/2/20/6/21
Distance traveled to reach a station (miles)	45.3 (211.80)
Mode to reach the station Drove or rent a car/Rode a bus/Walked/ Someone dropped me off/Bike/ Taxi, or ridesharing/Other	22/6/12/29/0/18/13
Trip Purpose Work/Social-Recreational/School/Other	6/83/8/3
Frequency in the last year 0/1-2/3-4/5-6/7-8/9-10/>10	40/30/13/5/4/2/6
Importance Rating of the Attributes Cost/Travel time/Comfort/ Safety/ Amenities /Flexibility of travel/ Convenient/Reliability/Ease of traveling	3.89(0.88)/3.68(0.93)/3.79(0.89)/4.06(0.95)/3.50(1.04)/3.66(0.89)/ 3.80(0.83)/4.04(0.87)/4.03(0.82)
Respondent prefers train over personal vehicle Yes/No	49/51
Respondent prefers personal vehicle over airplane Yes/No	68/34
Respondent believes train is safer than personal vehicle Yes/No	85/15
Respondent believes train is easier to travel by than personal vehicle Yes/No	65/35

As expected, most respondents owned one car or more, but they were still riding the train for medium-distance trips. Another remarkable finding was the mode used to reach a station (Figure 3.1). The most dominant mode of transportation for access and egress to the station in Indianapolis was the option of having someone dropping the passenger off to the station/picking the passenger up from the station (45% and 58%, respectively). The second option was driving or renting a car (25% and 21%, respectively) and the third option was using a taxi or a ride-sharing service (20% and 15%, respectively). A similar trend was observed for the rest of the stations in Indiana. This finding suggests that there is a possible gap into the first and last mile travel options for the riders and alternative options to fill this gap need to be considered. Contrary, the station located in Chicago, had a different trend of access and egress modes. Alternatives modes considered in "Other" such as Metra or metro were found as the most popular to reach the station. Due to those findings, the FMLM will be analyzed in the stations located in the state of Indiana.

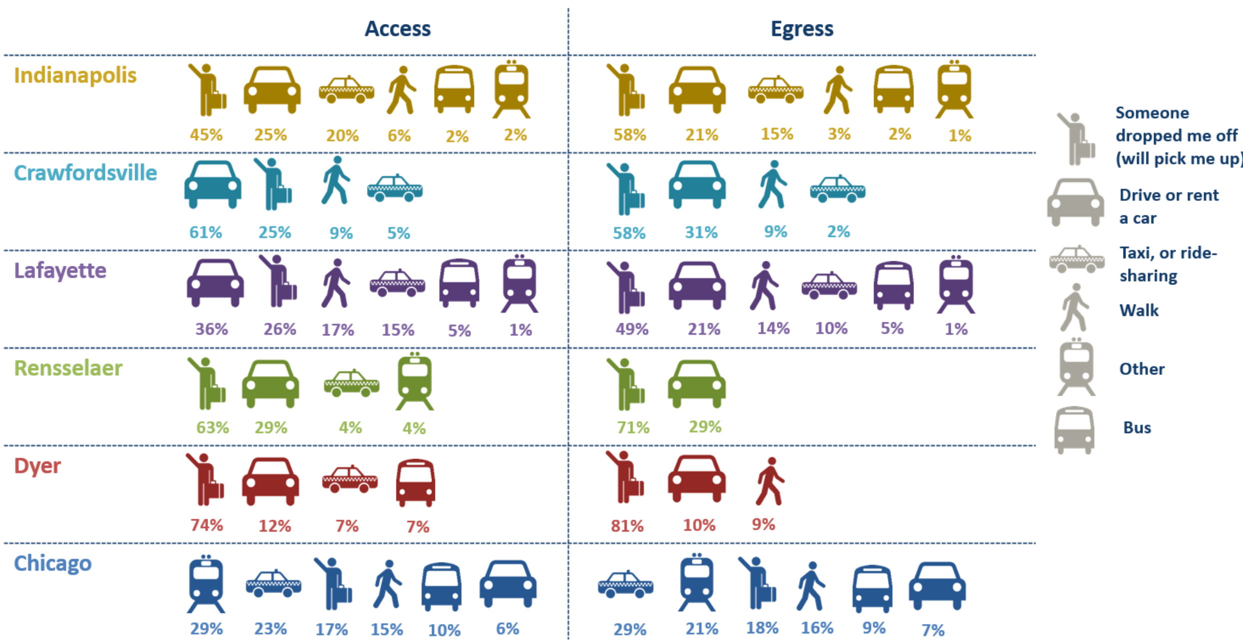


Figure 3.1 Area traveled by respondents to reach the station.

When a similar graphic is developed considering only the origin station of the respondents, different outcomes are observed in Figure 3.2. For instance people who rode the train from Chicago (as origin) would leave the train station mainly using "someone dropped me off" option or driving a car. This supported the idea that the FMLM problem in Indiana stations is a common difficulty observed by riders traveling from Chicago. Contrary, some passengers who took the train from one of the five Indiana stations used other modes such as buses or taxis to reach their final destination. Nonetheless, most of the passengers who originated from Indiana stations used “someone dropped me off” as option for the last mile of their trip.

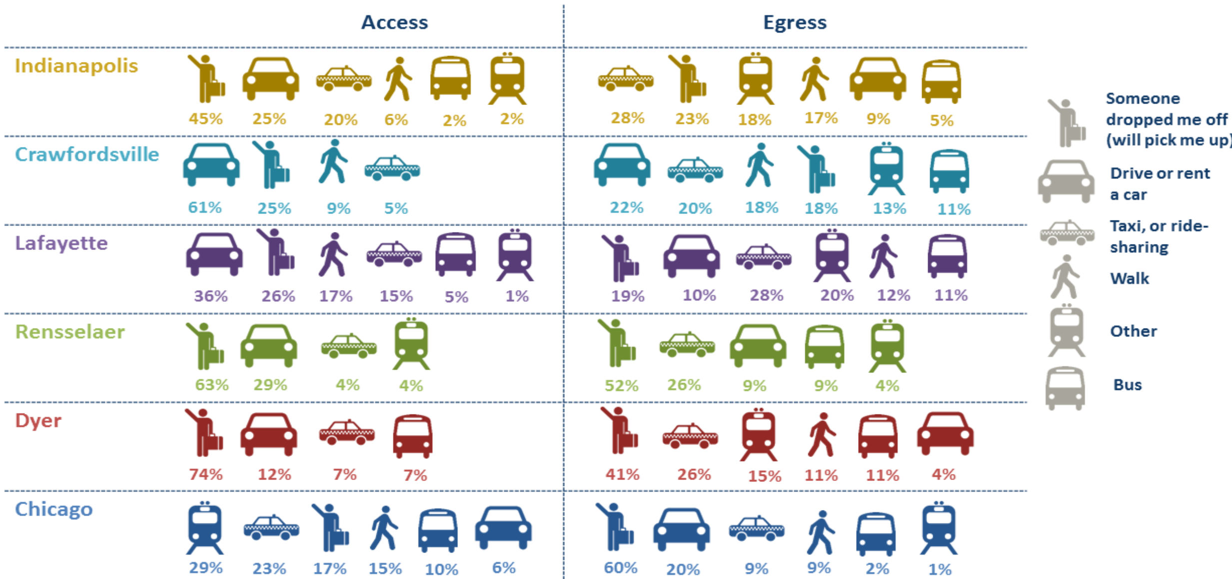


Figure 3.2 Area traveled by respondents to reach the station (Origin only).

Additionally, some passengers stated that they traveled more than 60 miles to reach a HST station. The area covered by a 60 miles buffer will be used in subsequent chapters to analyze the first and last mile problem. Figure 3.3 presents the area covered by different buffers around Indianapolis station. As it can be seen, 7% of respondents who reached Indianapolis station (21% of survey responses) traveled more than 60 miles. Appendix B presents the area covered for the different buffer for Chicago, Dyer, Rensselaer, Lafayette, and Crawford.

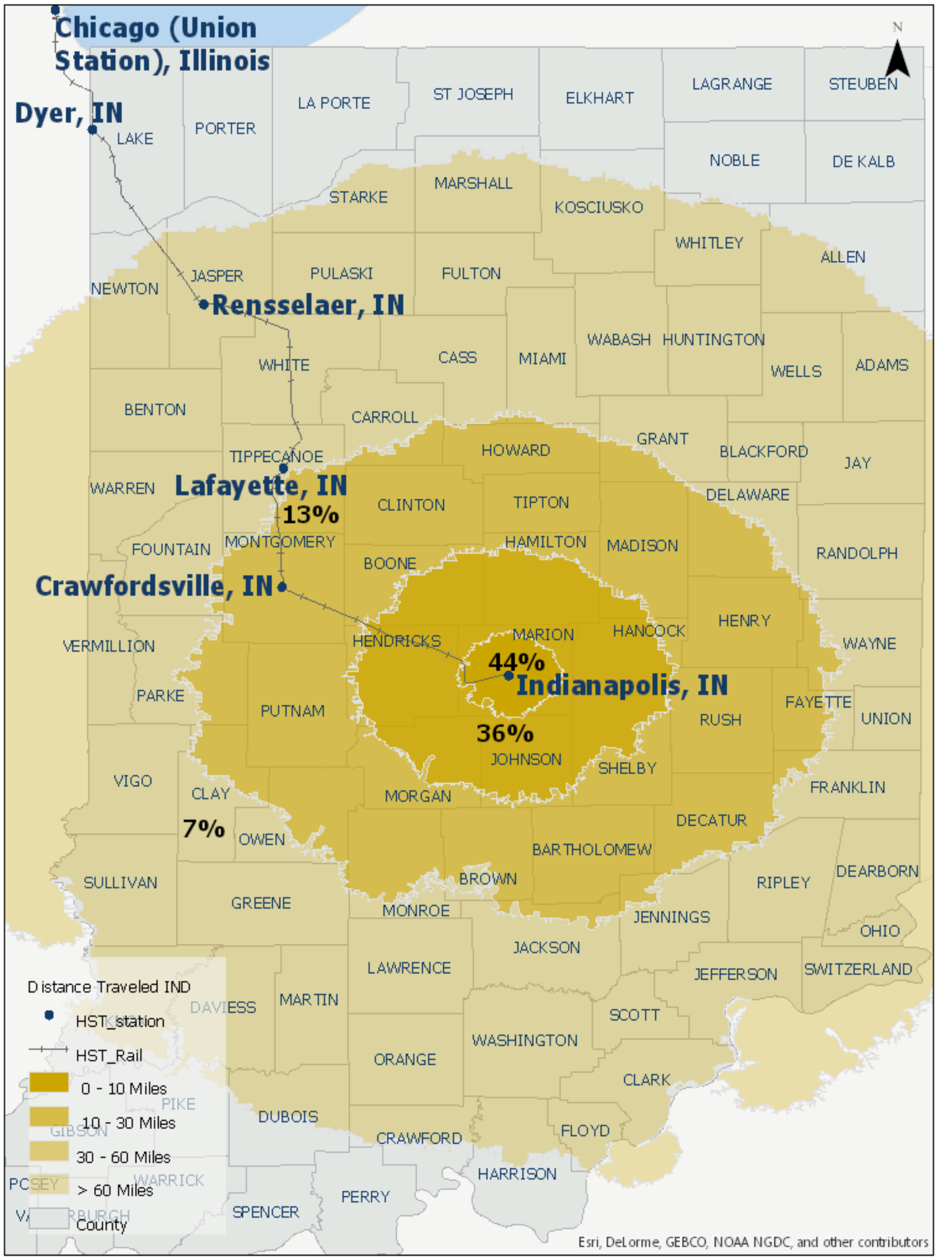


Figure 3.3 Area traveled by respondents to reach the station.

In order to address the research questions for this study, the on-board survey will be used as the primary data source, as it will be explained in following chapters. However, to conduct accessibility analysis, geographic information data, and other relevant data are needed, as discussed next.

3.2.2 Secondary Data

3.2.2.1 TIGER Files

TIGER products are developed by the US Census Bureau. These files are spatial extracts from the Census Bureau's MAF/TIGER database, containing features such as roads, railroads, and statistical geographic areas among others (U.S. Census Bureau, 2010b). TIGER means Topically Integrated Geographic Encoding and Referencing. For this study, the TIGER/Line with Selected Demographic and Economic Data will be used. These shapefiles contain geometry and selected attributes from the 2010 Census TIGER/Line Shapefiles and the 2010 Census Summary 1 Demographic Profiles for the U.S. An important characteristic about these data is the ability to extract the files and analyze them with a geographic information system.

3.2.2.2 GTFS Data

Another important component for the access analysis is the General Transit Feed Specification (GTFS). GTFS defines a common format to describe schedule, route, stop, fare, and calendar data for fixed route transit services. This format was developed by Google for sharing public transportation information. The data will be mainly taken from open sources such as GTFS Data Exchange, Google list of GTFS data, TransitFeeds, and Transit Land. To make available the GTFS into a geographic information system such ArcGIS, a tool developed allows to use the files. That tool "Display GTFS Route Shapes" needs to be linked to the ArcGIS toolbox to display the route shapes used, while another tool "Add GTFS to Network Dataset" will allow to display the transit system (Morgan, 2014).

3.2.2.3 Smart Location Data

In order to perform the access analysis, it was necessary to consider income data for the area of analysis. To that end, the Smart Location Database was used. This database is a free data product and service provided by the U.S. EPA Smart Growth Program (US Environmental Protection Agency, 2013). This data summarized numerous socio-demographic, employment, and built environment variables for every census block group (CBG) in the United States. This database includes employment data such as number of workers in CBGs, number of working earnings in different ranges, among others, based on Census Longitudinal Employer-Household Dynamics 2010 data.

4. THE ROLE OF THE FIRST AND LAST MILE ON THE MODE CHOICE OF INTERCITY PASSENGER RAIL RIDERS

4.1 Introduction

Passenger perception is fundamental for evaluating the performance of a transit service (Eboli & Mazzulla, 2011). For example, in a national study in 1996, intercity passenger rail travelers chose rail from different transportation alternatives available to them (Drea & Hanna, 2000). The factors that made rail the most appealing mode to passengers were comfort, speed, and, cost. It is also know that the use of consumer satisfaction surveys have helped prioritizing the future quality of service enhancement initiatives and in identifying the degree of accomplishment of previous initiatives. In another study, service frequency and accessibility were found as relevant factors considered when choosing a transportation mode (Tyrinopoulos & Antoniou, 2008). The identification of the most important factors that can be used in a competitor orientation, which is defined as "that a seller understands the short-term strengths and weaknesses and long-term capabilities and strategies of... competitors" (For discussion see Drea et al (2000), page 34), is important for increasing market share. Nowadays, on interstate markets, the main competitor for Amtrak is the automobile, with air travel as a secondary competitor (primarily for business travelers).

In view of the above, this chapter attempt to address the role of the FMLM in the mode choice of intercity passenger. This question is answered by compares different modes of transportations for medium or long distance travel (more than 50 miles from home to the furthest destination, according to BTS) in terms of the distance needed to reach an origin station. Automobile-driving alone, automobile-carpool, intercity bus, intercity train, and airplane were ranked in nine attributes to identify the preference of an intercity passenger. That list of transportation modes includes the modes available in the study area that could be competing modes of the passenger rail service. The results of the analysis are presented below.

4.2 Multi-attribute Attitude Model

Multi-attribute attitude models (MAM) have been used to investigate the beliefs, attitudes, and behaviors of passenger when they are choosing a mode to travel intercity. Also, it could be used analyze whether or not the distance to reach the initial station/stop may affect the decision of the riders to use a particular mode.

In the interest of identifying the preference of an intercity passenger, the analysis was performed in different steps. Firstly, four distances ranges were created from the survey responses given to the question *"Approximately how many miles did you travel to reach the station?".* The answer, given as a continuous variable, showed how much a rider needed to travel to reach the closest HST station. Due to the configuration of the model chosen, the data was divided into ranges. The initial range analyzed was defined using quartiles of the data collected. The decision to use quartiles was taken as there was no literature that defined the distance that an intercity passenger rail would impact. Additionally the use of quartiles ensures that enough data would be analyzed for each range and enough data would be use for the comparison of unequal proportions. Owing to the data, the percentages in the quartiles were not uniform, since the thresholds between each range was considered as an integer, some of the ranges encompass more than 25% or less than 25%. From the 593 valid responses about the distance to reach a station question, each range included 29.27%, 25.04%, 24.70% and 20.98% of responses, respectively. The four ranges defined through a quartile resulted in the following distances:

- Range 1: Riders who traveled less than 2 miles to reach a station,
- Range 2: Riders who traveled more than 2 miles but less than 7 miles to reach a station,
- Range 3: Riders who traveled more than 7 miles but less than 24 miles to reach a station,
- Range 4: Rides who traveled more than 24 miles to reach a station.

After that, the base case, composed by the four ranges, was used to develop four different MA models. The MAM provide the total average score (Total Rank) that corresponds to the estimated index (Eq. 1), in addition to the decomposed scores for each attribute.

$$A_j = \sum_{i=1}^{n} b_{ij} a_i \qquad (4.1)$$

where, for each individual, A_j represents the attitude toward brand j, in this case brand will take the form of a transportation mode, b_{ij} represents the rating of mode of transportation j on attribute i, a_i represents the importance of attribute i in forming an overall attitude toward the transportation mode, and n represents the number of attributes that a person will look at.

The attributes evaluated in each transportation mode were *cost*, *travel time*, *comfort*, *safety*, *amenities*, *flexibility of travel*, *convenient*, *reliability* and *ease of traveling*. Those nine attributes were taken from Pyrialakou (2016). She identified those as attributes that can characterize the transportation modes and would be considered during the mode choice. The average importance rating of the attributes was defined by asking the passenger to rate the level of importance of each attribute from not at all important, slightly important, moderately important, very important, and extremely important. For each of the four ranges of the base cases, the importance of the attributes was found. To test whether the results of the MAM are significantly different across different ranges, a one-tailed t-test for unequal sample size and unequal variance is used. The results for range 1 in each scenario are compared with those of range 2, range 3, and range 4. The results for the four ranges in the base (quartiles) scenario are presented next. Overall, it was found that train is preferred when passengers are traveling from the first quartile (closest area). Also, safety, reliability and ease of use were commonly placed in the top three of the most important factors to choose an intercity mode.

4.2.1 Range 1

This range includes the responses of passenger that traveled less than 2 miles to reach their origin stop. For this first range, the preferred mode to travel medium distance was the train. It was found that among the other options, train had a general rank of 126.05 out of the 225 (maximum value that a mode can get). This can be seen in Table 4.1 where the drive alone option appears to be the main competitor of the intercity train as it was expected.

Table 4.1 Attitude Model scores – Range 1 Quartiles (Base case) Scenario

	Train	***Drive Alone***	***Carpool***	***Airplane***	***Bus***
Total Rank	**126.05**	**125.52**	**113.59**	**113.15**	**96.49**
Safety	17.04	12.97	12.53	16.15	11.74
Reliability	14.33	16.42	14.11	13.70	11.79
Ease of use	15.90	13.91	12.66	12.33	11.54
Cost	15.07	11.73	14.57	8.79	13.55
Convenience	10.75	17.86	14.21	12.41	10.81
Comfort	15.83	13.50	11.31	11.85	9.09
Travel time	12.02	13.93	12.81	14.41	10.35
Flexibility	10.96	17.30	13.67	11.87	9.75
Amenities	14.15	7.90	7.72	11.63	7.87

Intercity Bus was the last option chosen by the riders who traveled less than 2 miles to reach a station. Carpool was the third preferred option following by plane mode. However, the total average scores of these two modes is fairly similar.

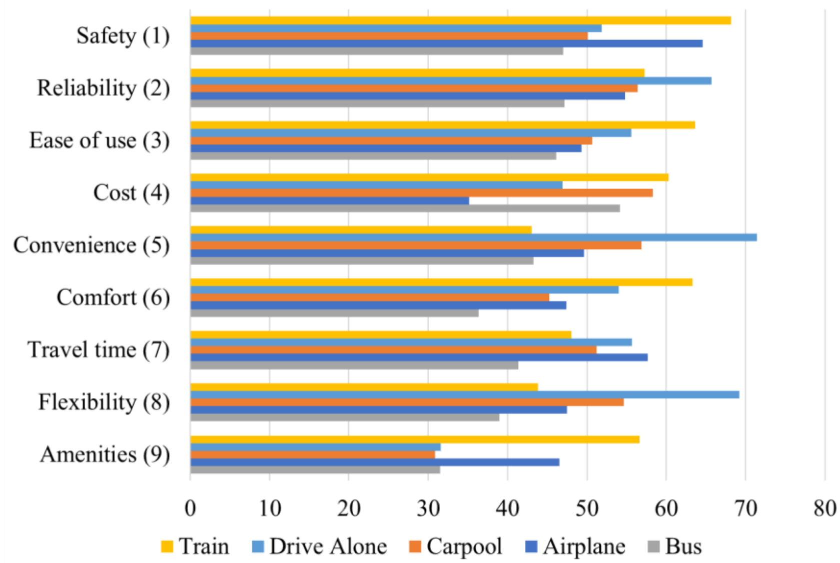


Figure 4.1 Average Score per Attribute – Range 1 Quartiles Scenario

The average score per attribute can be seen in Figure 4.1 where intercity train received high scores in *safety ease of use*, *comfort*, and *amenities*. Nevertheless, other factors such as *convenience* and *flexibility* were ranked low comparing with the driving alone option.

Table 4.2 Average Important Rating of Attributes – Range 1 Quartiles Scenario.

	Rank	***Mean score***	***St. Dev***	***Min***	***Max***
Safety	1	4.06	0.95	1	5
Reliability	2	4.04	0.87	1	5
Ease of use	3	4.03	0.82	1	5
Convenience	4	3.80	0.83	1	5
Comfort	5	3.79	0.89	1	5
Travel time	6	3.68	0.93	1	5
Flexibility of travel	7	3.66	0.89	1	5
Cost	8	3.89	0.88	1	5
Amenities (Wi-Fi, food, etc.)	9	3.50	1.04	1	5

Table 4.2 shows the importance of the attributes according to the respondents. In this case, *safety* was ranked as the most important attribute to choose an intercity mode. *Safety* is one of the attributes were train also presented a high rank. But, as showed in Table 4.1, *reliability* presented a low score for the train but it is one of the most important factors when a passenger is considering its options. *Amenities*, which is the less important factor, has a high rank value for intercity passenger train and a low value for driving alone.

4.2.2 Range 2

This range embraces the respondents that traveled more than 2 miles but less than 7 miles to reach their origin train station. In this group, driving alone was chosen as the most attractive mode of travel for medium distance trips. For this range, bus was again the most unpopular chose among the respondents. Plane and carpool were the third and fourth option chosen. Similarly, the value of the total average rank for those two modes was close.

Table 4.3 Multi-attribute Attitude Model scores – Range 2 Quartiles Scenario

	Drive Alone		***Train***		***Plane***		***Carpool***		***Bus***	
Total Rank	**132.70**	*****	**125.53**		**115.93**		**114.75**		**97.08**	
Reliability	17.96		14.67		14.14		14.95		12.53	
Ease of use	15.02	*	15.41	**	13.38		13.05		12.05	
Safety	14.24		17.26		15.95		12.98		12.06	
Convenience	18.23	*	10.97		12.41		14.11		11.13	
Flexibility	18.10		11.61		13.25		13.82		10.70	
Comfort	13.99		15.45	*	12.53		11.17		8.93	
Cost	11.62		14.24	**	8.15	*	13.48	**	12.35	**
Amenities	9.86	***	14.64		12.19		8.97	**	7.89	*
Travel Time	13.67		11.29	**	13.93		12.21		9.43	**

* Significant at 0.1 level, ** significant at 0.05 level, *** significant at 0.01 level.

Table 4.3 also shows the confidence level of the changes between range 1 and range 2 for the attributes considered in this analysis. For driving alone, the change of the value given to *amenities* between range 1 and range 2 was significant at the 1% level of confidence. In this case, the further the distance the respondents traveled to reach the station, the biggest the value given to *amenities* in a car. For intercity train, the comparison between range 1 and range 2 showed that *ease of use* had a significant change at the 5% level of confidence. In this respect, *ease of use* had a minor value for intercity train which means people who traveled more gave less value to this attribute. *Cost* in this range had a greater value for intercity comparing with range 1. This finding represents that passenger who traveled more distance to reach a station had a better perception of the *cost* for the train as intercity transportation than the ones who needed to travel less to reach a station.

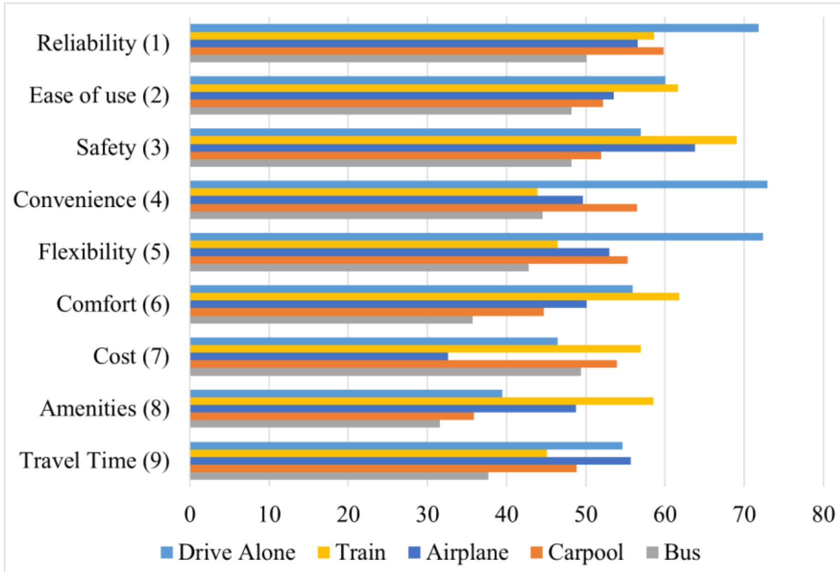


Figure 4.2 Average Score per Attribute – Range 2 Quartiles Scenario

The average score per attribute can be seen in Figure 4.2 where intercity train received high scores for *ease of use, safety, comfort, and amenities*. Nevertheless, other factors such as *convenient* and *flexibility* were ranked low comparing with the driving alone option.

Table 4.4 Average Important Rating of Attributes – Range 2 Quartiles Scenario

	Rank	***Mean score***	***St. Dev***	***Min***	***Max***
Reliability	1	4.22	0.74	1	5
Ease of use	2	4.10	0.77	1	5
Safety	3	4.02	0.92	1	5
Convenience	4	3.97	0.87	1	5
Flexibility of travel	5	3.88	0.88	1	5
Comfort	6	3.79	0.86	1	5
Cost	7	3.77	1.01	1	5
Amenities (Wi-Fi, food, etc.)	8	3.61	0.97	1	5
Travel time	9	3.57	0.97	1	5

For the respondents that traveled more than 2 miles and less than 7 miles to reach a station, *reliability* was the most important factor according to this ranking (Table 4.4). This time, the less important factor was *travel time*. For a medium distance trip, other important attributes for these respondents were *ease of use* and *safety*. Those two factors were ranked high for intercity train.

4.2.3 Range 3

Range 3 includes the respondents that traveled more than 7 miles but less than 24 miles to reach their origin station. This group of respondents chose again driving alone as their preferred option to perform a medium distance trip. Bus was again in the last position of the options. Plane and Carpool were ranked third and fourth. Table 4.5 shows the discomposed average score per each attribute. For the respondents in this range, *safety* was the most important factor to consider.

Table 4.5 Multi-attribute Attitude Model scores – Range 3 Quartiles Scenario

	Drive Alone		***Train***		***Plane***	***Carpool***		***Bus***	
Total Rank	**132.56**	*****	**130.59**	*****	**115.45**	**114.63**		**93.84**	
Safety	14.19	*	17.69		16.83	13.19		11.59	
Reliability	17.48		15.39		14.33	14.67		11.73	
Ease of use	14.60		15.95	*	11.95	12.92		11.09	
Convenient	18.90	*	12.02		12.72	14.52		10.64	
Comfort	14.42	*	15.63	*	12.54	11.58		8.61	
Flexibility	17.91		11.92		12.32	13.52		9.96	
Cost	11.84		14.72		9.07	13.48	*	12.21	**
Travel time	13.83		13.00		14.57	12.36		10.04	
Amenities	9.39	***	14.27		11.13	8.39	**	7.96	

* Significant at 0.1 level, ** significant at 0.05 level, *** significant at 0.01 level.

Table 4.5 also displays the significant changes between range 1 and range 3. The changes were tested through a t-test one tail for unequal variances. The comparison between those ranges showed that again *amenities* had a small value compared to range 1 for driving alone option at the 1% of confident level. For intercity train, *ease of use* was less ranked comparing to range 1. A similar pattern was seen when this comparison was made with range 2. Likewise, *comfort* was ranked higher for intercity train for the passenger in range 1 (less distance to reach the station) than passenger in range 3. The changes of the decomposed value for *comfort* were significant at 10%.

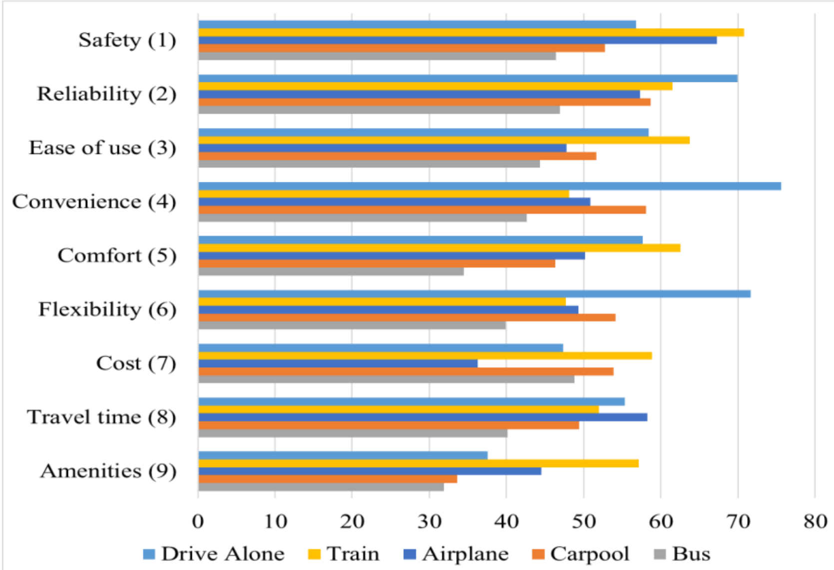


Figure 4.3 Average Score per Attribute – Range 3 Quartiles Scenario

According to Figure 4.3, *safety* was the most important attribute for the respondents who traveled more than 7 miles but less than 24 miles to reach a station. *Amenities* was again ranked as the least important. Those two factors were ranked high for intercity train.

Table 4.6 Average Important Rating of Attributes – Range 3 Quartiles Scenario

	Rank	***Mean score***	***St. Dev***	***Min***	***Max***
Safety	1	4.16	0.99	1	5
Reliability	2	4.16	0.83	1	5
Ease of use	3	4.06	0.84	1	5
Convenience	4	4.04	0.83	1	5
Comfort	5	3.91	0.90	1	5
Flexibility of travel	6	3.88	0.89	1	5
Cost	7	3.81	1.06	1	5
Travel time	8	3.73	0.90	1	5
Amenities (Wi-Fi, food, etc.)	9	3.57	1.07	1	5

4.2.4 Range 4

This range includes the riders that traveled more than 24 miles to reach their origin station. The results of the MA model showed that train was preferred among the respondents. Bus was the less desired mode between the respondents. Plane and Carpool were once more located in the third and fourth place. However, this time the difference between those two options was greater than in previous ranges.

Table 4.7 Multi-attribute Attitude Model scores – Range 4 Quartiles Scenario

	Train		***Drive Alone***		***Airplane***		***Carpool***		***Bus***	
Total Rank	**131.14**	*	**130.37**		**118.41**		**114.88**		**92.91**	
Safety	18.25		14.94	*	16.58		14.12	**	12.63	
Reliability	15.57		17.72	*	14.48		14.32		11.63	
Ease of use	15.93		13.88		13.35		12.11		10.99	
Convenient	11.16		18.06		12.73		14.05		9.86	
Comfort	16.38		14.18	*	13.19	*	11.68		8.41	
Cost	15.31		11.61		9.06		14.20		12.13	*
Flexibility	12.24		17.78		13.40		13.84		10.00	
Travel time	12.56		13.74		14.13		12.48		9.73	
Amenities	13.75	*	8.44	**	11.48		8.07		7.55	

* Significant at 0.1 level, ** significant at 0.05 level, *** significant at 0.01 level.

Table 4.7 indicates that *amenities* had a significant change for driving alone between range 1 and range 4. In this case, the further the distance traveled to reach the station, the biggest the value given to amenities for the driving alone option. *Amenities* was also identified as a significant change for train option. For the intercity train, the value given to *amenities* was higher when people needed to travel less to the reach a station.

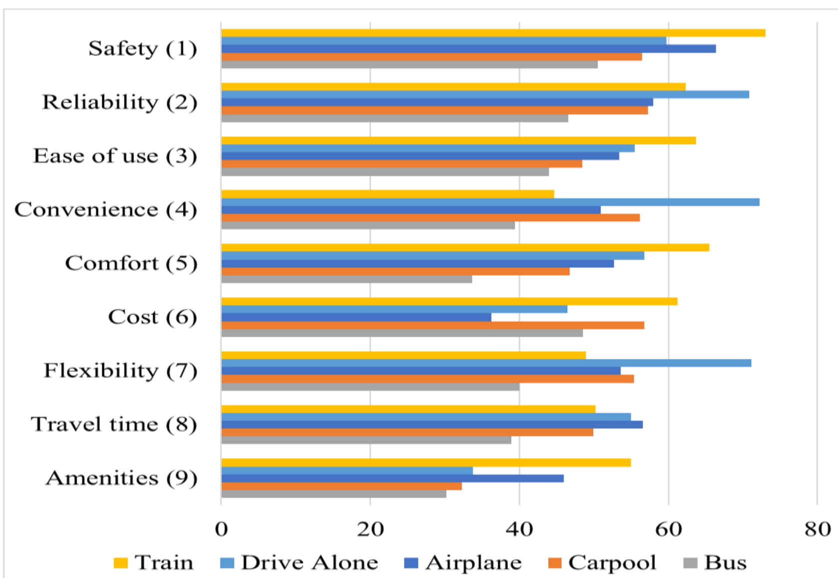


Figure 4.4 Average Score per Attribute – Range 4 Quartiles Scenario

The average score per attribute can be seen in Figure 4.4 where intercity train received high scores for *safety*, *ease of use*, *comfort* and *amenities*. Nevertheless, other factors such as *convenient* and *flexibility* were ranked low comparing with the driving alone option.

Table 4.8 Average Important Rating of Attributes – Range 4 (Base case)

	Rank	***Mean score***	***St. Dev***	***Min***	***Max***
Safety	1	4.28	0.85	1	5
Reliability	2	4.15	0.86	1	5
Ease of traveling	3	4.09	0.91	1	5
Convenience	4	3.96	0.86	1	5
Comfort	5	3.93	0.88	1	5
Cost	6	3.88	0.98	1	5
Flexibility of travel	7	3.84	0.91	1	5
Travel time	8	3.62	1.00	1	5
Amenities (Wi-Fi, food, etc.)	9	3.46	1.06	1	5

When choosing intercity mode of transportation, safety was the most important factor for respondents who traveled more than 24 miles. This factor was ranked high for intercity train. Nevertheless, other factors such as *reliability* and *convenience* were important for the riders but they did not have a high rank for the intercity train.

4.2.5 Comparison of Findings

Comparing all the ranges in the quartiles analysis, Figure 4.5 presents the decomposed MAM scores of the intercity train and drive alone alternatives, and the average importance score for each of the nine attributes for the four ranges of the base case scenario. To enable a comparison, the scores for both the attributes and the importance of the attributes have been brought to a common scale from a 0 to 100. The transformed scores now denote the percent of the maximum possible score (where the maximum possible importance score is 5 and the maximum possible decomposed score is 25). The most/least important factors and the perceived performance of the two alternatives with respect to those factors can be identified from the figure. In addition, because of this transformation, the transformed importance scores also correspond to the transformed maximum possible decomposed scores, given the importance score of the specific attribute. For example, the average importance score of the attribute "amenities" was 3.5. The transformed value was $((3.5-1)*100)/(5-1) = 62.5\%$. Because the maximum attribute rating is 5, the maximum possible decomposed score given the importance score would be 3.5*5=17.5 and the transformed maximum possible decomposed score would be (17.5*100)/25=70%.

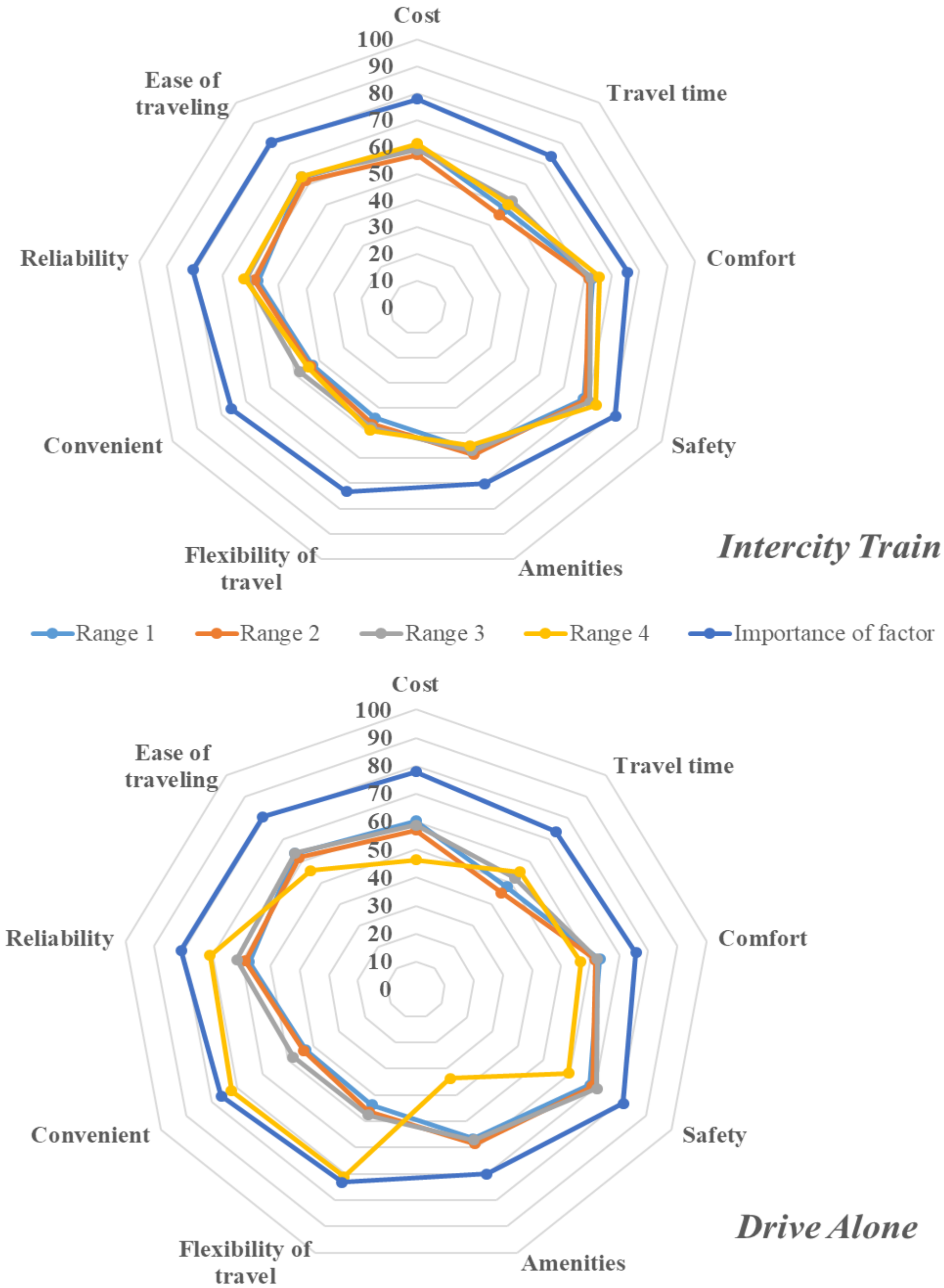


Figure 4.5 Average scores per attribute due to different ranges for Intercity Train and Driving Alone (scores are displayed as a percent of maximum possible score).

The findings suggest that *reliability* is important when choosing an intercity transportation mode. However, as Figure 4.5 shows, riders do not find intercity train a very reliable transportation mode. *Safety* was also identified as an important attribute considered in intercity mode choice decisions. For *Safety*, intercity train obtained a value close to the maximum possible value, and therefore, respondents had a positive perception of the train with respect to this factor. Generally,

Figure 4.5 shows that the train is considered a safe and relatively comfortable and with many available amenities mode, but has a lot of room for improvement in terms of convenience, flexibility of travel, and travel time. The findings also suggest that the scores are very similar for all ranges.

Figure 1 also shows the average value of the nine attributes for each different range for the driving alone options. For distances up to 24 miles to access the station, the scores of this alternative are similar to the trains. Interestingly, however, respondents who traveled more than 24 miles to access the station seemed to have a very different attitude towards driving alone. As Figure 1 shows, *amenities* had a very low score and safety a relatively low score compared to the maximum possible scores. Contrary, the alternative was ranked close to perfect with respect to *reliability*, *flexibility of travel*, and *convenience*.

4.2.6 Sensitivity Analysis

A sensitivity analysis was conducted to identify any changes in the MAM due to the changes in the ranges that are analyzed. All the alternative scenarios were tested against the quartiles (base case) scenario. Sixteen multi-attribute attitude models were developed and they can be seen in Appendix C. To ensure that the four ranges had at least 10% of the observations, the maximum distance decrease or increase from the quartiles scenario is chosen to be 50%. The scenarios considered include a 33% decrease and increase from the quartiles scenario (Scenarios 1 and 2 respectively) and a 50% decrease and increase from the quartiles scenario (Scenarios 3 and 4 respectively). Based on these percentage changes, the following four scenarios are used for the sensitivity analysis

- Scenario 1: 0 to 1.5 miles, 1.5 to 5.5 miles, 5.5 to 19.75 miles, and greater than 19.75,
- Scenario 2: 0 to 2.5, 2.5 to 8.5 miles, 8.5 to 28.25 miles, and greater than 28.25 miles,
- Scenario 3: 0 to 1 miles, 1 to 4 miles, 4 to 15.5 miles, and greater than 15.5 miles, and
- Scenario 4: 0 to 3 miles, 3 to 10 miles, 10 to 32.5 miles, and greater than 32.5 miles.

The results of those models showed that, the scores of some mode choice alternatives changed due to the variation in the distance considered in the range. The importance scores of the attributes also changed in some of the new scenarios. To test if those changes were significant, *t*-tests of unequal sample size and unequal variance were considered.

The results of the sensitivity analysis indicated, for the most part, that the MAM scores were not particularly sensitive to different range distances. The patterns emerged pertaining to the intercity train from the base case analysis (discussed in the previous section) were similar to the ones identified in the four scenarios considered in the sensitivity analysis. Therefore, the sensitivity analysis generally supported the choice of the quartiles (base case) scenario for the analysis of this thesis. It is noted that some differences in the order of preference of the modes were found in some of the ranges in a few scenarios. For example, respondents who traveled more than 32.5 miles rated intercity bus (which, in most ranges of all scenarios, has been consistently ranked as one of the least favorable alternatives) as the most attractive option for a medium distance trip. Those differences, however, where not consistent across the different scenarios considered.

4.3 Concluding Remarks

The main objective of this analysis was to identify how access to a passenger rail line is related to transportation mode preferences and mode choice factors. To that end, a multi-attribute attitude model (MAM) was estimated, and the results provided the total average score (total rank) that corresponds to the estimated index. Indices with a higher value are seen as the more attractive mode (or the more favorable the attitude towards the mode). The stated distance to access the station was first divided in quartiles ranges (from 0 to 2 miles, 2 to 7 miles, 7 to 24 miles, and more than 24 miles) to analyze the respective changes in the MAM index. The findings suggested that traveling by intercity train and driving alone were the most preferred modes for medium distance trips. This finding was anticipated for two reasons: first, because the survey was conducted on-board the HST and the respondents had already chosen to travel by intercity train when they were surveyed. Therefore, their preference of intercity passenger rail over other competing modes was expected. Second, it was also expected that driving alone would be one of the most preferred ways to travel because Indiana is generally an automobile-oriented state. For instance, data suggests that approximately 76 percent of U.S. commuters chose to drive alone in 2015 (Bureau of Transportation Statistics, 2015).

The use of distance serves as an indicator of access to the station and helps to comprehend the first and last mile journey. The results of the analysis suggested that intercity train is the most favorable mode for riders who traveled less than two miles to access a station. This finding implies that people who traveled less to access a station would be more likely to take the train, if they had the chance to do so. Furthermore, for these respondents (i.e., with high levels of access to the station), safety, ease of use, and reliability were identified as the most important factors in mode-choice decision making for intercity travel. The order of importance of these highly ranked three attributes and the rest six attributes varied as the level of access to the train varied. For the most part, however, the riders' opinions of the train's performance with respect to these attributes were similar regardless of how much they traveled to access the line. However, this was not the case for riders' perceptions of driving alone. Specifically, riders with the lowest level of access to the line (the ones who traveled more than 24 miles to a station) thought that driving alone was more difficult and less safe, but more reliable compared to the riders with the higher level of access to the line. The fact that driving alone obtained a low scores in those of the most important factors to choose a mode such as easy of travel and safety, make the train the preferred mode of passengers that were traveling from more than 24 miles. Another finding worth noting was that cost and travel time were not perceived as important attributes to consider when choosing a mode.

5. FACTORS AFFECTING THE FREQUENCY OF TRAVEL BY INTERCITY RAIL

5.1 Introduction

Apart from availability of and access to services, as studied in the previous chapter, there are different factors that make a transportation system more competitive and appealing to passengers, such as fare level, service frequency, the quality of the waiting environment, and in-vehicle amenities (Tilahun et al., 2016). Although, improving access to the stations (how to get to and leave from a station) might be cheaper, and overall, more cost effective than improvements to the actual train journey (Givoni & Rietveld, 2007), identifying the factors that affect the usage of intercity passenger rail can also help enhance the usage of passenger train service and potentially attract more riders. Understanding the factors that influence intercity travel is important for long-term transportation planning and to support quality of life in the study area for a changing population (Neely, 2016).

This chapter aims to examine the relationship between frequency of travel by intercity rail and (i) mode choice related factors, (ii) factors associated with access to a rail line, and (iii) passenger characteristics. To that end, this chapter presents an Ordered Probit model estimated to explore factors associated with trip frequency in intercity passenger rail service. This type of model was chosen due to the nature of the dependent variable, frequency of travel, which was recorded in the survey instrument as an ordinal variable.

5.2 Ordered Probit Model

Trip frequency is of count nature when it refers to the number of trips. However, frequency can also be captured in an ordinal variable when represented in discrete categories. The dependent variable herein (stated trip frequency in the year prior to the survey) is ordinal with seven response categories: 0 trips, from 1 to 2 trips, from 3 to 4 trips, from 5 to 6 trips, from 7 to 8 trips, from 9 to 10 trips, and more than 10 trips taken on the HST in the year prior to the survey (since August 15th, 2015) when a single trip counts as one trip and a round trip counts as two trips. Because of the discrete and ordered nature of the data collected, an ordered probit model is estimated to identify whether the factors described in the MAM (Chapter 4), apart from mode decisions, affect the trip frequency of intercity passenger rail riders. Ordered probit models are a broadly used approach to model ordinal variables (Jackman, 2000).

A model is derived by defining an unobserved variable *z* for modeling the ordinal ranking of data. This unobserved variable is usually specified by a linear function characterized by the following expression:

$$z = \beta X + \varepsilon \qquad (5.1)$$

where *X* is a vector of variable determining the discrete ordering for observation *n*, *β* is a vector of estimable parameters, and *ε* is a random disturbance (Washington, Karlaftis, & Mannering, 2010). Using Eq. 2 and assuming the *ε* is normally distributed across observations with mean equal to zero and variance equal to 1, an ordered probit model can be established with the probability of each category being selected characterized by the expressions

$$P(y = 1) = \Phi(-\beta X)$$
$$P(y = 2) = \Phi(\mu_1 - \beta X) - \Phi(\beta X) \qquad (5.2)$$
$$P(y = I) = 1 - \Phi(\mu_{x-1} - X)$$

Where *Φ* is the cumulative normal distribution. The variable *z* is characterized by the expression

$$z = \beta X_{ic} + \varepsilon_{ic} + \sigma_i = \qquad (5.3)$$

Denoting *i* as individuals and *c* as the observations created by each individual.

For the interpretation of the intermediate categories, marginal effects are computed. The marginal effects will provide the direction of the effects on the interior categories. The effects are computed as the difference in the estimated probabilities with the indicator variable changing from zero to one for the indicator variables, while the other variables are equal to the mean. The marginal effects are understood as a change in the outcome probability for each threshold category P(y=j) given a unit change in a variable x.

Defined the model to estimate, the first step developed was to calculate the histogram of the dependent variable in order to evaluate the distribution and number of rated values for each one of the five categories. As a rule of thumb, a minimum of 10% of observation for each category was needed to perform the analysis, however, there was not a sufficient number of observations for each category. Hence, merging of the individual categories was required for this analysis (Figure 5.1). Three categories of responses were considered. The first category described the respondents who took the HST for the first time at the time of the survey and had zero previous trips on board the HST the year prior to the survey. This category would be referred to as *new (or* *low frequency) riders*. The second category described passengers who took the HST one, two, three or four times in the year prior to the survey. This category describes the *medium frequency riders*. Finally, the last category describes the *high frequency riders*, i.e., respondents who traveled five times or more with the HST in the previous year. The decision to separate those categories was the fact that a one-way trip was defined as one trip and a round trip counts as two trips. In that sense, riders who have traveled equal to or less than two round trips are getting used to the train, however, they are no longer new passengers but they cannot be classified as frequent riders either. The histogram of responses for this new categorization is presented in Figure 5.2.

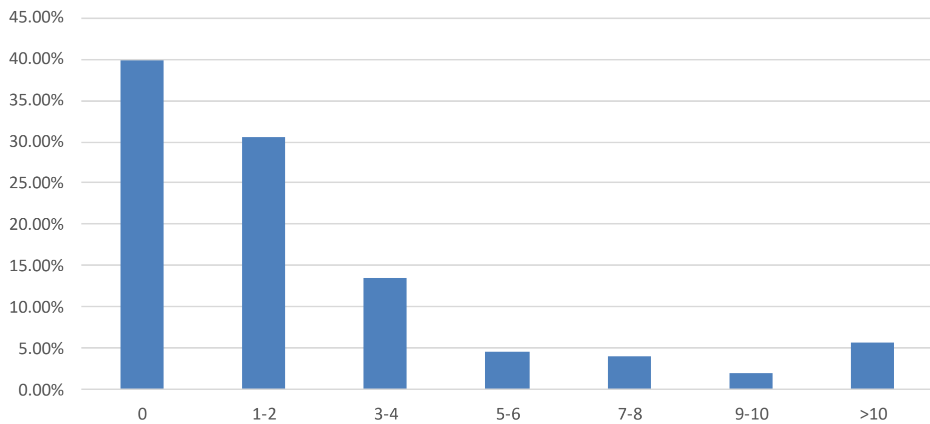


Figure 5.1 Histogram trip frequency in the previous year.

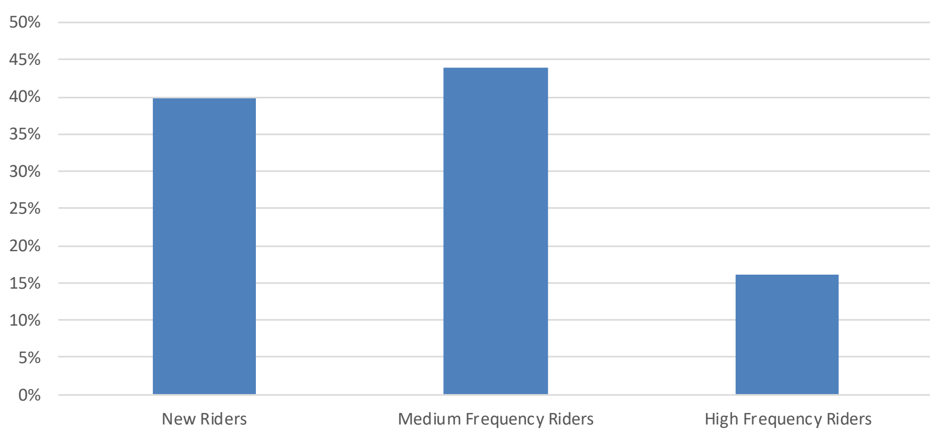


Figure 5.2 Histogram new categories for trip frequency in the prior year.

The descriptive statistics of other significant variables for this analysis are presented in Chapter 3. Table 3.2. Some of these indicator variables such as “Respondent prefers train over personal vehicle” were created by the author from the mode choice part of the survey to provide an enhanced understanding of the results of the model.

5.2.1 Results

An ordered probit model was estimated to identify attributes/factors that are associated with frequency of travel by intercity passenger rail. The model was estimated in NLOGIT 6.0 and the results are presented in Table 5.1. As Table 5.1 shows, the model had a low McFadden pseudo-r squared. This measure cannot be interpreted the same way as an r square (R^2) value in ordinary least square (OLS) regression because it is helpful for comparing models using the same sample, trying to estimate the same outcome (Louviere et al. 2000). However, the count R squared was higher (46%), demonstrated that the model was able to accurately predict the responses of approximately half of the observations. The NLOGIT Outputs can be found in Appendix D.

Table 5.1 Ordered Probit Estimation Results

Variable	**Parameter Estimate**	**Std. Error**	**p-value**
Constant	0.4685	0.0856	0.000
Respondent walked to reach a station	0.0005	0.0003	0.0881
Respondent who had origin at Indianapolis station and someone drop his/her off to access the station	-0.2756	0.0836	0.0010
Respondent who had origin at Chicago station and use public transportation	0.1414	0.0811	0.0812
Respondent prefers train over personal vehicle	0.0002	0.0003	0.4442
Respondent prefers personal vehicle over to airplane	-0.0009	0.0003	0.0023
Respondent believes train is safer than personal vehicle	0.0004	0.0003	0.0976
Respondent believes train is easier to travel with than personal vehicle	0.0004	0.0003	0.0876
Respondent took the HST for work or school	0.3919	0.1140	0.0006
Household size of three or greater	-0.2217	0.0830	0.0075
Respondent graduated from high school or less	-0.1875	0.1138	0.0995
Respondent drives more than 100 miles weekly	-0.1988	0.0977	0.0418
Threshold parameter for index			
Mu	1.3196	0.0567	0.0000
McFadden pseudo-R squared 0.0408 Count R squared 46% Log likelihood function -861.0075 Number of observations 879		Restricted log likelihood	-897.6301

*For all the variables 1 if yes, 0 if otherwise

Table 5.1 also shows that several factors related with a passengers' access to the route appear to be associated with trip frequency. Specifically, the findings suggest that respondents who reached the station on foot are more likely to be frequent riders. The ability to walk to the station implies that the respondent has access to a station close to their origin and no additional cost was added to their intercity trip. In addition, the findings suggest that riders who were dropped off by someone else in Indianapolis, one of the main origin station in this line (as seen in Table 3.2–Chapter 3), to take the train are less likely to be frequent riders. These finding may bring up two access issues: Indianapolis, is mainly an automobile oriented city (like most of the U.S. cities according to (U.S. Census Bureau, 2015)), where there is not a high coverage of public transit and people have low preference for public transit. Another access issue is that the HST departs at 6 am from the Indianapolis station, and as such, depending on the passenger's origin, it might not even be possible to use a bus to access the station because of the hours of transit operation. On the contrary, the findings of the model suggest that respondents who had origin at Chicago, a large city with a mature multimodal transportation system, and used public transportation to reach the station (bus, metro, Metra) are more likely to be frequent riders. The first and last mile problem to access the train station might not be as severe in cities similar to Chicago.

Apart from the factors directly related with access to the line, a number of factors pertaining to the transportation mode preferences and the attributes considered in mode choice decisions were explored in this model. Specifically, the MAM total rank index was used to compare the mode preferences. The findings suggest that respondents who preferred the train over driving alone are more likely to be frequent riders (though the variable was not strongly significant). Another finding was that respondents who prefer driving alone compared to airplane were less likely to be frequent riders. Furthermore, some of the decomposed MAM scores were found significantly associated with trip frequency. Specifically, the results showed that people who thought that train is safer and easier to use than driving alone are more likely to be frequent riders of the HST. The results of the MAM analysis suggested that *Safety* and *Ease of use* are two of the most important factors that respondents considered for their mode choice decisions for a medium distance trip. Therefore, it was anticipated that those factors would be significantly associated with frequency of travel by intercity passenger rail as well.

Finally, various demographic and travel behavior related characteristics of the passengers were found to be significantly associated with trip frequency. Specifically, the findings suggest that respondents who travel for work and school are more likely to use the train (medium or high frequency users). In a different question in the survey, most of the respondents stated that traveling on board the HST allows them to use their time more productively, which can be a drive for frequent travel of students and commuters. Also, this specific line serves different college towns and two main cities in the Midwest. Furthermore, the findings suggest that respondents in a household size larger than three people are less likely to travel frequently on the HST. A separate question on the survey showed that people were 37% neutral and 38% in disagreement when they were asked if traveling with a group (family, friend, etc.), using the HST to reach the destination would cost them less. That could be an explanation for the sign of this variable. Furthermore, respondents who were high school graduates at most and respondents who usually drive more than 100 miles weekly are also less likely to be frequent riders. Given that the HST is 196 miles long, a person able to drive more than a 100 miles a week would probably consider driving the entire way between Indianapolis and Chicago (181.5 miles).

5.2.1.1 Marginal Effects

In terms of evaluating the effect of the individual variables on the ordered Probit model a first but ambiguous evaluation can be developed by just observing the signs of the coefficients. It provides a general behavior on how the probability increases or decreases in the extreme categories. However, as Mannering et al. (2011) notes, a practical difficulty with ordered probability model is associated with the interpretation of the intermediate categories. Depending on the location of the thresholds, it is not necessarily clear the effect of a positive or negative β has on the probability of these interior choices. For addressing this problem, marginal effects can be calculated for each category. The marginal effects for the final model are presented in Table 5.2.

Table 5.2 Marginal Effects

Variable	**New Riders Y=1**		**Medium Frequency Riders Y=2**		**High Frequency Riders Y=3**	
Respondent walked to reach a station	-0.0002	*	0.58D-04	*	0.0001	*
Respondent who had origin at Indianapolis station and someone drop his/her off to access the station	0.1046	***	-0.0386	***	-0.0659	***
Respondent who had origin at Chicago station and use public transportation	-0.0531	*	0.0180	*	0.0351*	*
Respondent prefers train over personal vehicle	-0.0002		0.27D-04		0.5144	

Respondent prefers personal vehicle over to airplane	0.0003	***	-0.0001	***	-0.0002	***
Respondent believes train is safer than personal vehicle	-0.0002	*	0.57D-04		0.0001	
Respondent believes train is easier to travel with than personal vehicle	-0.0002	*	0.57D-04	*	0.0001	
Respondent took the HST for work or school	-0.1390	***	0.0298	***	0.1092	***
Household size of three or greater	0.0840	***	-0.0307	**	-0.0534	***
Respondent graduated from high school or less	0.0719		-0.0287		-0.0433	*
Respondent drives more than 100 miles weekly	0.07604	**	-0.02960	*	-0.04644	**

For all the variables 1 if yes, 0 if otherwise. *,**,*** Significant at 0.1 level, at 0.05 level, and at 0.01 level.

Table 5.3 Cross tabulation of predictions for the ordered probit model

`	Predicted Rank			
Actual	**0**	**1**	**2**	**Row sum**
0	42 (49%)	199(36%)	0 (0%)	241
1	36 (42%)	245(44%)	0 (0%)	281
2	8 (9%)	107 (19%)	0 (0%)	115
Column Sum	86	551	0	637

For the *Medium frequency Riders* category it was observed that passengers were more likely to ride the train if they access the station by foot. In this case, the sign of the coefficient followed the pattern showed by *High frequency riders.* Also, respondents who had origin in auto-dependent cities such as Indianapolis, were less likely to be *Medium frequency riders* as this factor was found significant at the 1% confident level. Contrary, passengers who departed from more multimodal transportation cities, and were described by the intermediate category, were also more likely to ride the train as the one described as *High frequency riders.* Likewise, the rest of the results in the intermediate category were alike with the results showed by the *High frequency riders'* category exposing that it is more likely that those *Medium frequency riders* would be part of that extreme category. The cross tabulation of predictions for different categories showed that medium frequency riders category would be better explained by the model than any other category considered. Also, the high frequency riders' category could not be explained by this model and further exploration should be considered to evaluate the factors that would predict better the amount of travel by high-frequency riders.

5.3 Concluding Remarks

An ordered probit model was estimated in order to recognize the factors affecting the level of usage of intercity rail passengers. Three categories were analyzed according to the frequency of travel in the year previous to the survey. New riders, medium frequency riders, and high frequency riders were the categories considered in this model. The results of the ordered probit suggested that safety and ease of use are important factors for intercity trip frequency as well, along with the mode used to access a station and participants' socio-demographic characteristics. The behavior of the internal category was evaluated using the marginal effects. This analysis showed that the medium frequency riders present a trend alike with the high frequency riders. In addition, some access-related variables seemed to be associated with trip frequency, though variables directly capturing this factor (such as mode use to reach the station) seemed to be overshadowed by other unobserved factors. One reason for this finding (or lack thereof) is that the frequency of travel by an intercity train might be more strongly associated with factors affecting the travel frequency itself (by whichever mode) rather than the choice of rail over another transportation alternative. Given that the percentage of induced travel for intercity trips is not expected to be very high, lack of access to the train would probably affect mode choice decisions and not decisions of taking a trip or not.

6. ASSESSING THE FIRST AND LAST MILE PROBLEM IN INTERCITY PASSENGER RAIL SERVICE

6.1 Introduction

The term first and last mile (FMLM) of a trip has been used to describe passenger travel with regards to getting to and from transit stops/stations. The FMLM problem has been addressed in different public transit contexts, mainly within urban areas. However, limited research efforts have been undertaken to examine the FMLM problem related to intercity passenger rail. The survey findings of this thesis indicated that there are passengers who travel from counties far away from a county with a station to take the train, which is discussed in Section 3.2.1. Moreover, it was found that most of the respondents either drove their own vehicle, rented a car, or were dropped off to reach a train station in Indiana. Unlike the results for passengers who boarded HST at the Chicago station, a small percentage of passengers who boarded the train at one of the five stations in Indiana used ridesharing services or public transportation. These findings suggest that there is a gap in the FMLM travel options for intercity rail passengers. Solving the FMLM problem would expand access to transportation systems and increase the number of possible passengers from remote communities, such as rural areas.

This chapter discusses the results of an accessibility analysis for the state of Indiana conducted to identify the areas in need of FMLM service where there are no public transportation services and reaching a station from a desired origin is expensive. To that end, a cost surface for the different modes available in the area of study was created to determine the average travel cost to the nearest station. The analysis was carried out in ArcGIS using the origin-destination area identified from an HTS on-board survey, transportation network information from the U.S. BTS, and general transit feed specification data from Google developers (Google Transit, 2016). Subsequently, some of the best strategies that were identified in the literature for addressing the FMLM problem were modeled around the stations (e.g., buses to/from the station, ridesharing) to examine how the accessibility would change after the implementation of a selected strategy. An area of 60 miles around the stations was considered for passengers who “drove alone” or were “dropped off/picked up by someone” at a station. For the walking option, a buffer of two miles around the station was considered. Further, the access to Lafayette and Indianapolis rail stations using public transit and ridesharing services were also examined in detail.

6.2 Methodology

Access has been measured in different ways, considering cost and travel time among other factors, to reach core services such as stops/stations, health centers, and healthy food. This thesis builds on available methodologies and expands on some limitation of the methods. For instance, access typically has to do with proximity to service; however, estimation of the access has used more complicated methods that are hard to replicate due to their complexity and data availability (Biba, Curtin, & Manca, 2010; Murray, 2003; Wu & Murray, 2005). Further, access to a service such as a health center, using different modes of transportation, has been studied in the past. However, some of these studies did not consider the cost of operating a motor vehicle; and access was also measured in terms of the frequency of transit service around the destination (Burns & Inglis, 2007; Mao & Nekorchuk, 2013). Moreover, other studies considered the use of minimum wage data to calculate the value of the time traveled which underestimates the actual transportation cost but may serve as measure of equity and as a guidance for policy and planning applications. (Bailey, 2015).

Considering the above past literature, the methodology proposed in this thesis allows estimating the cost to arrive at the closest station from census block groups (CBGs) by different travel modes including average hourly data. The modes considered are those reported in the on-board survey. Different cost rasters were created to analyze driving alone, being dropped off by someone else, walking, using public transit, and using a ridesharing system, such as Lyft (Figure 6.1). The cost calculations in ArcGIS are defined in the subsequent sections and can be easily replicated in the future using updated files available.

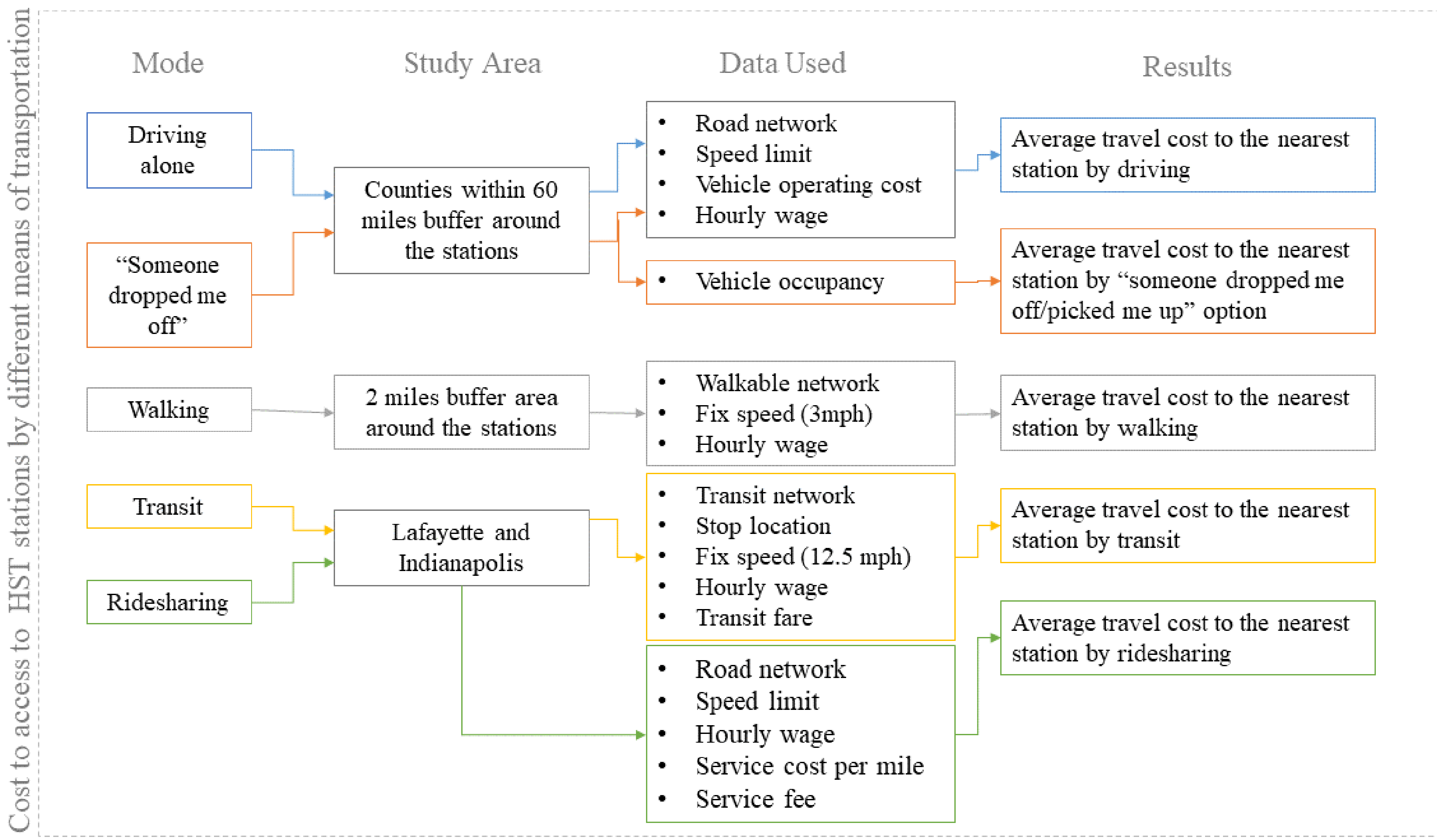


Figure 6.1 Methodology for Estimation the Cost to Access the HST Stations.

6.2.1 Data and Preliminary Steps

At the beginning of this analysis, the most recent year for which data were available was 2015 so most of the data references that year. The one exception is the hourly wage data that was taken from Smart Location Database which is based on the 2010 Longitudinal Employer-Household Dynamics (LEHD) Census. These data were updated by using the Consumer Price Index for 2015. The use of hourly wage data has not been explored before. Past studies considered the minimum federal wage to provide the value of the time spent in the different modes (Bailey, 2015; Burns & Inglis, 2007).

In this thesis, the hourly wage was estimated by using data from the LEHD Census at the CBG level. These data provide the number of workers earning between three ranges of monthly wages per month. It also includes the number of workers per CBG. The three ranges (low: $1,250 or less; medium: more than $1,250 but less than $3,333; and high: more than $3,333) were used to calculate the average earnings per CBG. In that sense, for the low range, the average value or earnings per month was taken as $1,250, for the medium range $2,291, and for the high range $3,333. The average values from the low and high ranges were taken as the maximum and minimum of each in order to compensate for over- and under-estimations. The hourly wage then was calculated by dividing the value obtained previously by the average number of hours worked monthly in Indiana. According to the U.S. Bureau of Labor Statistics (BLS), Indiana workers typically work 38.7 hours a week or approximately 154.8 hours a month (U.S. BLS, 2016).

The values that resulted from the previous step were $8.07 per hour for the low range, $14.80 per hour for the medium range, and $21.53 for the high range. As each CBG has workers from each of the three ranges, the average hourly wage of the CBG was calculated with the following formula:

$$Hourly\ wage = \frac{(a*Low)+(b*Med)+(c*High)}{TotalWK} \qquad (6.1)$$

where *Low* represents the number of workers classified in the low range income, *Med* represents the number of workers classified in the medium range of income, *High* represents the number of workers in the highest range of income, a=$8.07 per hour, b=$14.80 per hour, c=$21.53 per hour, and TotalWK represents the total number of workers at the CBG. This value was unique for each CBG in the analysis.

According to the U.S. BTS, the average cost per mile of operating a motor vehicle was 57.1 cents per mile in 2015 (BTS, 2016). This value was used to calculate the driving alone cost and the carpooling cost to the nearest station, as explained in the next sections.

An important part of this analysis came from the use of TIGER/Line files from the U.S. Census Bureau 2010. (U.S. Census Bureau, 2010c). These files served to build the road network of primary and secondary streets that was used to perform the spatial analysis. Other files were obtained through the IndianaMap website and included the railroad and train station locations and the county boundaries. The 2010 census urban/rural indicator also was used. The U.S. Census Bureau's urban-rural classification identifies the individual urban and rural areas in the U.S. According to this classification, areas must meet minimum population density requirements to be classified as urban. The file used in this analysis uses R or U as the indicators to show whether the CBG is classified as rural or urban (U.S. Census Bureau, 2010a). All the shape files obtained from the different sources were projected from their original geographic coordinate system (either UTM or D North American 1983) to a projected coordinate system (NAD 1983 Indiana State Plane East FIPS 130, meters) to perform the accessibility analysis. A summary of the data sources is presented in Table 6.1.

Table 6.1 Data Sources

Data	Source	Year
County Boundaries	U.S. Census Bureau 2010 (U.S. Census Bureau, 2010)	2010
Road and Walkable Network (TIGER/Line files)	U.S. Census Bureau 2010 (U.S. Census Bureau, 2010c)	2010
Rural/Urban Information	U.S. Census Bureau 2010 (U.S. Census Bureau, 2010a)	2010
Average cost per mile of operating a motor vehicle	Bureau of Transportation Statistics (U.S. Bureau of Transportation Statistics, 2016)	2015
Hourly Wage by Census Block Group	Longitudinal Employer-Household Dynamics (LEHD) Census (U.S. Environmental Protection Agency, 2013)	2010
Transit routes and stops (GTFS)	Google Transit Data Feed (Google Transit, 2016)	2015
Rail stations and line	National Transportation Atlas Database (NTAD) (Bureau of Transportation Statistics, 2015)	2015

Transit Fare Information	CityBus website and IndyGo website (CityBus, 2017; IndyGo, 2017)	2015
Transit Speed Information	Indy Go (IndyGo, 2010)	2010
Lyft Cost Information	Lyft Website (Inc, 2017)	2017

6.2.2 Average Travel Cost Estimation to the Nearest Station for "driving alone" and "Someone dropped me off/picked me up" options

The use of ArcGIS allowed achieving most of the following processes within the *Spatial Analyst* toolbox. Other tools used to perform this analysis were found in the *Conversion* toolbox. To initiate the analysis, all the shape files were referenced to the same coordinate system using the *Project* tool. The driving alone and someone dropped me off/picked me up analyses were performed around 60 miles from each station. The value of this buffer was taken from the survey responses where passengers stated that they traveled 60 miles, sometimes more, to reach a HST station (Figure 3.2).

In order to calculate the cost to reach a station by driving alone, the *Cost Distance* tool in ArcGIS Pro was used. To begin the process, the road network shape file field labeled MTFCC, which indicates the road type classification, was used to determine the speed limit of the road (refer to Table 6.2). The vehicular trail, walkway/pedestrian trail, private service vehicle road, internal U.S. Census Bureau use, and bike path or trail were excluded from this analysis.

Table 6.2 Road Type Classifications (U.S. Census Bureau, 2015)

MTFCC	***Feature Class***	***Description***	***Speed Limit (mph)***	***Drive time (0.001 hour/mile)***
S1100	Primary road	Generally divided highways distinguished by interchanges	55	11
S1200	Secondary road	Main arterial	40	16
S1400	Local neighborhood road, rural road, city street	Paved non-arterial road, usually 2-lane	30	21
S1500	Vehicular trail	Unpaved dirt trail	--	NoData
S1630	Ramp	Entry to or exit from limited access road	25	25
S1640	Service drive	Gives access to structures along a limited-access highway	30	21
S1710	Walkway/pedestrian trail	Restricted from vehicular traffic	--	NoData

S1730	Alley	Service road generally at the rear of buildings	10	62
S1740	Private service vehicle road	Privately maintained for service purposes	--	NoData
S1750	Internal U.S. Census Bureau use	Internal U.S. Census Bureau use	--	NoData
S1780	Parking lot road	Main vehicular route through a paved parking area	10	62
S1820	Bike path or trail	Restricted from vehicular traffic	--	NoData

After this classification system was defined, the road network shape file was transformed to a raster by using the *Polyline to Raster* tool. This raster was reclassified in order to convert each cell to a cost value; however, in this analysis, this value was indicated in time. The values used to reclassify this raster depended on the first part of the formula given below and are given in Table 6.2 according to the type of road. The second part of that formula addresses the value of the vehicle operating cost (VOC):

$$TC_d = \frac{\frac{100}{s}}{1609.34} XHourlyWage + \frac{c}{1609.34} \qquad (6.2)$$

where TC_d= transportation cost by driving or being dropped off in cents per meter, Hourly Wage= average hourly wage by CBG zone in dollars, s=speed limit in mph, c=cost of operating a motor vehicle in cents, and 1609.34 is a conversion factor between meters and miles. Since the hourly wages were per CBG, the least accumulative travel time to reach a station from a desired point was calculated as a cost surface by using the *Cost Distance* tool. The units assigned to the cost raster can be any type of cost desired: dollar cost, time, energy expended or a unitless system that derives its meaning relative to the cost assigned to other cells (ESRI, 2016). In this case, after having the results for the cost distance analysis, these values were assigned to the specific zone by using the *Zonal Statistics* tool and then were multiplied by each CBG average hourly wage. A similar analysis was performed to identify the cost of operating a motor vehicle from a particular zone to the station. For the operating cost, the value of the raster was fixed at 35 (0.001 cents/m) due to an integer number being needed to execute the *Cost Distance Tool.* After obtaining the least accumulative cost to reach the station, these results were plugged into the CBG areas by using the *Zonal Statistics* tool. Once both values were converted to the cost (in dollars), they were totaled, and that result provided the estimated cost to travel from each of the CBGs analyzed to the nearest

HST station. The parking cost was not considered in the initial analysis, but it was considered in a subsequent analysis when parking was assumed to be $17/day. This value was taken from the daily average parking cost around the Indianapolis railroad station (Amtrak, 2017).

For the someone dropped me off/picked me up analysis, the same methodology was used to calculate the least travel time to get to a station and the least accumulative operating cost. However, as this trip would involve at least two people in the car, it was assumed that two people were part of this trip. Then, the travel time spent to get to the station was multiplied by two average hourly wages. The results of the previous steps were totaled with the cost obtained for operating a motor vehicle, which resulted in the estimated cost to travel from a CBG to the nearest HST station.

6.2.3 Average Travel Cost Estimation to the Nearest Station by Walking

For the walking alternative analysis, all the travel speeds were assumed to be the same (3 mph, according to (Bailey, 2015; Burns & Inglis, 2007)); however, the pedestrians and bicycle paths were included and interstates were omitted from the analysis. All other road types that were not considered in the driving analysis were also omitted in this analysis. Due to the findings in Section 4.2, this analysis also considered a buffer of two miles around the stations to estimate the cost of travel by walking. This distance was observed as the range within which passengers traveled to reach a station according to the survey results where 94% of respondents who stated that they walked to reach a station were traveling up to two miles. The equation used to calculate the walking road raster is as follows:

$$TC_w = \frac{100}{w*1609.34} * Hourly\ Wage \quad (6.3)$$

where TC_w = transportation cost by walking in cents per meter, Hourly Wage=average hourly wage by CBG in dollars, w=walking speed in mph, and 1609.34 a conversion factor between meters and miles. Similar to the driving analysis, the network was assigned an integer value of 21 which represents the travel time required to traverse each cell. After the *Cost Distance* tool was used to find the least accumulative time to reach a station, this value was plugged into the CBGs by using *Zonal Statistics* tool. Then, the mean value of time obtained was multiplied by the average hourly wage in each CBG to find the estimated cost to walk to a station.

6.2.4 Average Travel Cost Estimation to the Nearest Station by Transit

For the transit analysis, some constraints were established at the beginning of the analysis. One of the major concerns about the HST is its schedule. The hours that the HST leaves and arrives to Indiana stations are not convenient to take advantage of the public transportation available in Lafayette and Indianapolis, which are the only two areas that have service around the station.

The Greater Lafayette area (Lafayette and West Lafayette) is served by CityBus, the operating name for the public transportation corporation. CityBus was established as a municipal corporation in 1971 (CityBus, 2017). A review of the CityBus schedule in Lafayette showed that most of the services that served the CityBus Center (the principal transfer terminal of the system) on weekdays could reach the Amtrak Station by 7:36 am when the HST is scheduled to leave. However, the area is only served by four routes at 9:40 pm (estimated arrival time of the HST). Also, on Sundays, when the HST has the same schedule of arrival and departure times, the station area is not served by public transportation at any of those times.

Turning to Indianapolis, IndyGo is the public transportation provider in this area. IndyGo operates 31 bus routes throughout Marion County and provides nearly 10 million passenger trips a year (IndyGo, 2017). Even though IndyGo is the largest public transportation provider in the state of Indiana, there are no routes that offer transportation to passengers to arrive by 5:30 am to the Indianapolis station in order for the passenger to check-in and be ready to board the train at the 6:00 am departure time. The HST also arrives in Indianapolis at 12:00 am when there are no buses scheduled to depart from Union Station.

Considering the study areas, the first analysis was conducted within the available services around the Lafayette station when the train departs/arrives. Later, the second analysis was performed to determine how the accessibility would change if all the services around the station were available at the departure and arrival times of the HST. In the second analysis, the Indianapolis area also was studied using the routes that actually could provide service to the station at different times of the day.

Because of the multimodal nature of transit, the analysis in this thesis was conducted in three parts: walking to a bus stop, taking the bus to the stop nearest to the station, and walking from the bus stop to the station. One of the issues in this analysis was the location of the bus stops since they were not necessarily in the same raster cell as the walking or transit network. However, the cells with bus stops were added to the walking and bus raster networks by using the *Near* tool.

The transit and walking network shape files were converted into rasters by using the *Polyline to Raster* tool. It is important to note that the cells to which no bus routes were assigned were initially given a value of zero by the tool; however, it was necessary to reclassify them as *NoData*. The transit line and bus stops rasters were later added and the travel time per unit distance of traversing each cell was calculated with the first part of the following equation:

$$TC_b = \frac{100}{b * 1609.34} * Hourly\ wage + fare \quad (6.4)$$

where $TC_b =$ travel cost by bus in cents per meter, Hourly Wage=average hourly wage by CBG in dollars, b=the bus speed, assumed to be 12.5 mph (IndyGo, 2010), and 1609.34 is a conversion factor between meters and miles, which was converted to an integer number in order to use the value of time in the *Cost Distance* tool. The value of the bus rate was not considered as a raster because this value is fixed (does not depend on the miles traveled) and is added to the final total cost.

After each of the cost rasters was calculated in terms of time traveled (walking to a bus stop, taking the bus to the stop nearest to the station, and walking from the bus stop to the station), they were assigned to the CBGs by using the *Zonal Statistics tool*. After those three values of time were determined, they were multiplied by the average hourly wage in that CBG. Finally, all three costs were totaled, and the value of the fare ($1 for Lafayette and $1.75 for Indianapolis) was added to the final cost.

6.2.5 Average Travel Cost Estimation to the Nearest Station Using Ridesharing Services

The final analysis was performed for ridesharing systems. This mode is currently available in the greater Lafayette and Indianapolis areas. The presence of Uber and Lyft in these areas has facilitated the movement of people. One of the advantages of this mode is that it does not have a fixed schedule, and a person can request it at any time. However, the service is constrained to the usage of a smartphone and the availability of the rideshare drivers in the area.

The ridesharing service chosen to perform this analysis was Lyft. The differences between Lyft and Uber mainly pertain to their prices and popularity; however, at the time this thesis was being developed (August 2017), a partnership between Amtrak and Lyft was available. The partnership agreement provides a $5 discount for each of the first four Lyft rides by using a promo code. Although the author did not have any information about the usage of this discount, it is an alternative already implemented and worthy of study. The charges for Lyft in Indianapolis and

Lafayette are summarized in Table 6.33 (Inc, 2017). It is also important to mention that ridesharing services have different types of services, such as Original Lyft and Lyft Plus. The service used to perform this analysis was Original Lyft due to its wider availability in the area. The Original Lyft allows the passenger to ride solo or with up to three friends. By using the mobile app, the availability of the Lyft service was tested from the different areas (rural or urban) in the counties where the service was operating at that time.

Table 6.3 Lyft charges according to the area (Source: (Inc, 2017))

Charge (Dollars)	***Indianapolis***	***Lafayette***	***All Counties***
Service Fee	$2.15	$2.20	$3.00
Cost Per Mile	$0.81	$1.20	$1.56
Cost Per Minute	$0.15	$0.20	$0.20
Base Fare	$1.25	$2.00	$2.00
Minimum Fare	$3.00	$2.60	$4.00

For the analysis of the area defined in the driving alone exploration, the highest prices of the Lyft services found in Indiana were used since some of the trips would be longer in the case of travel from a county far away from the station. For this analysis, it was also assumed that drivers will be available at the time a passenger will request the service. After the previous conditions were set, two rasters were created from the network. The following formula is a combination of the driving alone analysis, and the transit analysis in the sense that the speed limit of the network is used, and there is a cost per mile but there is also a fare involved in the analysis.

$$TC_{rs} = \frac{\frac{100}{s} X HourlyWage}{1609.34} + \frac{m}{1609.34} + fare \quad (6.5)$$

where TC_{rs} = travel cost by bus in cents per meter, Hourly Wage=average hourly wage by CBG in dollars, *s*=speed limit in mph, *m*=cost per mile in cents, fare=service fee in dollars, and 1609.34 is a conversion factor between meters and miles. For this analysis, two rasters were created: one raster capturing the least travel time to reach the station and the other capturing the cost per mile due to the service. The former was multiplied by the hourly wage data and totaled with the latter. The service fee was also considered in this analysis, but was not involved in the first part of the previous equation as this value is fixed (it does not depend on the miles traveled) and it was added to the final total cost.

As the hourly wages were defined per CBG, the least accumulative travel time to reach a station from an origin was calculated as a cost surface by using the *Cost Distance* tool. In this case, after having the results for the cost distance analysis, this value was assigned to the specific zone by using the *Zonal Statistics* tool and then was multiplied by each CBG average hourly wage. A similar analysis was performed to identify the cost per mile for using the service from a particular zone to the station. For the service cost, the value of the raster was fixed to 50 for Indianapolis and 75 for Lafayette (0.001 cents/m) as an integer number was needed to execute the *Cost Distance* tool. After obtaining the least accumulative cost to reach the station, these results were plugged into the CBG shape by using the *Zonal Statistics* tool. Once both values were converted to a cost, they were totaled. Finally, the value of the service fee ($2.15 for Lafayette, $2.20 for Indianapolis, and $3.00 for all other counties) was added to the final cost, which provided the estimated cost to travel from each of the CBGs analyzed to reach the nearest HST station by ridesharing services.

6.3 Travel Cost Analyses Results

The results of the travel cost analyses generated the one-way average trip cost from a CBG to the stations. By observing the travel cost, it was possible to find areas for which the cost is higher to reach a HST station.

The first analysis carried out was the driving alone option. This option was the preferred one for HST passengers to reach and leave a station. Figure 6.2 shows the average cost per each CBG by driving alone. The average travel cost by this mode varied from $0.03 to $89.60. As expected, for CBGs located close to rail stations, it costs less to reach a station (up to $10.36 dollars). Some of the passengers that drove to reach the station from CBGs classified in the third and the fourth quantile, incurred more than the average cost of a trip on the HST from Indianapolis to Chicago, which is $38. Additionally, most of the CBGs that are located outside the first quantile around the stations are classified as rural, and 86% of the area of study is classified as rural. People located in that area would have to spend between $10.37 and $89.60 to reach a station. Note that the average travel cost does not include daily parking fees, which can vary between stations. For instance, the daily parking fee around the Indianapolis station is $17 on average. When this cost was accounted for, the highest cost to reach the station from the closest area increased from $10.36 to $27.36 (Figure 6.3). For passengers traveling from the CBGs located in the fourth quantile, this option cost between $67.06 and $106.60 when it was assumed they would travel by rail and return on the same day. When parking was considered, driving alone was an expensive option to reach the station from the closest areas. For the passenger located in the second, third, and fourth quantiles, the cost to reach the station was higher than an average trip on the HST when the parking cost was added to the results.

When passengers were “dropped off/or picked up” at a HST station, the cost to reach the station was higher than driving alone because these trips are performed by at least two people, Therefore, the minimum amount that passengers traveling from the first quantile (closest to the station) would spend increased around 30% from $0.30 to $0.43 when parking was not considered in the driving alone option. Similarly, the maximum amount spent by someone who traveled from the furthest counties (those located in the fourth quantile) increased around 31%, from $89.60 to $130.60. In this scenario, some of the passenger traveling from the third quantile and passengers coming from or going to the fourth quantile would spend more resources than the average cost of a ticket from Indianapolis to Chicago ($38), which was the most expensive alternative. Similar to the driving alone option, 86% of this area is categorized as rural. People coming from rural areas would have to spend between $15.01 to $130.60 to reach or leave a station by this option. When comparing this option with the driving alone option, which includes parking cost, the someone dropped me off/picked me up option was less expensive for passengers who needed to travel from the first and second quantile to reach the station.

The results of the walking analysis are presented in Figure 6.5. For the different stations, passengers located in the closest quantile would spend up to $6.62 to reach a station. The minimum value was found to be $1.48. When compared with the minimum value found for the driving alone option when parking was not considered and the someone dropped me off/picked me up option, the cost of walking was about five times higher and 3.5 times higher, respectively. When the parking cost was considered, walking became a less expensive option to reach the station compared to the driving alone option from the CBGs around the stations. Additionally, the Indianapolis station only resulted with CBGs around the station classified in the first, second, and third quantile. These findings were due to the division of the geography units; however, they also represented the urban structure of the city of Indianapolis, where the sidewalks are better connected in the downtown area. A similar trend was seen for the Lafayette station, where there are no CBGs located in the highest quantile in the area of analysis (passenger spending more than $23.97 to reach a station by walking). The CBGs considered in the two-mile buffer did not belong to the rural classification for any of the station locations but Rensselaer, located in Jasper County.

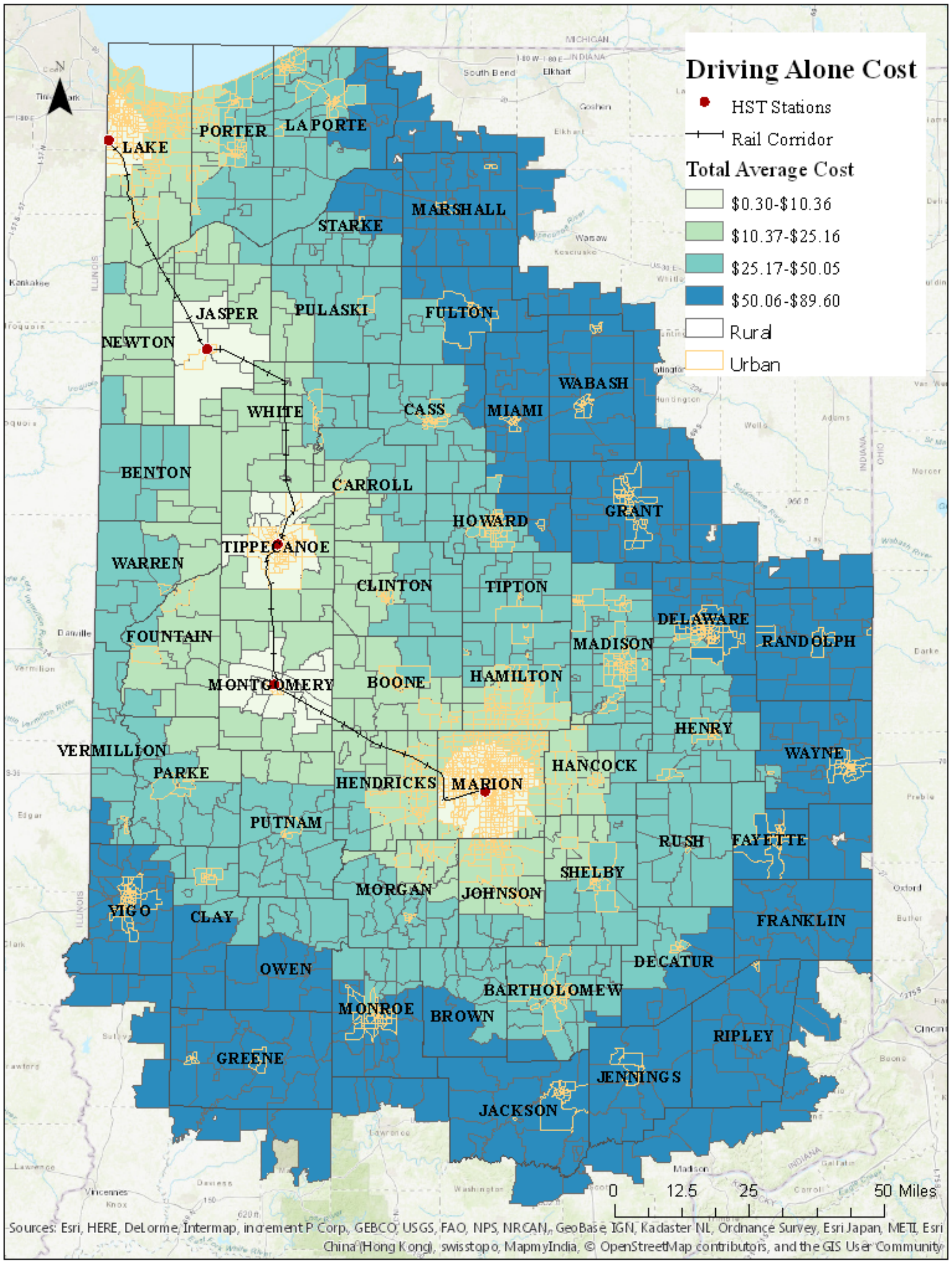


Figure 6.2 Average Travel Cost to the Nearest Station by Driving Alone

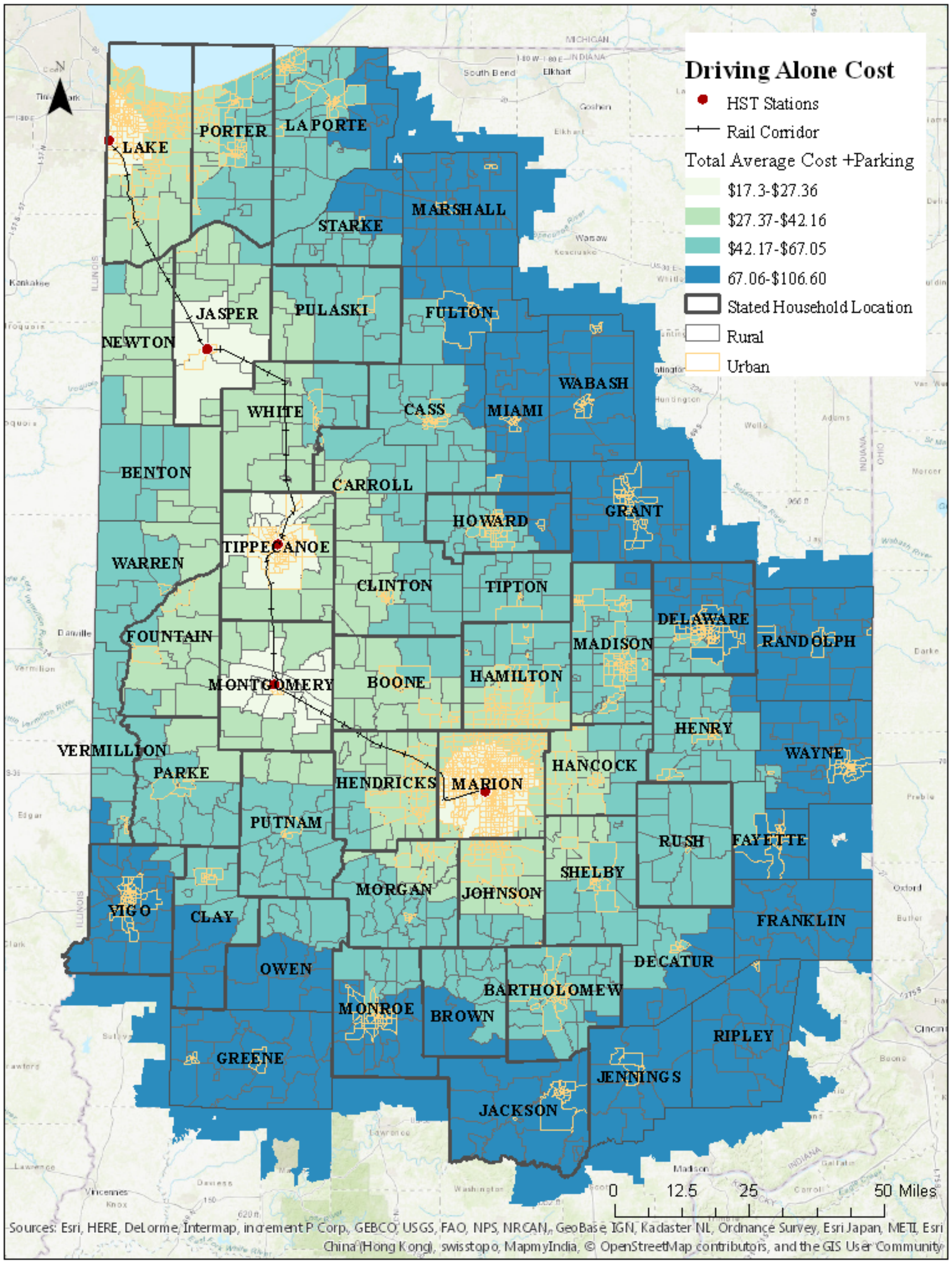


Figure 6.3 Average Travel Cost to the Nearest Station by Driving Alone – Parking included.

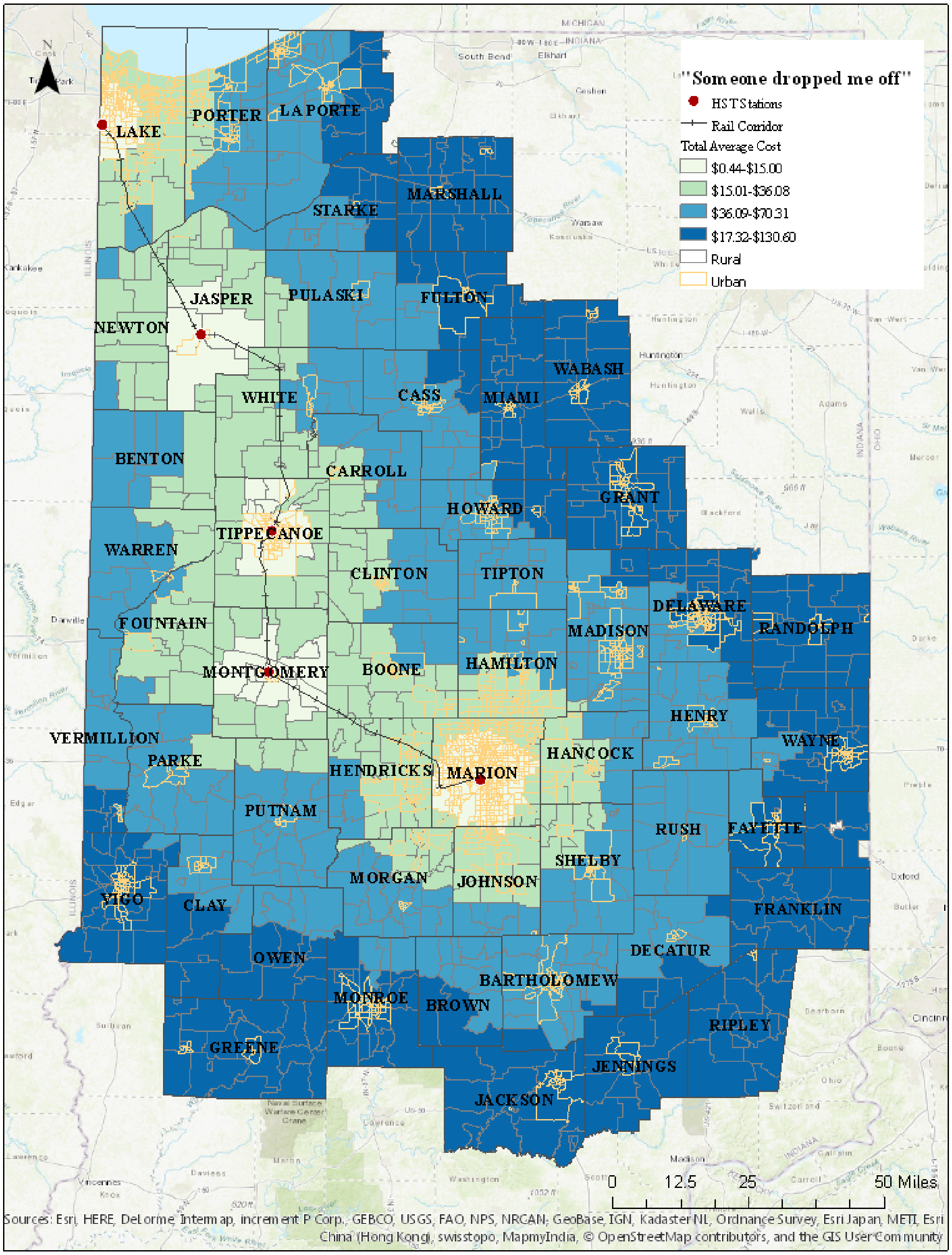
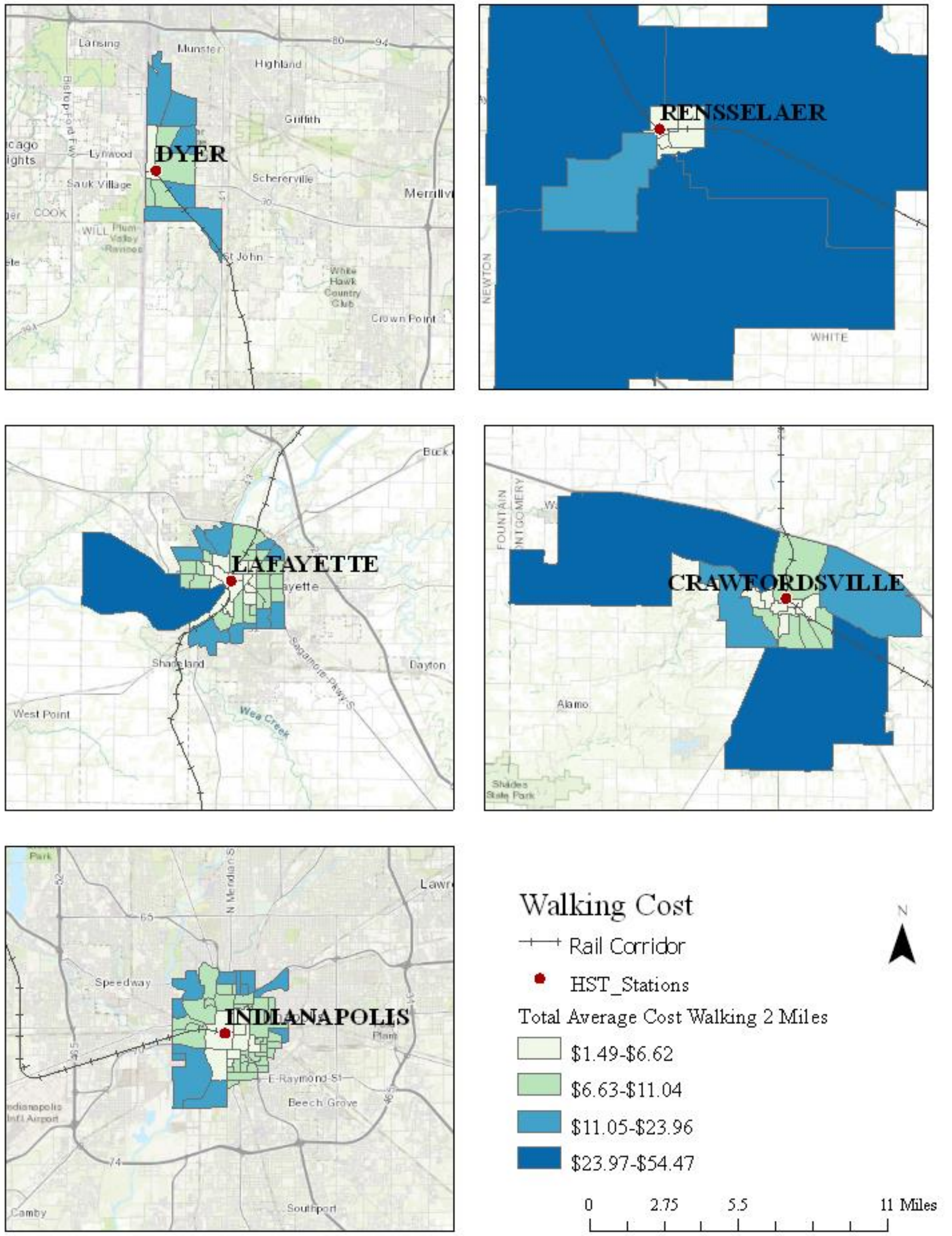


Figure 6.4 Average Travel Cost to the Nearest Station by “Someone dropped me off” option



Sources: Esri, HERE, DeLorme, Intermap, increment P Corp., GEBCO, USGS, FAO, NPS, NRCAN, GeoBase, IGN, Kadaster NL, Ordnance Survey, Esri Japan, METI, Esri China (Hong Kong), swisstopo, MapmyIndia, © OpenStreetMap contributors, and the GIS User Community

Figure 6.5 Average Travel Cost to the Nearest Station by Walking

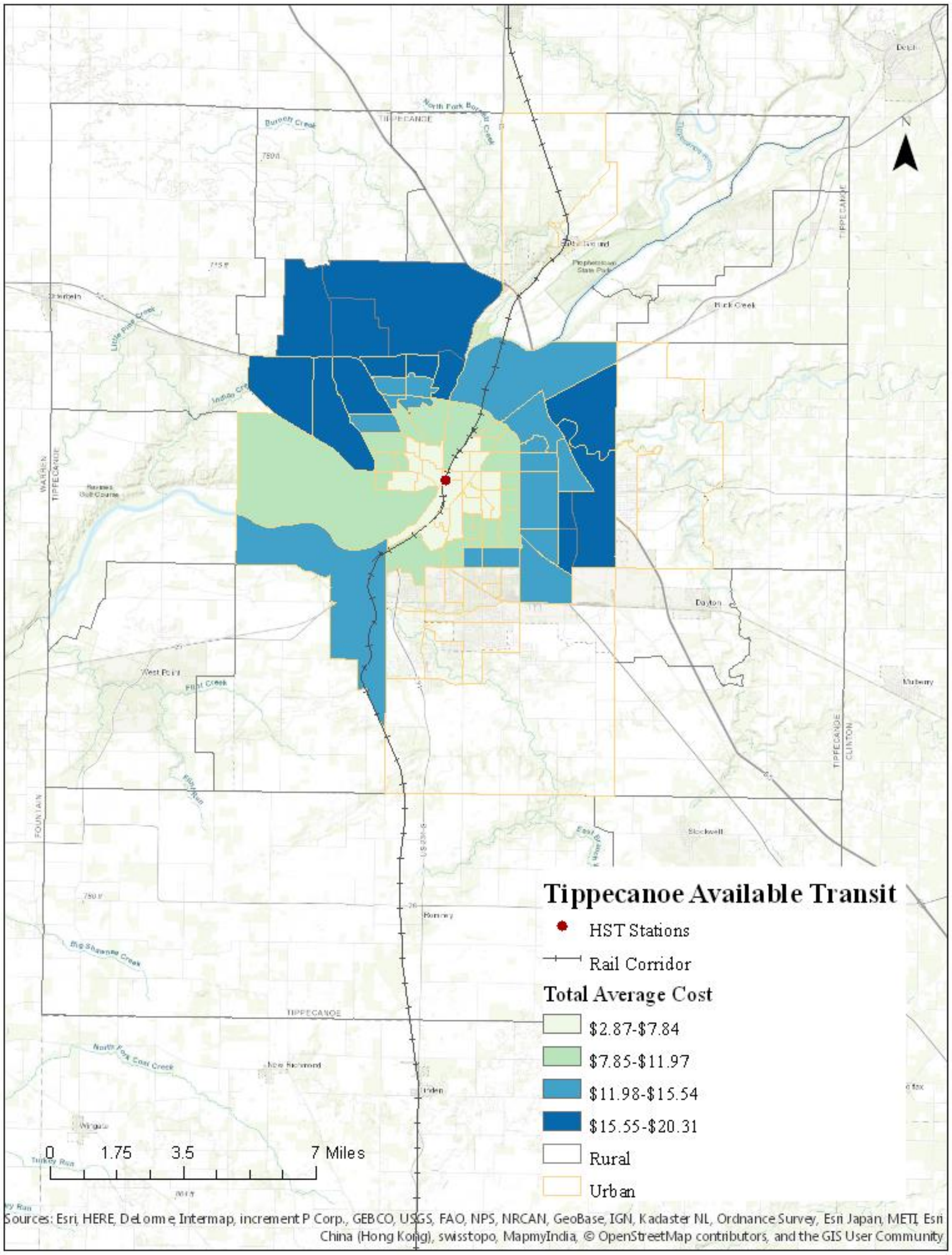


Figure 6.6 Average Travel Cost to the Nearest Station by Transit (Tippecanoe County)

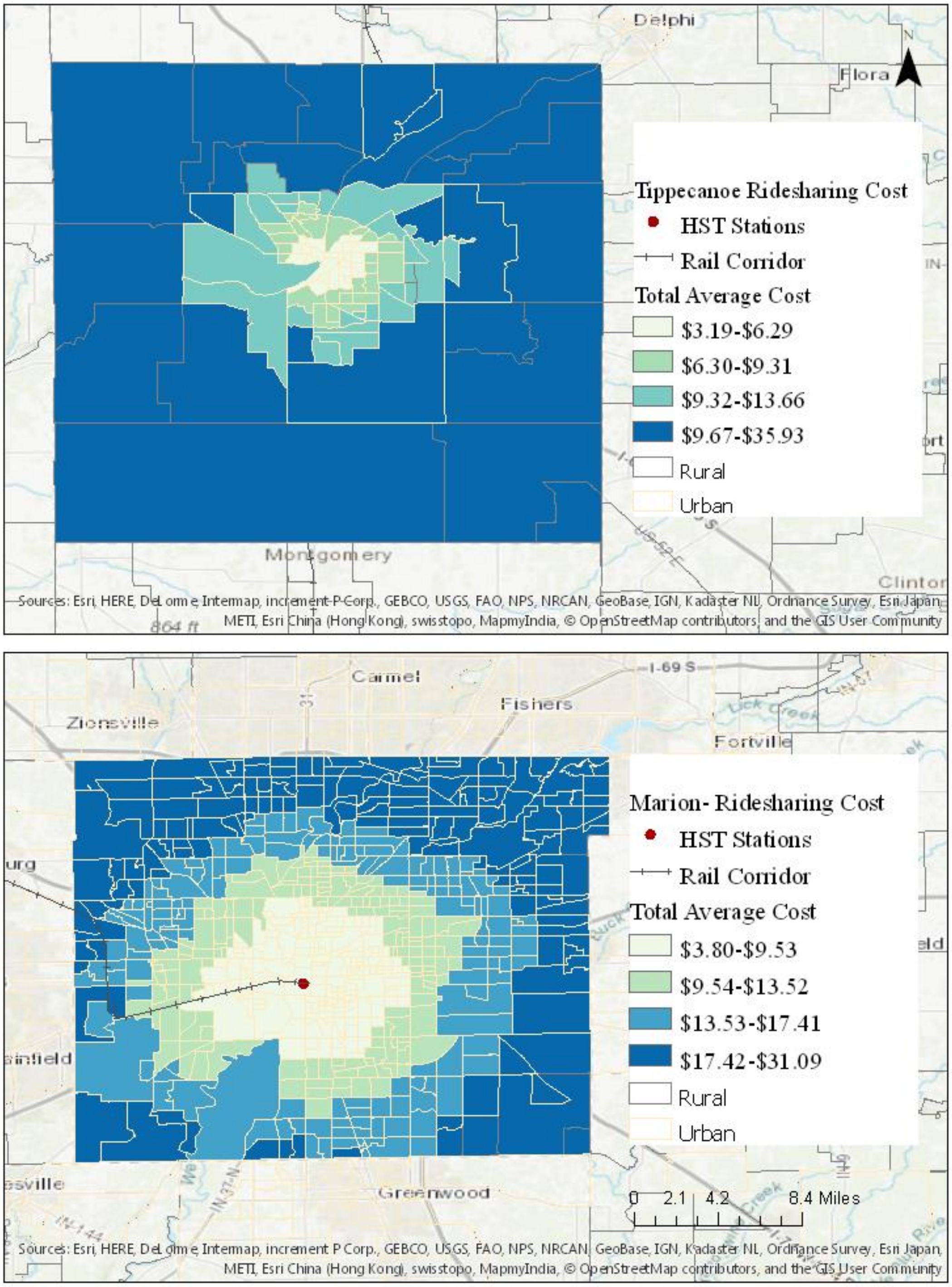


Figure 6.7 Average Travel Cost to the Nearest Station by Using Ridesharing Services

For the transit mode, only Lafayette and Indianapolis have public transportation systems. However, their current schedules constrain the use of transit as a FMLM service for the area. There are only four routes available to serve the Lafayette area at the times the train departs or arrives to that station, and Indianapolis does not have any routes that can serve the station at the times the train stops. For that reason, the current scenario was only analyzed for the Lafayette area, but a hypothetical scenario is explored for Indianapolis in the next subsection. As presented in Figure 6.6, the available service merely covers the Lafayette and West Lafayette area. This means that only 15% of the Tippecanoe County area is covered by this service. From that area, 14% is classified as rural according to the 2010 Census Urban Rural classification. Additionally, some areas classified as urban would not be covered by transit in the current service. The minimum cost to reach the Amtrak station in the Lafayette area is $2.87. This cost is due to the system fare ($1) and the cost to walk from the closest bus stop location to the station. It is worth mentioning that this analysis assumed everyone pays a fare when using the Lafayette CityBus services. This clarification is made due to the special agreements between CityBus and Purdue University and some other academic institutions around the area. In comparison with the walking mode, the results show that transit allows reaching a wider area at a similar cost. For example, the closest quantile from the station covers an approximately 1.5 mile area around the rail station if a passenger chooses to travel by transit. However, if the passenger is walking to the station, the closest quantile covers approximately 0.8 miles around the station for a similar cost. On the other hand, if transit is compared either with the driving alone or someone dropped me off/picked me up options, the former makes it possible to cover a wider area at a less expensive cost. If the minimum costs between these options and the transit option are compared, a passenger traveling by transit would spend around nine times more than a passenger choosing to drive when the parking cost is not considered. If the parking cost is considered, transit would become a better option to reach the station from the closets CBGs.

A special analysis was carried for Tippecanoe and Marion County in terms of ridesharing services. The stations located in those two counties are the only ones that are served by Lyft, the ridesharing service analyzed in this option to reach/leave the station. The outcomes show that the minimum cost to reach a station for Tippecanoe and Marion County are $3.19 and $3.40, respectively. These costs are considerably higher than for previous modes (driving alone, someone dropped me off/picked me up, and walking) because they have a service fee. However, they can provide service to a wider area compared to the transit analysis for a similar cost as shown in Figure 6.5. Marion County is mainly covered by CBGs classified as urban areas. However, the use of ridesharing services as a FMLM feeder serves also the 9.4% of the area that is classified as a rural in this county. Similarly, ridesharing services make possible the access of 32% of area classified as rural in Tippecanoe County. These results make the ridesharing service option preferable for reaching the station from those areas compared to transit. Additionally, when the cost of parking was considered in the driving alone option, ridesharing became less expensive for reaching the station in Tippecanoe and Marion Counties.

6.3.1 Scenario-based Analysis

This section presents the results of hypothetical scenarios related to the availability of transit in Tippecanoe and Marion Counties and the availability of ridesharing services in the entire study area. Table 6.4 presents the descriptive statistics of the available and proposed services considering the conditions given. If transit were available at the time the HST arrives in Indianapolis, a passenger who is riding from this county would have a minimum cost of $3.52 to reach the station. This can also be seen in Figure 6.9 where Marion County appears to be totally covered by transit. The maximum cost that a passenger traveling from a place located in the fourth quantile would face is $48.98. The implementation of this alternative would increase the coverage of this mode to 100% in this area and would also provide accessibility to those who live in the 9.4% classified as rural in this county.

Table 6.4 Summary Statistics for Cost Analysis.

	Average of Total Cost	***Max of Total Cost***	***Min of Total Cost***	***Std. Dev of T. Cost***
Driving Alone	$30.56	$89.60	$0.30	$22.28
Someone Drop me off	$43.50	$130.60	$0.44	$31.41
Walking	$10.06	$54.47	$1.49	$7.29
Lyft Tipp.[1]	$11.45	$35.93	$3.19	$7.26
Lyft Marion[1]	$13.85	$31.09	$3.93	$5.37
Lyft All area[2]	$68.88	$199.23	$4.71	$50.52
Transit Laf[1]	$12.95	$20.31	$2.87	$4.27
Transit Laf[12]	$11.07	$29.42	$2.81	$6.10
Transit Marion[12]	$31.74	$63.78	$4.78	$11.52

[1]Average Travel Cost for the areas where service is currently available. [2] Average Travel Cost estimated using a hypothetical scenario.

On the other hand, Figure 6.9 also shows the case study of Tippecanoe County if their whole transit system would be in operation at the times the HST arrives or departs to/from the Lafayette station. In this case, the minimum cost would decrease around 2% from $2.87 to $2.81. This change is due to the availability of more routes. Moreover, this alternative covers 37% of the total area of the county, of which 47% is classified as rural. This is an increment of 12% from the actual service. Even though all the routes are considered, in the case of Tippecanoe County, it is not possible to obtain 100% coverage due to the current design of the public transit system; therefore, the remaining 68% of rural areas would not be able to access to the station by using transit as a FMLM service.

Besides the transit analysis, a ridesharing system analysis was included in this hypothetical scenario analysis. By assigning the most expensive values of cost per mile and service fees from the Lyft services available around Indiana, it was possible to create a cost raster that would cover the entire area of study. In that sense, a passenger who travels from the closest quantile to reach a station would spend a minimum amount of $4.71 dollars. Even though this minimum is more expensive that the minimum seen for the current service, the area covered would be wider. When compared with the driving alone and the someone dropped me off/picked me up options, a person using a ridesharing service would spend around two times (more when parking cost is not considered) and 1.5 more than using those options. However, a passenger who uses a ridesharing service to reach the station would not need to pay for parking and it would only depend on the availability of Lyft drivers in the area, which is assumed in this analysis. When parking is considered, ridesharing services become a less expensive option for passengers that need to reach the station from the closest area (first quantile). Additionally, people located in the third and fourth quantiles would spend more to reach the station than the average value of the ticket from Indianapolis to Chicago, which is $38. Additionally, the availability of ridesharing services in all the counties in the study area would provide access to the 86% of areas classified as rural. People located in those areas would have to spend between $23.19 and $199.23 to reach a station, but it would not depend on whether he/she has or can drive a car or whether someone else is willing to give her/him a ride to reach one of the HST stations.

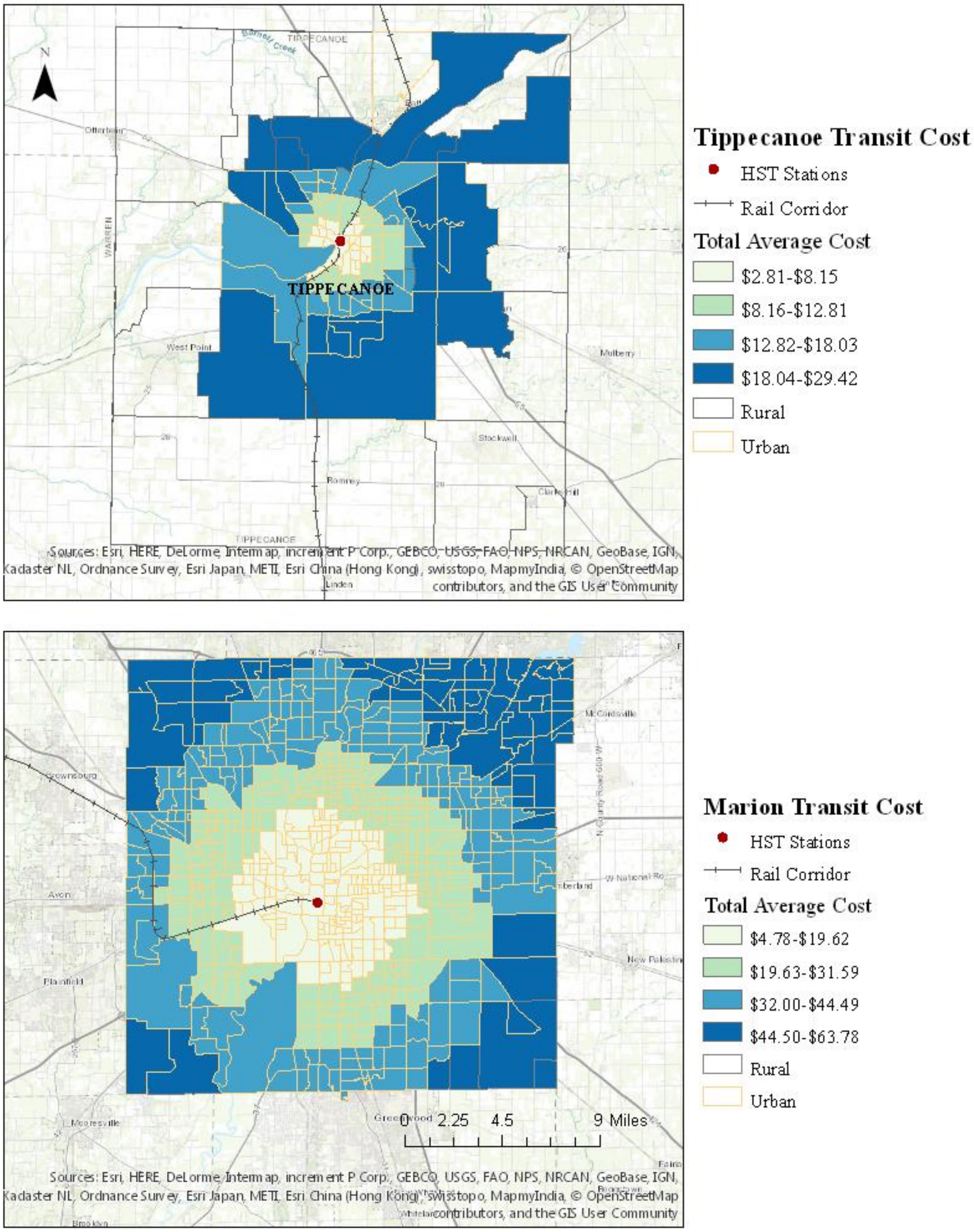


Figure 6.8 Average Travel Cost to the Nearest Station by Transit (Tippecanoe and Marion Counties)

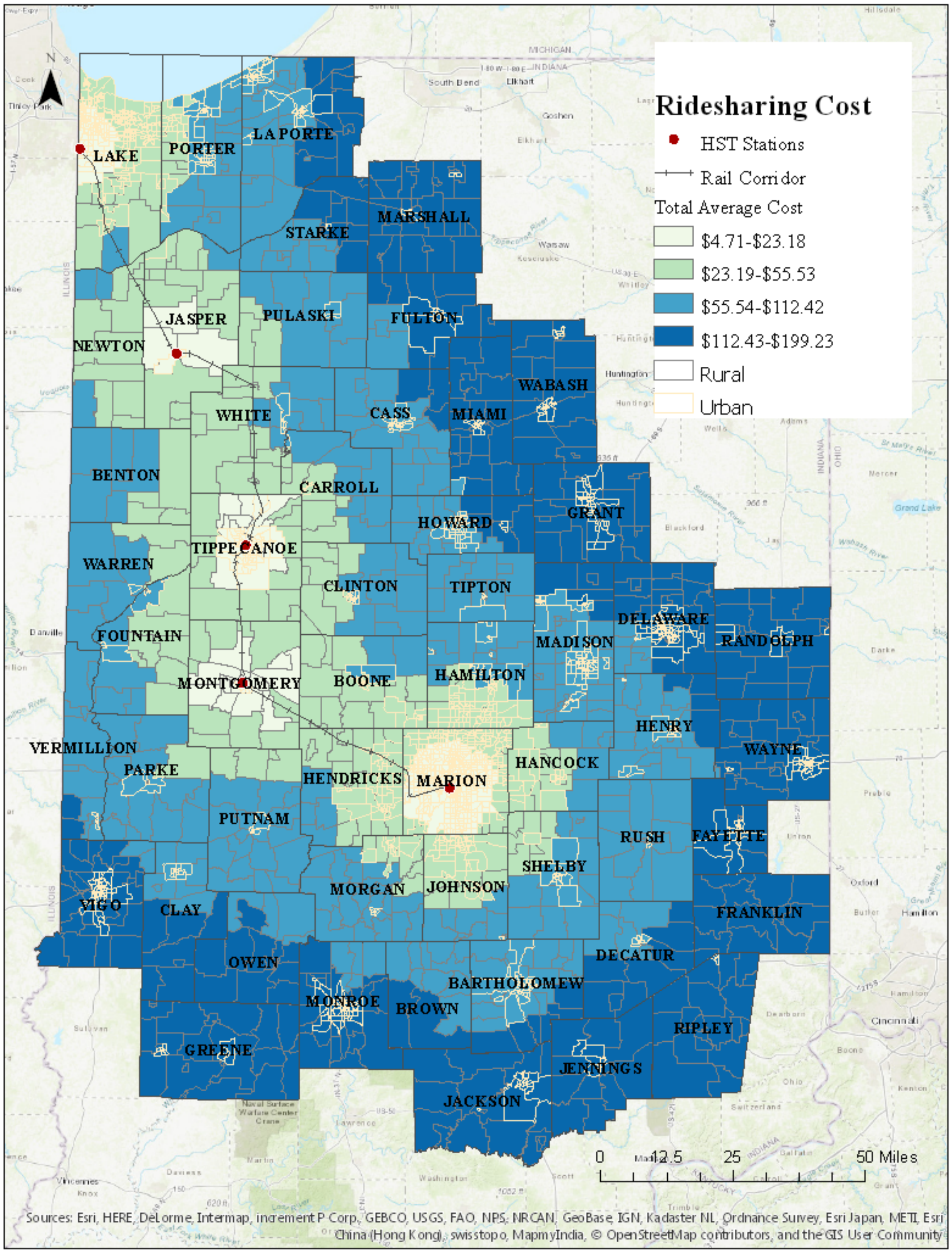


Figure 6.9 Average Travel Cost to the Nearest Station by Ridesharing (All Counties)

6.4 Concluding Remarks

The current alternatives to reach and leave HST stations were evaluated by performing an accessibility analysis using cost rasters. As driving alone was the preferred mode among the survey respondents, a gap in first and last mile travel options was identified. After considering all the options currently available in the different areas studied in this thesis, it was found that the least expensive options to reach the station were the driving alone and someone dropped me off/picked me up options since other services such as public transit and ridesharing services are not available in the whole area of study. This option also allows passengers located in rural areas to reach the stations. However, it was found that when the cost of parking around the station was considered, driving alone became the preferred option only for passengers located in the second, third, and fourth quantile of the analysis. Furthermore, the analysis indicates that only Lafayette could be served by transit using the current HST schedule, but this alternative only covers 15% of the Tippecanoe county area. When Marion County was considered as having available transit service, the entire county was served and the maximum amount that a passenger would spend to reach the station was $63.78. This alternative, however, was more expensive than driving alone even when parking cost was considered. Also, the someone dropped me off/picked me up option was less expensive compared to transit but the later might provide service to passengers who did not own a car or who cannot be given a ride to the station. In terms of rural passengers' accessibility, transit would only provide access to those passengers located the rural areas within Tippecanoe and Marion County. For the former county, only 32% of the area classified as rural would be served be transit if the whole routes area available at the times the HST arrives or departs. At that point, all the other counties included in the analysis, (except for Tippecanoe and Marion, are in need of a transit service that could provide passengers the possibility to reach the HST stations.

The option of ridesharing was also considered. The ridesharing option appeared to be more expensive than the most popular modes that HST passengers reported using to reach/leave the station (driving alone or someone dropped me off/picked me up options); however, this service would provide a wider flexibility. Furthermore, passengers who have certain constraints to reach the station, such as disabilities, or passengers who do not own a car or do not have a driver's license, would be able to reach the station for a comparable train with transit if we looked at the areas where it is available. Passengers coming from areas classified as rural around the study area (86% of the study area) would also benefit from a ridesharing service to reach a HST station. Finally, alternatives to subsidize this alternative could be explored by the HST operation team, such as the one already available between Amtrak and Lyft. Those alternatives might be available as well for the counties further away from the rail stations, which are located mainly in urban counties.

7. CONCLUSIONS, LIMITATIONS AND RECOMMENDATIONS

7.1 Summary and Contributions of this Thesis

This thesis addressed the lack of research to date pertaining to intercity passenger rail and the need for a better understanding of the factors affecting the level of usage as well as the magnitude of the FMLM problem that intercity passengers face. Referring back to the research objectives, this thesis had two main purposes: 1) identify the different factors that influence the mode choice of commuters for medium distance travel and the level of usage of an intercity passenger rail service and 2) explore strategies to attract a wider number of passengers by addressing the FMLM problem. To accomplish these objectives, the HST system served as a case study. An on-board survey was the main data source for this thesis; and secondary data included, but was not limited to, geographic information files and the GTFS.

The first objective involved two research questions. For the first question, "*What role does the FMLM play in the mode choice of medium distance passengers?"* for which a MAM was used. This model was estimated for four distance ranges (quartiles of the distance traveled to reach the station according to the survey responses) defined as a proxy of access to the station in order to understand the FMLM problem. The MAM results showed that the intercity train was the most favorable mode for passengers who traveled less than two miles to access a station. This finding suggests that people who traveled less to access a station would be more likely to take the train if they had the chance to do so. However, it was also found that the intercity train was the preferred mode for respondents who traveled more than 24 miles. This finding emerged due to the score given to the most important factors considered to estimate the MAM in that range. In this sense, factors such as safety, ease of use, and reliability were identified as the most important factors in mode choice decision-making, and the rank obtained for those three factors was fundamental for the total MAM estimations. While an intercity train was highly ranked for safety and ease of use by respondents who traveled more than 24 hours, the driving alone option was ranked low with regards to those factors but high in other factors such as reliability. The sensitivity analysis displayed that the factors and order of the preferred modes did not change drastically when other distances were considered in the ranges. Due to those findings, it was concluded that the role of the FMLM in the mode choice of medium distance passengers is also influenced by the passenger's perspective of other important factors, such as safety, ease of use, and reliability, which depend on the mode itself.

In a similar vein, the second research question, "*What is the relationship between frequency of travel by intercity rail and (i) mode choice-related factors, (ii) factors associated with access to a rail line, and (iii) passenger characteristics?*" served to identify the factors that affect the level of usage of intercity passenger trains, which was tested using an ordered probit model. This model was primarily chosen due to the nature of the data. Three categories to describe the frequency with which passengers were traveling on the HST were defined as new passengers, medium frequency passengers, and high frequency passengers. The ordered probit model suggested that safety and ease of use are important factors for intercity trip frequency as well, along with the mode used to access a station and the passengers' socio-demographic characteristics. In addition, some access-related variables seemed to be associated with trip frequency; however, the variables directly capturing this factor seemed to be overshadowed by other unobserved factors. One reason for this finding (or lack thereof) is that the frequency of travel by an intercity train might be more strongly associated with factors affecting the travel frequency itself (by whichever mode) rather than the choice of rail over another transportation alternative. Given that the percentage of induced travel for intercity trips is not expected to be very high, the lack of access to the train would probably affect mode choice decisions but not decisions as far as whether or not to take a trip.

The second objective of this thesis was to explore strategies to attract a wider number of passengers by addressing the FMLM problem. To that end, a third research question was proposed: "*Which strategies are the most advantageous for accessing an intercity passenger rail service??*" The *Spatial Analyst* tool in ArcGIS Pro was used to compute the travel cost to the closest station by producing rasters in which the cells were associated with a specific cost for the most prevalent modes in the survey results. Additionally, areas classified as rural were included in the analysis in order to study the needs of such areas in transportation modes that would allow them to reach and leave the station. The current alternatives to reach and leave a station were evaluated by performing an accessibility analysis. Driving alone was the preferred mode to reach/leave a station. Therefore, a FMLM problem was identified along the HST line. The only option that was considered on the survey but excluded from the analysis was "rode a bike" to reach the station. The exclusion of this mode was based on the low percentage of respondents in the survey, as seen in Figure 3.1. After considering the options currently available in the study area, it was found that the least expensive options to reach the station were the driving alone (when parking is not considered) and someone dropped me off/picked me up options since other services such as public transit and ridesharing services were not available in the whole area of study. In the alternative scenario when parking cost was considered for the driving alone option, the results of this thesis show that driving alone was the most advantageous strategy in term of cost for those CBGs that are farther away from the station. On the other hand, the CBGs located around the stations belonging to the first and second quantile of analysis, would find better options to reach the station in other modes such as someone dropped me off/picked me up and ridesharing.

Furthermore, it was found that only Lafayette could be served by transit based on the current HST schedule; however, this alternative only serves 15% of the Tippecanoe County area. When Marion County (Indianapolis) was considered with its available transit service, the entire county could be served and the maximum amount that a passenger would spend to reach the station was $63.78. This alternative, however, was more expensive than driving alone even when parking was considered. Additionally, the someone dropped me off/picked me up option to reach the station was less expensive, but transit might provide service to passengers who do not own a car or cannot get a ride to the station. Correspondingly, all the other counties included in the analysis (except for Tippecanoe and Marion) are in need of a transit service that could provide passengers the opportunity to reach the HST stations.

The option of ridesharing was also considered. The ridesharing option was more expensive than the most popular modes that HST passengers reported using to reach/leave the station (driving alone or someone dropped me off/picked me up options); however, this service would provide wider flexibility. Furthermore, passengers who have certain constraints to reaching the station, such as disabilities, or passengers who do not own a car or do not have a driver’s license would be able to reach the station for a comparable cost than using transit where it is available. Finally, alternatives to subsidize this service could be explored by the HST operation team, such as the one already available between Amtrak and Lyft. Those alternatives could be available for the counties farther away from the stations located mainly in urban counties.

The methodology proposed in this thesis can help state transportation agencies identify the most important factors for a passenger when choosing an intercity transportation option and devise related strategies to increase the number of passengers on intercity trains. Factors such as safety and ease of use could be more effectively marketed to attract new passengers that might have a lower level of access to a train. The proposed methodology is easily replicable, yet the inferences of this thesis are limited to the case study of the Indiana HST. Nevertheless, the results of this thesis may have extensive implications for planning strategies that can enhance access to rail stations. Also, the strategies proposed to address the FMLM could provide insights to intercity passenger rail service providers that can help attract a larger number of passengers. Regional improvements in transportation would require not only eliminating inefficiencies in the existing services, but also extending coverage to reach a greater proportion of the total population, such as rural areas, if ridesharing services were to be available to them.

7.2 Limitations

The main limitation of this thesis pertains to the survey data. Even though the primary source of information was a survey collected on-board the HST, the survey was not specifically designed to fulfill the needs of the research questions addressed above. The accuracy of the data also was a limitation, especially to address the second research question, due to frequency it is usually modeled using count data models; however, the frequency data were collected in bins, which makes an ordered model a better fit for the analysis. Additionally, the insights found in the first objective did not provide strong evidence that higher accessibility affects mode choice and level of usage. Even though different variables related to access were found to be significant in the models developed, there was not a specific question that addressed this issue in the survey. Another limitation of the data was found when the third question was analyzed. The data used to answer this question led to many assumptions made in this thesis. For example, the speed limits used were generalized to the road type due to the specific speeds for road segments not being available. Another issue found in this analysis was the availability of a network to perform the analysis using the *Network Analysis* tool instead of the *Spatial Analysis* tool in ArcGIS Pro. The use of this package of tools would allow taking into consideration the actual speed of the network and the transit frequency. It also would be able to account for the added time and cost of transfers between buses if that was necessary. These changes could increase the correctness of the travel cost estimate. Additionally, there were no data that described whether sidewalks were present in the area studied. Assumptions therefore were made to perform the walking analysis as well as the transit analysis part that included walking.

In addition, note that the survey data were collected on-board the HST only, and therefore the sample included only passengers of the train and not a sample of the general population. To address any potential selectivity bias issues, future work could explore the same research question for a similar intercity train service and include a wider sample with respondents from counties around the stations. Although the methodology is robust and can be useful for researchers elsewhere, the HST service is somewhat unique in terms of its operations (frequency, inconvenient schedule, and four day/week service), which can make the findings not transferable to other corridors. Most of the respondents of the on-board survey stated that they were using the train for leisure trips, which also infers that the results cannot be transferable to commuter intercity rail. However, to apply this methodology to other type of rail services, changes need to be made in the survey instrument to address specific points of interest on those lines and additional aspects could be considered in the MAM model, such as environmental factors, which were not included in the survey in this thesis. Changes in the survey questionnaire would not modify the methodology if the questions used to develop the model herein are asked in a similar manner.

7.3 Future Research

In addition to overcoming the limitations described above, recommendations for expanding this work are as follows. First, in order to further explore the relationship between access to a rail line and trip frequency, it is recommended that future studies assess whether access to intercity passenger rail is associated with the frequency of travel by rail, controlling for travel intensity. Additionally, the analysis of the third question was based on current services and their expansion. Future research could hypothesize the existence of more futuristic FMLM feeder strategies, such as autonomous vehicles. This assumption would need to be modeled into transit simulation software and integrated land-use models. The usage of autonomous vehicles as a FMLM feeder could save the cost of an additional person traveling in the car to reach the station. The use of a ridesharing company, such as Uber or Lyft, in their pool mode (dynamically creating a route based on where customers are and where they want to go) as a FMLM problem solution also could be an interesting analysis for a wider region such as the one explored herein. The use of micro-transit (IT-enabled private multi-passenger transportation services) options from the counties that do not have a station could also be explored to analyze the potential of this service to increase ridership. Future research also could address seasonal factors when analyzing the available modes around the stations. Modes such as walking or transit are preferred for people if the weather allows them to use these modes as a FMLM server.

REFERENCES

ACE. (2017). Infrastructure Report Card. American Association of Railroads.

Amtrak, P. (2017). Indianapolis Station Parking - Compare Prices and Reserve Parking - Amtrak. Retrieved November 28, 2017, from https://parking.amtrak.com/indianapolis-station-parking

Bailey, A. (2015). Low income, supermarket accessibility, and the transportation network: A multimodal analysis identifying areas of poor accessibility and intervention strategies in Indianapolis, Indiana.

Biba, S., Curtin, K. M., & Manca, G. (2010). A new method for determining the population with walking access to transit. International Journal of Geographical Information Science, 24(3), 347–364. https://doi.org/10.1080/13658810802646679

Bureau of Transportation Statistics. (2015). National Transportation Atlas Database. Retrieved November 26, 2017, from http://osav-usdot.opendata.arcgis.com/datasets?keyword=Rail

Burns, C. M., & Inglis, A. D. (2007). Measuring food access in Melbourne: Access to healthy and fast foods by car, bus and foot in an urban municipality in Melbourne. Health & Place, 13(4), 877–885. https://doi.org/10.1016/j.healthplace.2007.02.005

Ceder, A. (2007). Public transit planning and operation: theory, modelling and practice (1. ed). Burlington, MA: Elsevier Butterworth-Heinemann.

Ceder, A., Net, Y., & Coriat, C. (2009). Measuring Public Transport Connectivity Performance Applied in Auckland, New Zealand. Transportation Research Record: Journal of the Transportation Research Board, 2111, 139–147. https://doi.org/10.3141/2111-16

Chandra, S., Bari, M. E., Devarasetty, P. C., & Vadali, S. (2013). Accessibility evaluations of feeder transit services. Transportation Research Part A: Policy and Practice, 52, 47–63. https://doi.org/10.1016/j.tra.2013.05.001

Cheyne, C., & Imran, M. (2010). Exploring the potential for shared transport services in non-metropolitan cities of New Zealand. In Proceedings of the 32nd Australasian transport research forum (ATRF). Accessed October. Retrieved from http://atrf.info/papers/2009/2009_Cheyne_Imran.pdf

Cho, H. D. (2013). The factors that affect long-distance travel mode choice decisions and their implications for transportation policy. University of Florida.

CityBus. (2017). CityBus of Lafayette Indiana > About > About CityBus. Retrieved October 28, 2017, from http://www.gocitybus.com/About/About-CityBus

Dell'Amico, M., & Hadjidimitriou, S. (2012). Innovative Logistics Model and Containers Solution for Efficient Last Mile Delivery. Procedia - Social and Behavioral Sciences, 48, 1505–1514. https://doi.org/10.1016/j.sbspro.2012.06.1126

Dixon, K. K., Sarasua, W. A., Daniel, J., & Mazur, G. D. (2001). Tool for rural and statewide multimodal transportation planning. Journal of Computing in Civil Engineering, 15(4), 275–283.

Drea, J. T., & Hanna, J. B. (2000). Niche marketing in intrastate passenger rail transportation. Transportation Journal, 39(3), 33–43.

Eboli, L., & Mazzulla, G. (2011). A methodology for evaluating transit service quality based on subjective and objective measures from the passenger's point of view. Transport Policy, 18(1), 172–181. https://doi.org/10.1016/j.tranpol.2010.07.007

ESRI. (2016). Understanding cost distance analysis—Help | ArcGIS for Desktop. Retrieved October 6, 2017, from http://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-analyst-toolbox/understanding-cost-distance-analysis.htm

Foth, N., Manaugh, K., & El-Geneidy, A. M. (2013). Towards equitable transit: examining transit accessibility and social need in Toronto, Canada, 1996–2006. Journal of Transport Geography, 29, 1–10. https://doi.org/10.1016/j.jtrangeo.2012.12.008

Givoni, M., & Rietveld, P. (2007). The access journey to the railway station and its role in passengers' satisfaction with rail travel. Transport Policy, 14(5), 357–365. https://doi.org/10.1016/j.tranpol.2007.04.004

Google Transit. (2016). ¿Qué es la GTFS? | Google Transit | Google Developers. Retrieved October 28, 2017, from https://developers.google.com/transit/gtfs/

Governors' Institute. (2017). How and Why of Measuring-Access to opportunity: A guide to performance management.

Gutiérrez Gallego, J. A., Naranjo Gómez, J. M., Jaraíz-Cabanillas, F. J., Ruiz Labrador, E. E., & Jeong, J. S. (2015). A methodology to assess the connectivity caused by a transportation infrastructure: Application to the high-speed rail in Extremadura. Case Studies on Transport Policy, 3(4), 392–401. https://doi.org/10.1016/j.cstp.2015.06.003

Hadas, Y. (2013). Assessing public transport systems connectivity based on Google Transit data. Journal of Transport Geography, 33, 105–116. https://doi.org/10.1016/j.jtrangeo.2013.09.015

Hadas, Y., & Ranjitkar, P. (2012). Modeling public-transit connectivity with spatial quality-of-transfer measurements. Journal of Transport Geography, 22, 137–147. https://doi.org/10.1016/j.jtrangeo.2011.12.003

Hallett, L. F., & McDermott, D. (2011). Quantifying the extent and cost of food deserts in Lawrence, Kansas, USA. Applied Geography, 31(4), 1210–1215. https://doi.org/10.1016/j.apgeog.2010.09.006

Handy, S. L., & Niemeier, D. A. (1997). Measuring accessibility: an exploration of issues and alternatives. Environment and Planning A, 29(7), 1175–1194.

Inc, L. (2017). Cities. Retrieved October 6, 2017, from https://www.lyft.com/cities

IndyGo. (2010). Comprehansive Operational Analysis (No. Part 2).

IndyGo. (2017). About IndyGo. Retrieved October 28, 2017, from https://www.indygo.net/about-indygo/

Iwan, S., Kijewska, K., & Lemke, J. (2016). Analysis of Parcel Lockers' Efficiency as the Last Mile Delivery Solution – The Results of the Research in Poland. Transportation Research Procedia, 12, 644–655. https://doi.org/10.1016/j.trpro.2016.02.018

Jackman, S. (2000). Model for Ordered Outcomes. Retrieved June 21, 2017, from http://web.stanford.edu/class/polisci203/ordered.pdf

Kaplan, S., Popoks, D., Prato, C. G., & Ceder, A. (Avi). (2014). Using connectivity for measuring equity in transit provision. Journal of Transport Geography, 37, 82–92. https://doi.org/10.1016/j.jtrangeo.2014.04.016

King, D. A. (2016, October 6). What Do We Know About the “First Mile/Last Mile” Problem for Transit? Retrieved November 18, 2016, from https://transportist.org/2016/10/06/what-do-we-know-about-the-first-milelast-mile-problem-for-transit/

Kuby, M., Barranda, A., & Upchurch, C. (2004). Factors influencing light-rail station boardings in the United States. Transportation Research Part A: Policy and Practice, 38(3), 223–247. https://doi.org/10.1016/j.tra.2003.10.006

Lane, B. W. (2012). On the utility and challenges of high-speed rail in the United States. Journal of Transport Geography, 22, 282–284. https://doi.org/10.1016/j.jtrangeo.2012.01.030

Liang, X., Correia, G. H. de A., & van Arem, B. (2016). Optimizing the service area and trip selection of an electric automated taxi system used for the last mile of train trips. Transportation Research Part E: Logistics and Transportation Review, 93, 115–129. https://doi.org/10.1016/j.tre.2016.05.006

Lin, T. (Grace), Xia, J. (Cecilia), Robinson, T. P., Goulias, K. G., Church, R. L., Olaru, D., … Han, R. (2014). Spatial analysis of access to and accessibility surrounding train stations: a case study of accessibility for the elderly in Perth, Western Australia. Journal of Transport Geography, 39, 111–120. https://doi.org/10.1016/j.jtrangeo.2014.06.022

Lindgren Jr, J. H., & Konopa, L. J. (1980). A comparative analysis of multiattribute attitude models. Journal of the Academy of Marketing Science, 8(4), 374–389.

Liu, S., & Zhu, X. (2004). An Integrated GIS Approach to Accessibility Analysis. Transactions in GIS, 8(1), 45–62. https://doi.org/10.1111/j.1467-9671.2004.00167.x

Lovett, A., Haynes, R., Sünnenberg, G., & Gale, S. (2002). Car travel time and accessibility by bus to general practitioner services: a study using patient registers and GIS. Social Science & Medicine, 55(1), 97–111. https://doi.org/10.1016/S0277-9536(01)00212-X

Martens, K., Golub, A., & Robinson, G. (2012). A justice-theoretic approach to the distribution of transportation benefits: Implications for transportation planning practice in the United States. Transportation Research Part A: Policy and Practice, 46(4), 684–695. https://doi.org/10.1016/j.tra.2012.01.004

Mayhew, S. (2009). A Dictionary of Geography (4th ed.). Oxford University Press. Retrieved from http://www.oxfordreference.com/view/10.1093/acref/9780199231805.001.0001/acref-9780199231805

McCray, T., & Brais, N. (2007). Exploring the Role of Transportation in Fostering Social Exclusion: The Use of GIS to Support Qualitative Data. Networks and Spatial Economics, 7(4), 397–412. https://doi.org/10.1007/s11067-007-9031-x

McLafferty, S. L. (2003). GIS and Health Care. Annual Review of Public Health, 24(1), 25–42. https://doi.org/10.1146/annurev.publhealth.24.012902.141012

Metro. (2014a). First_Last_Mile_Strategic_Plan.pdf. Southern California Association of Governments - SCAGSounds.

Metro. (2014b, March). First Last Mile Strategic Plan. Los Angeles County Metropolitan Transportation Authority (Metro).

Moniruzzaman, M., & Páez, A. (2012). Accessibility to transit, by transit, and mode share: application of a logistic model with spatial filters. Journal of Transport Geography, 24, 198–205. https://doi.org/10.1016/j.jtrangeo.2012.02.006

Morgan, M. (2014). Yay, transit! - Using GTFS Data in ArcGIS Network Analyst. Retrieved April 10, 2017, from http://transit.melindamorang.com/

Muñoz-Villamizar, A., Montoya-Torres, J. R., & Vega-Mejía, C. A. (2015). Non-Collaborative versus Collaborative Last-Mile Delivery in Urban Systems with Stochastic Demands. Procedia CIRP, 30, 263–268. https://doi.org/10.1016/j.procir.2015.02.147

Murphy, S. F. and C., Transit Cooperative Research Program, Transportation Research Board, & National Academies of Sciences, Engineering, and Medicine. (2016). Shared Mobility and the Transformation of Public Transit. Washington, D.C.: Transportation Research Board. Retrieved from http://www.nap.edu/catalog/23578

Murray, A. T. (2003). A coverage model for improving public transit system accessibility and expanding access. Annals of Operations Research, 123(1), 143–156.

Murray, A. T., Davis, R., Stimson, R. J., & Ferreira, L. (1998). Public Transportation Access. Transportation Research Part D: Transport and Environment, 3(5), 319–328. https://doi.org/10.1016/S1361-9209(98)00010-8

Neely, S. (2016). Factors influencing mode choice for intercity travel from northern New England to major northeastern cities. The University of Vermont and State Agricultural College. Retrieved from http://search.proquest.com/openview/49764e9dd0a827956eef63058bf236ef/1?pq-origsite=gscholar&cbl=18750&diss=y

Nutley, S. (2003). Indicators of transport and accessibility problems in rural Australia. Journal of Transport Geography, 11(1), 55–71.

O'Sullivan, D., Morrison, A., & Shearer, J. (2000). Using desktop GIS for the investigation of accessibility by public transport: an isochrone approach. International Journal of Geographical Information Science, 14(1), 85–104. https://doi.org/10.1080/136588100240976

Patterns, L. D. T., & Choice, M. (2006). America on the Go. Retrieved from https://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/america_on_the_go/long_distance_transportation_patterns/pdf/entire.pdf

PRIIA. (2008). PRIIA Overview.

Pyrialakou. (2016). Assessing Public Transportation Options for Intercity Travel in U.S. Rural and Small Urban Areas: A Multimodal, Multiobjective, and People-Oriented Evaluation. PhD thesis.

Pyrialakou, V. D., Gkritza, K., & Fricker, J. D. (2016). Accessibility, mobility, and realized travel behavior: Assessing transport disadvantage from a policy perspective. Journal of Transport Geography, 51, 252–269. https://doi.org/10.1016/j.jtrangeo.2016.02.001

Pyrialakou, V. D., & Gkritza, N. (2016). Exploring the Opinions of Passenger Rail Riders: Evidence From the Hoosier State Train. In 2016 Joint Rail Conference (p. V001T08A002–V001T08A002). American Society of Mechanical Engineers. Retrieved from http://proceedings.asmedigitalcollection.asme.org/proceeding.aspx?articleid=2528485

Rao, D. K. L., & Jayasree, K. (2003). Rural Infrastructure Planning with emphasis on road network connectivity by Coplanar Concurrent Theory. In Map India Conference. Citeseer. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.488.2096&rep=rep1&type=pdf

Rietveld, P. (2000). The accessibility of railway stations: the role of the bicycle in The Netherlands. Transportation Research Part D: Transport and Environment, 5(1), 71–75.

Scott, R. A., & NZ Transport Agency. (2010). Demand responsive passenger transport in low-demand situations. Wellington, N.Z.: NZ Transport Agency.

Shaheen, S., Meyn, M., & Wipyewski, K. (2003). US shared-use vehicle survey findings on carsharing and station car growth: Obstacles and opportunities. Transportation Research Record: Journal of the Transportation Research Board, (1841), 90–98.

Sinha, K. C., & Labi, S. (2007). Transportation decision making: principles of project evaluation and programming. Hoboken, N.J: John Wiley.

Song, L., Cherrett, T., McLeod, F., & Guan, W. (2009). Addressing the Last Mile Problem: Transport Impacts of Collection and Delivery Points. Transportation Research Record: Journal of the Transportation Research Board, 2097, 9–18. https://doi.org/10.3141/2097-02

Tilahun, N., Thakuriah, P. (Vonu), Li, M., & Keita, Y. (2016). Transit use and the work commute: Analyzing the role of last mile issues. Journal of Transport Geography, 54, 359–368. https://doi.org/10.1016/j.jtrangeo.2016.06.021

Tyrinopoulos, Y., & Antoniou, C. (2008). Public transit user satisfaction: Variability and policy implications. Transport Policy, 15(4), 260–272. https://doi.org/10.1016/j.tranpol.2008.06.002

US Bureau of Labor Statistics. (2016). Persons at work in agriculture and nonagricultural industries by hours of work. Retrieved June 28, 2017, from https://www.bls.gov/cps/cpsaat19.htm

US Bureau of Transportation Statistics. (2015). Commute Mode Share: 2015. Retrieved August 1, 2017, from https://www.bts.gov/content/commute-mode-share-2015

US Bureau of Transportation Statistics. (2016). Table 3-17: Average Cost of Owning and Operating an Automobile(a) (Assuming 15,000 Vehicle-Miles per Year) | Bureau of Transportation Statistics. Retrieved June 22, 2017, from https://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/national_transportation_statistics/html/table_03_17.html

U.S. Census Bureau. (2010a). 2010 Census Urban and Rural Classification and Urban Area Criteria. Retrieved November 4, 2017, from https://www.census.gov/geo/reference/ua/urban-rural-2010.html

US Census Bureau. (2010). Cartographic Boundary Shapefiles - Counties. Retrieved November 26, 2017, from https://www.census.gov/geo/maps-data/data/cbf/cbf_counties.html

U.S. Census Bureau. (2010b). TIGER Products. Retrieved April 10, 2017, from https://www.census.gov/geo/maps-data/data/tiger.html

U.S. Census Bureau. (2010c). TIGER/Line® Shapefiles and TIGER/Line® Files. Retrieved October 28, 2017, from https://www.census.gov/geo/maps-data/data/tiger-line.html

U.S. Census Bureau. (2015). American FactFinder - Results. Retrieved October 31, 2017, from https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_15_1YR_S0801&prodType=table

US Environmental Protection Agency, O. (2013, April 5). Smart Growth [Collections and Lists]. Retrieved November 3, 2017, from https://www.epa.gov/smartgrowth

Velaga, N. R., Beecroft, M., Nelson, J. D., Corsar, D., & Edwards, P. (2012). Transport poverty meets the digital divide: accessibility and connectivity in rural communities. Journal of Transport Geography, 21, 102–112. https://doi.org/10.1016/j.jtrangeo.2011.12.005

Washington, S. ., Karlaftis, M. ., & Mannering, F. . (2010). Statistical and Econometric Methods for Transportation Data Analysis. (2da ed.). Boca Raton, FL: Chapman & Hall.

Welch, T. F., & Mishra, S. (2013). A measure of equity for public transit connectivity. Journal of Transport Geography, 33, 29–41. https://doi.org/10.1016/j.jtrangeo.2013.09.007

Wilkie, W. L., & Weinreich, R. P. (1972). Effects of the Number and Type of Attributes Included in an Attitude Model: More Is Not Better (pp. 325–340). Presented at the Proceedings of the Third Annual Conference of the Association for Consumer Research, Chicago, IL: Association for Consumer Research.

Wu, C., & Murray, A. T. (2005). Optimizing Public Transit Quality and System Access: The Multiple-Route, Maximal Covering/Shortest-Path Problem. Environment and Planning B: Planning and Design, 32(2), 163–178. https://doi.org/10.1068/b31104

Wygonik, E., & Goodchild, A. V. (2016). Urban form and last-mile goods movement: Factors affecting vehicle miles travelled and emissions. Transportation Research Part D: Transport and Environment. https://doi.org/10.1016/j.trd.2016.09.015

Yap, M. D., Correia, G., & van Arem, B. (2016). Preferences of travellers for using automated vehicles as last mile public transport of multimodal train trips. Transportation Research Part A: Policy and Practice, 94, 1–16. https://doi.org/10.1016/j.tra.2016.09.003

APPENDIX A. SURVEY

The Hoosier State train (that is, the Amtrak train that runs four times per week between Indianapolis and Chicago, with stops in Indianapolis, Crawfordsville, Lafayette, Rensselaer, Dyer, and Chicago) is a joint partnership between Iowa Pacific Holdings, Indiana Department of Transportation, Amtrak, and the Cities of Crawfordsville, Lafayette, West Lafayette/Tippecanoe County and Rensselaer since 2015. The joint partnership has resulted in improvements in train performance, reliability, and in onboard amenities, such as Wi-Fi, hot meal services, snacks and beverages. Please take a few minutes to tell us what you think about the Hoosier State train.

☐ Business Class ☐ Coach Class

SECTION 1

1.1 Trip characteristics and experience with the Hoosier State train

1. In which station did you board the Hoosier State train?

☐ Indianapolis ☐ Lafayette ☐ Dyer
☐ Crawfordsville ☐ Rensselaer ☐ Chicago

2. Approximately how many miles did you travel to reach the train station? ______________mi

3a. How did you reach the station?

☐ Drove private car / rental car to the station
☐ Someone dropped me off at the station
☐ Rode a bus
☐ Rode a bicycle
☐ Walked to the station
☐ Took a taxi or ride-sharing service (Uber, Lyft, etc)
☐ Other, please specify_________________

3b. If you drove to reach the station, where did you park?

☐ At the station's parking lot
☐ At a friend's house
☐ On a street near the station
☐ At a parking garage near the station
☐ On a street far (more than a mile) from the station
☐ At a parking garage far (more than a mile) from the station
☐ Other, please specify_________________

4. In which station are you planning to get off the Hoosier State train?

☐ Indianapolis, ☐ Lafayette ☐ Dyer
☐ Crawfordsville ☐ Rensselaer ☐ Chicago

5. Approximately how many miles do you need to travel from the station that you will arrive at to reach your final destination? ______mi

6. How do you plan to reach your final destination when you will get off the train?

☐ Drive private car / rental car
☐ Someone will pick me up
☐ Ride the bus
☐ Ride a bicycle
☐ Walk
☐ Take a taxi or ride-sharing service (Uber, Lyft, etc)
☐ Other, please specify_________________

7. How many times approximately have you taken the Hoosier State train since August 15th, 2015 not including this trip (a single trip counts as one trip and a round trip counts as two trips)?

0_____ 1- 2_____ 3-4_____ 5-6_____ 7-8______ 9-10_____ > 10 ______

8. What is the purpose of your trip today?

- ☐ Work
- ☐ School
- ☐ Social/Recreational
- ☐ Other, please specify ____________

9. Have you ever taken this train as part of a tour or a large group (boys/girls scouts, alumni association, etc.)?

Yes ___ No ___

10. Have you ever used any of the following discounts (Please select all that apply)?

- ☐ Kids ride discount
- ☐ Seniors save 15%
- ☐ 10% off for AAA members
- ☐ 15% off student travel
- ☐ Military personnel and families save 10%
- ☐ Other , please specify__________________
- ☐ Save on group and convention travel
- ☐ Indiana bicentennial promotion (save 15%)
- ☐ Government employee savings (save 20%)
- ☐ Save with a Veterans Advantage card
- ☐ 10% Savings for NARP members

SECTION 2

Please answer the following questions based on your perceptions of passenger rail. There are no right or wrong responses; we are merely interested in your personal opinions. In your responses to the following questions, please share the thoughts that come immediately to mind.

2.1 Ease of using the Hoosier State train

1. My interaction with the ticketing system of the Hoosier State train (Amtrak) is easy and understandable.

Not applicable (Did not buy the ticket by myself) _ Strongly Disagree __ __ Neutral __ Agree__ Strongly Agree__

2. My interaction with the information system (such as Amtrak app, electronic information boards and other systems providing real-time trip information) of the Hoosier State train (Amtrak) is easy and understandable.

Strongly Disagree __ Disagree __ Neutral __ Agree__ Strongly Agree__

3. It is easy for me to reach the closest Hoosier State station from my house.

Strongly Disagree __ Disagree __ Neutral __ Agree__ Strongly Agree__

4a. It is easy for me to park my personal vehicle (car, motorcycle, etc.) near the Hoosier State train station.

Not applicable (I do not own a personal vehicle) __ Strongly Disagree __ Disagree __ Neutral __ Agree__ Strongly Agree

4b. There is enough parking availability near the Hoosier State train station that I use.

Not applicable (I do not own a personal vehicle) __ Strongly Disagree __ Disagree __ Neutral __ Agree__ Strongly Agree__

5a. It is easy for me to access the platform at the Hoosier State train station.

Strongly Disagree __ Disagree __ Neutral __ Agree__ Strongly Agree__

5b. The platform is easily accessible for passengers with disabilities.

Not applicable (I do not have an opinion)___ Strongly Disagree __ Disagree __ Neutral __ Agree__ Strongly Agree__

6. It is easy for me to travel with the essentials for my trip purposes (carry-on luggage, etc.).

Strongly Disagree __ Disagree __ Neutral __ Agree__ Strongly Agree__

7. There is enough available space to store my luggage on the train.

Strongly Disagree __ Disagree __ Neutral __ Agree__ Strongly Agree__

8. The changes in the amenities (e.g., Wi-Fi, hot meal services, snacks and beverages) in the Hoosier State train make my trip more pleasant.

Strongly Disagree __ Disagree __ Neutral __ Agree__ Strongly Agree__

9. It is easy for me to travel with my pet on the Hoosier State train

Not applicable (I do not have a pet) __ Strongly Disagree __ Disagree __ Neutral __ Agree__ Strongly Agree__

10. It is easy for me to find travel brochures related to Indiana destinations at the Hoosier State train stations.

Strongly Disagree __ Disagree __ Neutral __ Agree__ Strongly Agree__

11 Traveling with the Hoosier State train is easy for me.

Strongly Disagree __ Disagree __ Neutral __ Agree__ Strongly Agree__

2.2 Usefulness of the Hoosier State train

1. Using the Hoosier State train would enable me to reach my destination faster.

Very unlikely__ Unlikely __ Neutral __ Likely__ Very likely__

2. Taking the Hoosier State train would make my trip safer.

Very unlikely__ Unlikely __ Neutral __ Likely__ Very likely__

3. Using the Hoosier State train would enable me to use the time it takes to reach my destination more productively.

Very unlikely__ Unlikely __ Neutral __ Likely__ Very likely__

4. When I am traveling alone, using the Hoosier State train to reach my destination would cost me less.

Very unlikely__ Unlikely __ Neutral __ Likely__ Very likely__

5. When I am traveling with a group (family, friends, etc.), using the Hoosier State train to reach my destination would cost me less.

Very unlikely__ Unlikely __ Neutral __ Likely__ Very likely__

6. I find the Hoosier State train useful for my traveling purposes.

Strongly Disagree __ Disagree __ Neutral __ Agree__ Strongly Agree__

2.3 Your thoughts about the Hoosier State train

1. If more people used the Hoosier State train, it would be good for the environment.

Strongly Disagree __ Disagree __ Neutral __ Agree__ Strongly Agree__

2. If more people used the Hoosier State train, it would contribute to the reduction of traffic congestion in Indiana.

Strongly Disagree __ Disagree __ Neutral __ Agree__ Strongly Agree__

3. If more people took the Hoosier State train, it would enhance economic development in Indiana.

Strongly Disagree __ Disagree __ Neutral __ Agree__ Strongly Agree__

4. The State of Indiana should invest funding to support the Hoosier State service.

Strongly Disagree __ Disagree __ Neutral __ Agree__ Strongly Agree__

5. How likely is it that the Hoosier State schedule will be convenient for your travel purposes?

Very unlikely__ Unlikely __ Neutral __ Likely__ Very likely__

6. How likely is it that you can reach your destination on time using the Hoosier state train?

Very unlikely__ Unlikely __ Neutral __ Likely__ Very likely__

2.4 Using the Hoosier State train in the future

1. I intend to travel with the Hoosier State train in the next month.

Strongly Disagree __ Disagree __ Neutral __ Agree__ Strongly Agree__

2. I expect to travel with the Hoosier State train in the foreseeable future.

Strongly Disagree __ Disagree __ Neutral __ Agree__ Strongly Agree__

3. Higher gas prices would make it more likely that I would take the Hoosier State train in the future.

Strongly Disagree __ Disagree __ Neutral __ Agree__ Strongly Agree__

4. Higher parking costs would make it more likely that I would take the Hoosier State train in the future.

Strongly Disagree __ Disagree __ Neutral __ Agree__ Strongly Agree__

5. The availability of a bike-car would make it more likely that I would take the Hoosier State train in the future.

Not applicable (I do not have a bike) __ Strongly Disagree __ Disagree __ Neutral __ Agree__ Strongly Agree__

SECTION 3

3.1 Mode choice

In the following table, please place a check mark on the level of importance each attribute has when choosing a transportation mode for a medium-distance trip [between 3-5 hours travel].

Attribute	Not at all Important	Slightly Important	Moderately Important	Very Important	Extremely Important
a. Cost					
b. Travel time					
c. Comfort					
d. Safety					
e. Amenities (Wi-Fi, food, etc.)					
f. Flexibility of travel (be able to go wherever I want to go)					
g. Convenient/flexible schedule					
h. Reliability (not being late)					
i. Ease of traveling (minimize the effort required to travel)					

Now, please imagine that you are trying to choose between driving alone, carpool (sharing ride), intercity bus, intercity train (such as the Hoosier State train), or airplane for a medium-distance trip [between 3-5 hours travel]. For each of the following transportation modes, rate each attribute by using a score from 1 to 5 where 1 = poor, 2 = fair, 3 = neutral, 4 = good, and 5 = very good.

Attribute	Automobile-Drive Alone	Automobile-Carpool	Intercity Bus (e.g., Greyhound)	Intercity Train (e.g., Amtrak)	Airplane
a. Cost	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
b. Travel time	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
c. Comfort	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
d. Safety	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
e. Amenities (Wi-Fi, food, etc.)	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
f. Flexibility of travel (be able to go wherever I want to go)	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
g. Convenient/flexible schedule	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
h. Reliability (not being late)	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
i. Ease of traveling (minimize the effort required to travel)	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5

- Whether I go to work or go shopping, I almost always travel by car.

Strongly Disagree __ Disagree __ Neutral __ Agree__ Strongly Agree__

SECTION 4

4.1 Now a few last demographic questions

1. Are you male __ or female __?

2. Do you have a disability (or impairment) that may affect your travel needs or experience?

Yes ___ No ___ I prefer not to answer__

3. What is your age range? 18-24 __ 25-34 __ 35-44 __ 45-54__ 55-64__ 65 and over__

4. What describes best your employment situation?

Work full time__ Work part time__ Currently unemployed __ Student__ Retired __ Homemaker___
Other, please specify________

5. Please indicate your approximate annual household income before taxes. (Include total income of all adults living in your household.)

Under $25,000__ $25,000-$49,999__ $50,000-$74,999__ $75,000-$99,999__ $100,000-$149,999__ $150,000 and over__

6. What is your highest level of education?

Grade school or less__ Some high school__ High school graduate__ Technical training beyond high school__

Some college__ College graduate__ Graduate school__

7. Including yourself, how many persons are in your household? One__ Two__ Three__ Four__ Five or more__

8. Please indicate the number of children in your household under the age of 18.

None__ One__ Two__ Three__ Four or more__

9. How many personal vehicles (including cars, trucks, motorcycles, etc.) does your household have access to or own?

None__ One__ Two__ Three__ Four or more__

10. In a typical week, how many miles do you drive your personal vehicle?

I do not own a personal vehicle__ 5-99__ 100-299__ 300-499__ 500-1,000__ More than 1,000__

11. Do you live in Indiana? Yes__ No__

If no, which state do you live in? ________________

12. In which Indiana county is your house located?

I do not live in Indiana__ Jasper__ Lake__ Marion__ Montgomery__ Tippecanoe__

Bartholomew__ Hamilton__ Hancock__ Hendricks__ Johnson__ Madison__ Monroe__ Morgan__ Newton__ Porter __ Putman__ If other, please specify ________________

13. In which city is your house located?

I do not live in Indiana__ Crawfordsville__ Dyer__ Indianapolis__ Lafayette or West Lafayette__ Rensselaer__

Other, please specify ________________

Thank you for your participation!

APPENDIX B. DISTANCE TO REACH THE STATION

The following figures represent the distance that riders were willing to travel to reach a station in order to take the train. To represent the distance, 4 buffers were created for each station: 0-10 miles, 10-30 miles, 30-60 miles and more than 60 miles. The highest proportion of respondents that took the train traveled short distances (less than 10 miles) in order to reach the respective station.

In specific, 44% of respondents that took the train from Indianapolis traveled less than 10 miles to reach the station, 36% of them traveled between 10-30 miles and 13% of them traveled between 30-60 miles. 7% of respondents traveled more than 60 miles in order to take the train from Indianapolis.

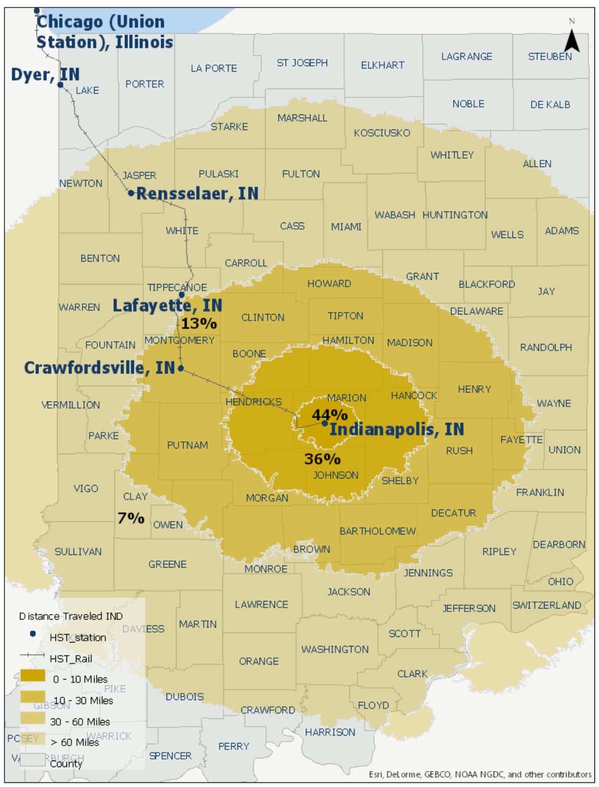


Figure A.1 Distance traveled to reach Indianapolis station.

79% of the respondents that took the train from Lafayette traveled less than 10 miles, 13% of them traveled between 10-30 miles, 2% of them traveled between 30-60 miles and 6% of respondents traveled more than 60 miles in order to take the train.

53% of respondents that took the train from Dyer traveled less than 10 miles to reach the station, and 47% of them traveled between 10- 30 miles.

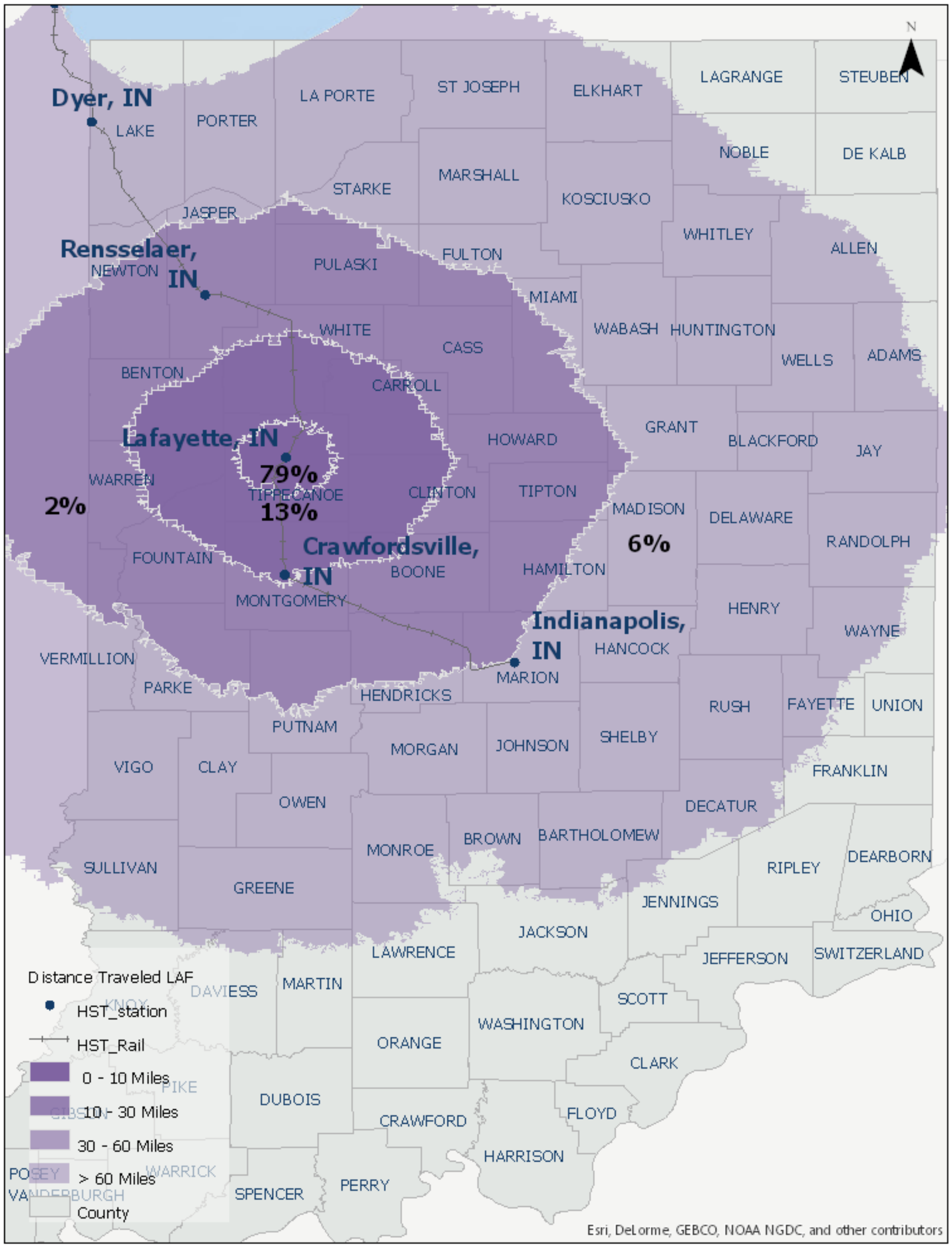


Figure A.2 Distance traveled to reach Lafayette station.

69% of respondents that took the train from Rensselaer, traveled less than 10 miles to reach the station, 8% of them traveled between 10- 30 miles and 23% of them traveled between 30-60 miles.

48% of the respondents traveled less than 10 miles, 15% of them traveled between 10-30 miles, 25% of them traveled between 30-60 miles, and 13% of respondents traveled more than 60 miles in order to take the train from Crawfordsville.

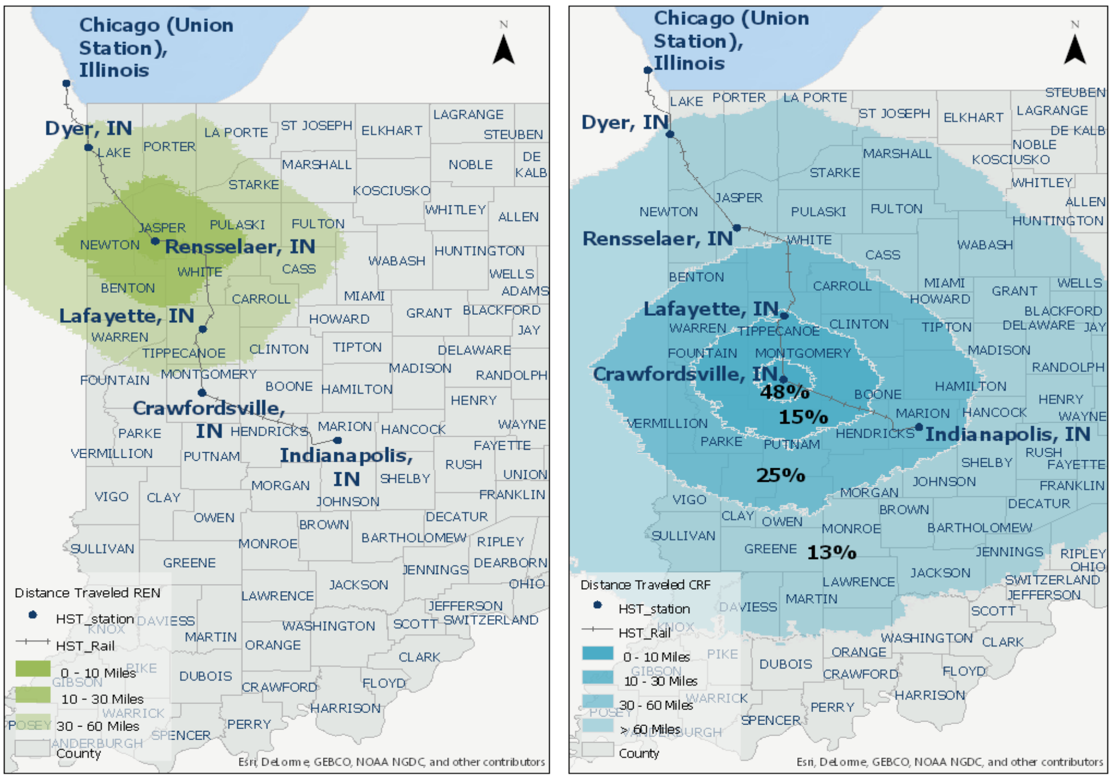


Figure A.3 Distance traveled to reach Rensselaer (left map) and Crawfordsville (right map) stations.

APPENDIX C. SENSITIVITY ANALYSIS

Scenario 1

The first scenario considered a reduction of 33% of the distance from the base scenario. The ranges resulted as:

1 Range 1:0 to 1.5 miles,
2 Range 2: 1.5 to 5.5 miles,
3 Range 3: 5.5 to 19.75 miles, and
4 Range 4: greater than 19.75

The result of this scenario were summarized and presented in Table Table C.1. The first range analyzed showed that driving alone was the preferred mode of transportation for the riders that traveled less than 1.5 miles to reach a station. This is not alike with the results of range 1 in the base case. However, the results for the following three ranges (Range 2, Range 3, and Range 4) were similar to the ones found in the base scenario. For this scenario, respondents in the first three ranges prefer to drive alone while the respondents who traveled a further distance to reach a station chosen to travel by intercity train.

Table C.1 Multiattribute Attitude Model scores – Scenario 1.

	Drive Alone	Train	Carpool	Airplane	Bus
Range 1	125.12	124.76	111.92	111.59	98.82
Range 2	131.06	126.23	114.75	117.36	95.05
Range 3	130.52	129.51	115.82	113.19	95.54
Range 4	131.41	131.52	117.84	114.52	92.80

Figure Figure C.1 shows the discomposed scores. As mentioned, driving alone was the preferred option among the different ranges. Also, *reliability* was the most important factor identified in three out of four ranges. *Amenities* was the less important factor also seen in three of the ranges. Nonetheless, the three most important factors among all the ranges were *reliability*, *safety*, and *ease of use*.

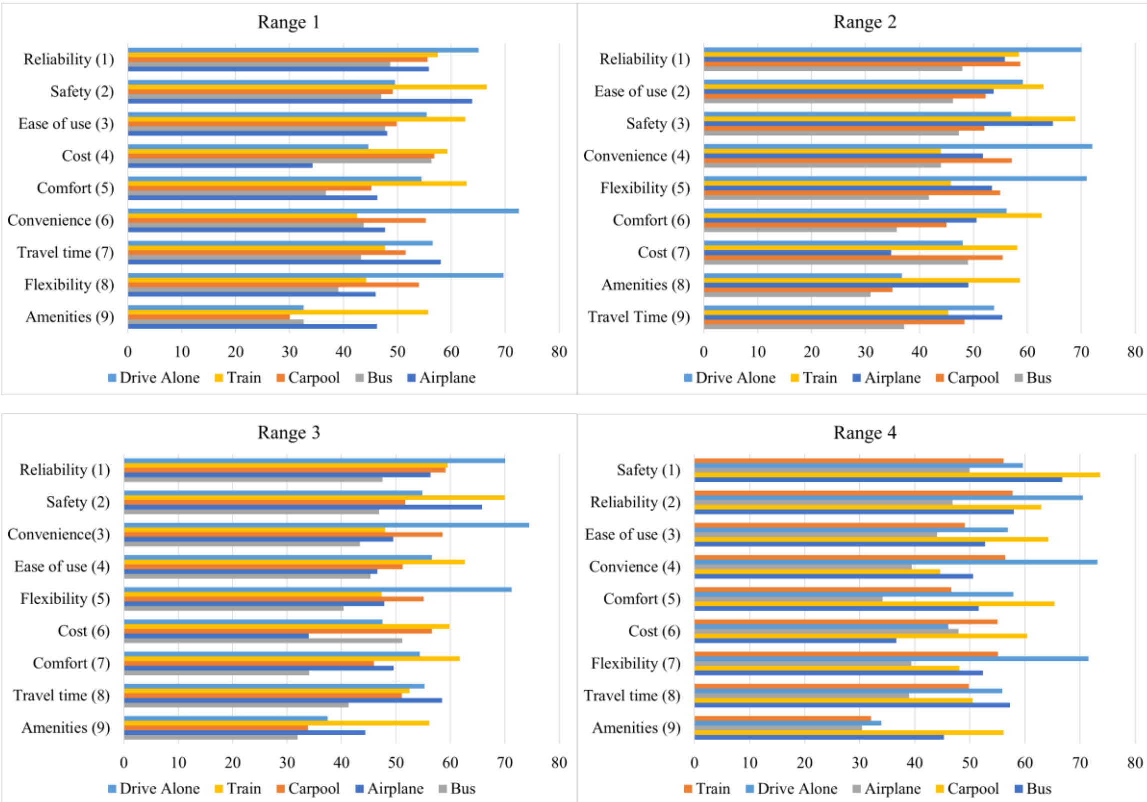


Figure C.1 Average Score per Attribute – Scenario 1

Table C.2 Average Important Rating of Attributes – Scenario 1.

	Range 1		***Range 2***		***Range 3***		***Range 4***	
	Rank	***Mean Score***	***Rank***	***Mean Score***	***Rank***	***Mean Score***	***Rank***	***Mean Score***
Cost	4	3.94	7	3.77	6	3.86	6	3.84
Travel time	7	3.75	9	3.52	8	3.75	8	3.66
Comfort	5	3.81	6	3.79	7	3.83	5	3.95
Safety	2	4.05	3	4.03	2	4.09	1	4.31
Amenities (Wi-Fi, food, etc.)	8	3.48	8	3.61	9	3.54	9	3.48
Flexibility of travel	8	3.66	5	3.81	5	3.89	7	3.84
Convenient	6	3.80	4	3.92	3	4.02	4	3.97
Reliability	1	4.09	1	4.13	1	4.14	2	4.17
Ease of traveling	3	3.98	2	4.12	4	3.98	3	4.13

The analysis results of the important rating of attributes were summarized in Table Table C.2. The attributes changes in some of the ranges. For instance, *cost* was located as the fourth most important attribute for passenger who traveled less than 1.5 miles. However, it was located around the six position for passenger who traveled further than 1.5 miles to reach a station. For those passenger, *convenient* and sometimes, *comfort* were more important than *cost*. *Travel time* was again ranked for most of the ranges as not important factor.

Scenario 2

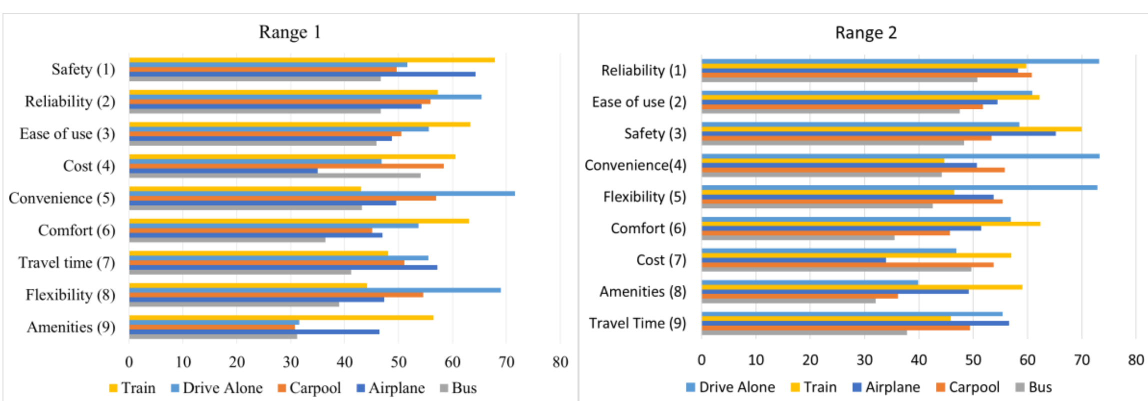
The second scenario considered an increase of 33% of the distance from the base scenario. The ranges resulted as:

1 Range 1:0 to 2.5 miles,
2 Range 2: 2.5 to 8.5 miles,
3 Range 3: 8.5 to 28.25 miles, and
4 Range 4: greater than 28.25

The result of this scenario were summarized and presented in Table C.3. The first range analyzed showed that intercity train was the preferred mode of transportation for the riders that traveled less than 1.5 miles to reach a station. This is alike with the results of Range 1 in the base case. The results for the following three ranges (Range 2, Range 3, and Range 4) were similar to the ones found in the base scenario. For this scenario, respondents in the closest and furthest ranges prefer to use intercity train while the respondents who traveled intermediate distance to reach a station chosen to travel by driving alone.

Table C.3 Multiattribute Attitude Model scores – Scenario 2.

	Drive Alone	Train	Carpool	Airplane	Bus
Range 1	125.18	125.97	113.24	112.49	96.11
Range 2	134.44	126.86	115.56	118.41	97.11
Range 3	131.74	129.29	114.99	114.51	94.02
Range 4	129.80	131.63	114.05	117.8911	92.93



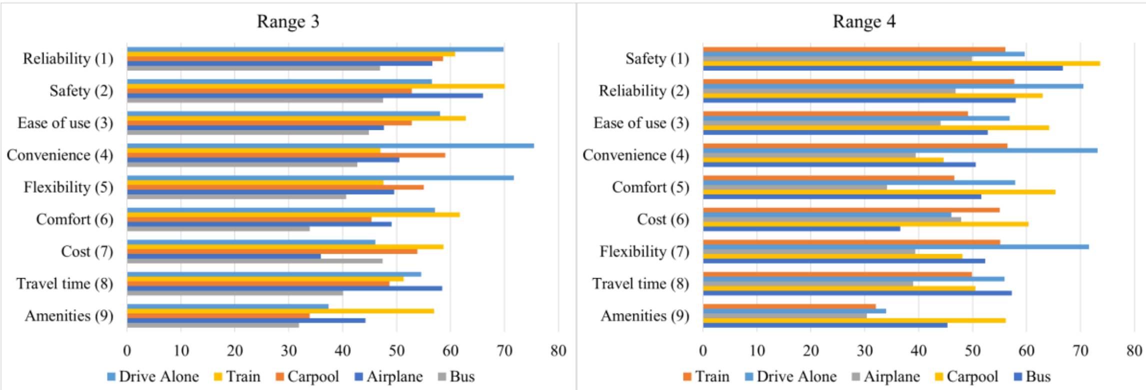


Figure C.2 Average Score per Attribute – Scenario 2

The average score per attribute was shown in Figure Figure C.2. Intercity train received a high score related to *safety*, *ease of use*, and *comfort* for most of the ranges. However, other important factors such as *reliability* and *convenience* were not high for train comparing to other modes such as driving alone.

Table C.4 Average Important Rating of Attributes – Scenario 2.

	Range 1		***Range 2***		***Range 3***		***Range 4***	
	Rank	***Mean Score***	***Rank***	***Mean Score***	***Rank***	***Mean Score***	***Rank***	***Mean Score***
Cost	4	3.90	7	3.77	7	3.80	6	3.89
Travel time	7	3.67	9	3.61	8	3.70	8	3.62
Comfort	6	3.78	6	3.83	6	3.88	5	3.95
Safety	1	4.04	3	4.07	2	4.15	1	4.27
Amenities (Wi-Fi, food, etc.)	9	3.49	8	3.65	9	3.56	9	3.43
Flexibility of travel	8	3.66	5	3.88	5	3.89	7	3.84
Convenient	5	3.81	4	3.97	4	4.04	4	3.95
Reliability	2	4.02	1	4.27	1	4.16	2	4.11
Ease of traveling	3	4.02	2	4.11	3	4.08	3	4.08

For this scenario, *safety* and *reliability* were the most important factors when a passenger is choosing an intercity transportation mode. These factors change from the first place to the second in the ranges analyzed. *Amenities* was most of the time the less important attribute according to respondents in three of the fourth ranges (Table Table C.4).

Scenario 3

The third scenario considered a reduction of 50% of the distance from the base case. The ranges resulted as:

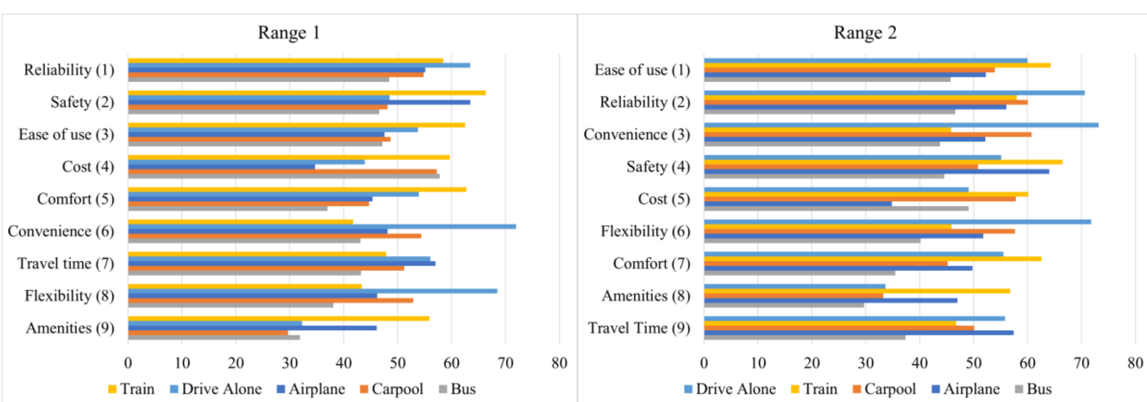
1 Range 1:0 to 1 miles,
2 Range 2: 1 to 4 miles,
3 Range 3: 4 to 15.5 miles, and
4 Range 4: greater than 15.5

The result of this scenario were summarized and presented in Table Table C.5. The first range analyzed showed that intercity train was the preferred mode of transportation for the riders that traveled less than 1 mile to reach a station. This is alike with the results of range 1 in the base case. The results for the following two ranges (Range 2, and range 3) were similar to the ones found in the base scenario. Range four this time resulted in driving alone as their preferred mode of intercity travel. For this scenario, respondents in the closest range prefer to travel by intercity train while the respondents who traveled further distance to reach a station chosen to travel by driving alone.

Table C.5 Multiattribute Attitude Model scores – Scenario 3.

	Drive Alone	**Train**	**Carpool**	**Airplane**	**Bus**
Range 1	123.14	124.68	110.44	98.36	110.93
Range 2	131.20	126.71	117.42	116.38	93.09
Range 3	131.22	128.19	114.66	116.03	97.26
Range 4	131.45	131.04	114.17	116.87	92.62

The average score per attribute showed again that reliability is one of the most important factors considered by the riders. In this scenario, range 2 (who traveled more than 1 mile but less than 4 miles) chose ease of use as the most important factor. This attribute had a high score for train.



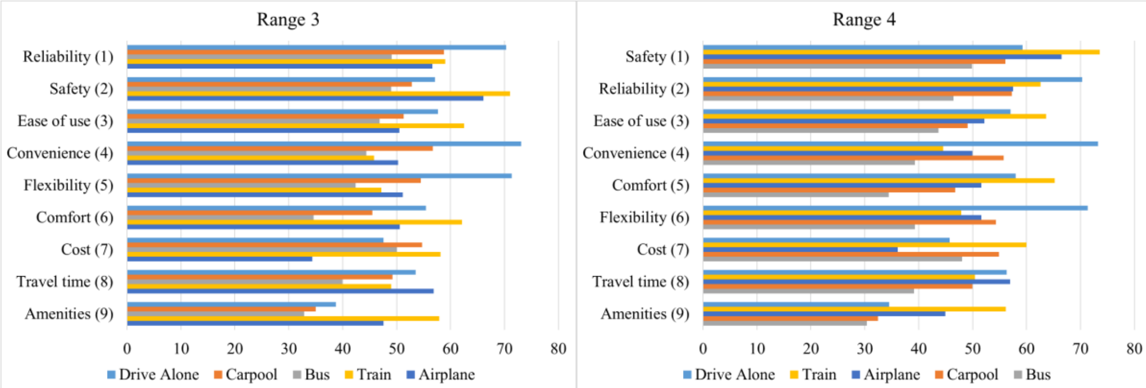


Figure C.3 Average Score per Attribute – Scenario 3

Amenities was again the less important factor in three out of four ranges. This is akin with the base case and previous scenarios. Again, the *cost* attribute was located as fourth in the closest range but it did not seem important with further distance analyzed. Still, *travel time* was positioned as a less important factor among the ones ranked.

Table C.6 Average Important Rating of Attributes – Scenario 3.

	Range 1		***Range 2***		***Range 3***		***Range 4***	
	Rank	***Mean Score***	***Rank***	***Mean Score***	***Rank***	***Mean Score***	***Rank***	***Mean Score***
Cost	4	3.95	5	3.86	7	3.81	7	3.80
Travel time	7	3.76	8	3.58	8	3.64	8	3.65
Comfort	5	3.79	7	3.75	6	3.85	5	3.94
Safety	2	4.04	4	3.93	2	4.13	1	4.30
Amenities (Wi-Fi, food, etc.)	9	3.47	9	3.53	9	3.64	9	3.47
Flexibility of travel	8	3.62	6	3.82	5	3.88	6	3.82
Convenient	6	3.78	3	3.94	4	3.98	4	3.96
Reliability	1	4.08	2	4.11	1	4.17	2	4.14
Ease of traveling	3	3.97	1	4.12	3	4.05	3	4.11

Scenario 4

The third scenario considered an increase of 50% of the distance from the base scenario. The ranges resulted as:

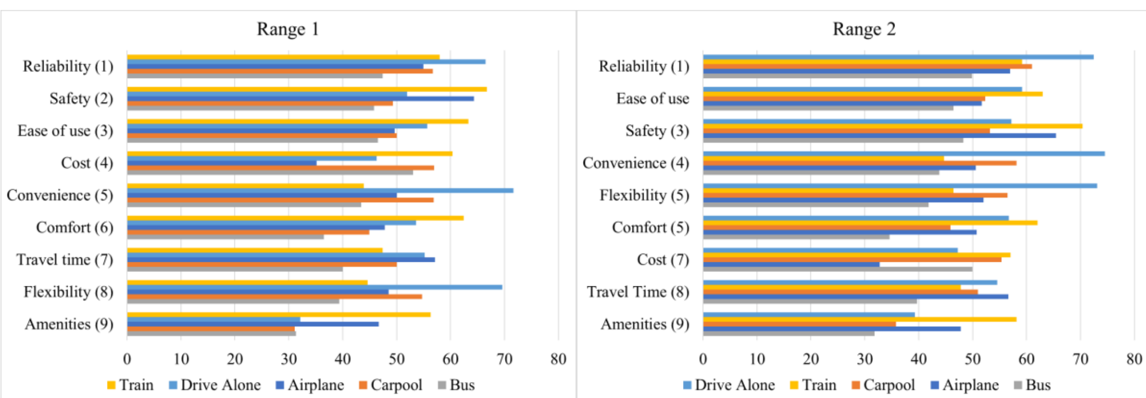
1 Range 1:0 to 3 miles,
2 Range 2: 3 to 10 miles,
3 Range 3: 10 to 32.5 miles, and
4 Range 4: greater than 32.5

The result of this scenario were summarized and presented in Table Table C.7. The first range analyzed showed that intercity train was the preferred mode of transportation for the riders that traveled less than 3 mile to reach a station. This is alike with the results of range 1 in the base case. The results for the following three ranges (Range 2, range 3, and range 4) were similar to the ones found in the base scenario. For this scenario, respondents in the closest and furthest range prefer to travel by intercity train while the respondents who traveled from intermediate distances to reach a station chosen to travel by driving alone.

Table C.7 Multiattribute Attitude Model scores – Scenario 4.

	Drive Alone	Train	Carpool	Airplane	Bus
Range 1	125.67	125.79	112.68	113.57	95.89
Range 2	133.62	127.22	117.33	116.15	96.61
Range 3	132.32	131.50	115.67	119.52	94.26
Range 4	129.36	130.75	110.48	113.24	130.75

Figure C.4 shows the average score per attribute for the four ranges in this scenario. *Safety*, and *ease of use* were high ranked high for intercity train. Also, *comfort* and *amenities* had high values for this mode.



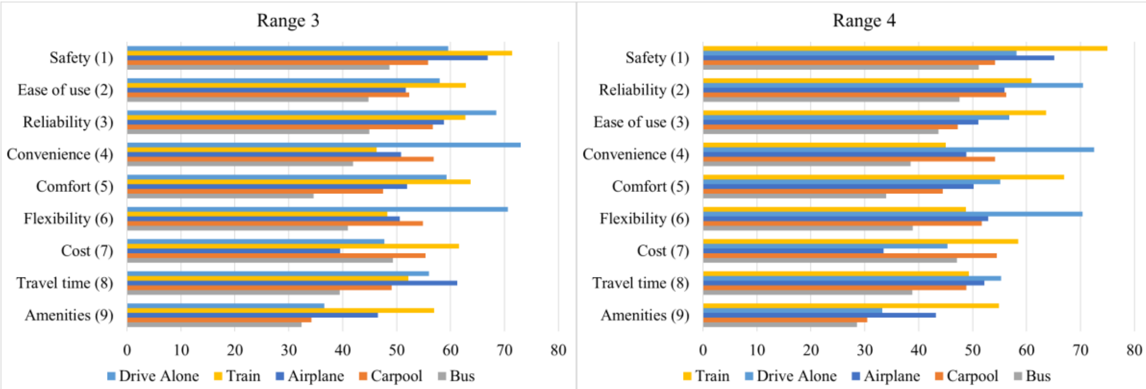


Figure C.4 Average Score per Attribute – Scenario 4

In this scenario, *amenities* was the less important factor in all the ranges analyzed. Also, *safety* and *reliability* were again rating as important factors to choose a transportation mode. *Cost* was again fourth in range 1 but it is not same important in the further ranges. *Travel time* was also not around the most important factors to consider in mostly all of the ranges.

Table C.8 Average Important Rating of Attributes – Scenario 4.

	Range 1		***Range 2***		***Range 3***		***Range 4***	
	Rank	***Mean Score***	***Rank***	***Mean Score***	***Rank***	***Mean Score***	***Rank***	***Mean Score***
Cost	4	3.89	7	3.80	6	3.85	7	3.74
Travel time	8	3.64	8	3.65	8	3.75	8	3.49
Comfort	6	3.75	6	3.86	5	3.94	5	3.89
Safety	3	4.01	3	4.08	1	4.26	1	4.25
Amenities (Wi-Fi, food, etc.)	9	3.50	9	3.62	9	3.56	9	3.39
Flexibility of travel	7	3.69	5	3.91	7	3.83	6	3.83
Convenient	5	3.83	4	4.02	4	3.94	4	3.97
Reliability	1	4.05	1	4.25	3	4.09	2	4.12
Ease of traveling	2	4.02	2	4.11	2	4.10	3	4.03

T-test

To test if the changes in preference of the modes were significant, a t-test one tail unequal sample size and unequal variance was considered. The null hypothesis considered is the value obtained in range 1 is different from the one obtained in the following ranges. The test was performed for the five scenarios: base scenario, scenario 1, scenario 2, scenario 3, and scenario 4. The goal of this analysis was to determine whether the distance would affect the change in the value obtained as a Total Average Score through a MA model in each mode studied. Driving alone, carpooling, intercity bus, intercity train, and airplane were tested for each of the ranges. All the values obtained in each range were compared with the range 1. The base scenario of changes resulted to have a significant change at 10% level of confidence between the values given for driving alone comparing to the Range 1 (0 to 2 miles) and Range 3 (7 to 24 miles) as seen in Table Table C.9.

Table C.9 Results of T-test for car mode –Base case

	Range 1	***Range 2***	***Range 3***	***Range 4***
Mean	125.52	132.69	132.55	130.36
Observations	145	120	114	103
df		235	234	209
t Stat		-1.11	-1.29	-0.80
P(T<=t) one-tail		0.13	0.10	0.21
t Critical one-tail		1.65	1.65	1.65
P(T<=t) two-tail		0.27	0.20	0.42
t Critical two-tail		1.97	1.97	1.97

The results for carpooling mode did not change significantly according to the t-test results. As seen in Table Table C.10, the change of the total score rank between the different ranges did not result significant when those were compared to the first range.

Table C.10 Results of T-test for carpool mode –Base case

	Range 1	***Range 2***	***Range 3***	***Range 4***
Mean	113.59	114.74	114.63	114.88
Observations	141	118	112	103
df		214	204	186
t Stat		0.56	0.31	0.27
P(T<=t) one-tail		0.29	0.38	0.40
t Critical one-tail		1.65	1.65	1.65
P(T<=t) two-tail		0.57	0.76	0.79
t Critical two-tail		1.97	1.97	1.97

Intercity bus had a significant change at the 10% level of confidence when the values where evaluated from Range 1 (0 to 2 miles) to Range 4 (more than 24 miles). The comparison of the Range 1 with the other Ranges did not result in significant changes.

Table C.11 Results of T-test for intercity bus –Base case

	Range 1	*Range 2*	*Range 3*	*Range 4*
Mean	96.49	97.08	93.84	92.91
Observations	141	115	108	104
df		251	238	212
t Stat		0.792845	0.95941	1.591935
P(T<=t) one-tail		0.214308	0.169163	0.056445
t Critical one-tail		1.650947	1.651281	1.652073
P(T<=t) two-tail		0.428617	0.338326	0.112889
t Critical two-tail		1.96946	1.969982	1.971217

The changes between ranges for intercity train did not result to be significant according to t-test results. The hypothesis that the distance will affect the selection of the train as an intercity transportation mode was not confirm through this analysis. However, as shows in Table Table C.11, the values between Range 2 and Range 3 change 5 points. This means that people who traveled more distance to reach a station valued more the nine attributes studied for a train mode.

Table C.12 Results t-test intercity train mode – base case

	Range 1	*Range 2*	*Range 3*	*Range 4*
Mean	126.05	125.53	130.59	131.135
Observations	150	119	116	105
df		251	209	193
t Stat		0.711	-0.348	-0.158
P(T<=t) one-tail		0.238	0.3633	0.437
t Critical one-tail		1.650	1.652	1.652
P(T<=t) two-tail		0.477	0.727	0.874
t Critical two-tail		1.969	1.971	1.972

The airplane mode did not resulted in significant changes between the ranges (Table C.13)

Table C.13 Results t- test airplane mode – base case

	Range 1	*Range 2*	*Range 3*	*Range 4*
Mean	113.15	115.93	115.45	118.40
Observations	147	118	110	109
df		254	233	211
t Stat		-0.16	-0.078	0.218
P(T<=t) one-tail		0.436	0.468	0.413
t Critical one-tail		1.650	1.651	1.652
P(T<=t) two-tail		0.873	0.937	0.826
t Critical two-tail		1.969	1.970	1.971

The analysis carried for the base case was also conducted for the alternative 4 scenarios. For driving alones, the changes in the MAM total score were found significant in the scenario 2 between Range 1 and Range 2 at the 10% of confidence level. Scenario 3 was found with significant changes at the 10% level of confidence between Range 1, and Range 3 and, between Range 1 and Range 2 (Table C.14). Scenario 4 between Range 1 (0 to 3 miles) and Range 2 (3 to 10 miles) was also found to have a significant change. Those changes were significant at the 10% level of confidence. Significant changes were also seen in the bus mode at 10% level of significance. The changes are significant at that level in all cases between Range 1 (From 0 to 2, o to 2.5, 0 to 1.5, 0 to 1, and 0 to 3 miles) and Range 4 (Greater than 24, greater than 19.75, greater than 28.55, greater than 15.5, and greater than 32 miles).

Table C.14 Results t- test driving alone – Scenario 3.

	Range 1	***Range 2***	***Range 3***	***Range 4***
Mean	125.67	133.62	132.31	129.35
Observations	76	120	76	135
df		177	161	157
t Stat		1.094	-1.430	-1.572
P(T<=t) one-tail		0.137	0.077	0.058
t Critical one-tail		1.649	1.654	1.654
P(T<=t) two-tail		0.142	0.154	0.117
t Critical two-tail		1.973	1.974	1.975

The result of the different t-test did not find significance changes between Total Score Rank for the intercity train mode. The train mode was not identify as having a significance change when the distance from the station was increasing. This do happens when other modes are considered such as driving alone or bus. If a person need to reach a bus station that is close to its origin of travel, this mode would be more likely to be taken than if the person needs to travel from further destinations. This can be seen for the total score rank given to bus for the Range 1 and Range 4 in the different scenarios. The opposite would befall with the driving alone option. The further the person would need to travel, the highest the score for this mode would be. That means that if a person would need to travel more than 24 miles to reach a station, as the base case hypnotized, this person would prefer driving alone than a person who is closer to the station.

APPENDIX D. ORDERED PROBIT MODEL OUTPUTS

Model Output

```
-----------------------------------------------------------------------------
Ordered Probability Model
Dependent variable XX1
Log likelihood function -861.00748
Restricted log likelihood -897.6301
Chi squared [11](P=.000) 73.24525
Significance level 0.000
McFadden Pseudo R-squared 0.0407992
Estimation based on N= 879, K=13
Inf.Cr.AIC=1748 AIC/N=1.989
Underlying probabilities based on Normal
--------+--------------------------------------------------------------------
        |                  Standard            Prob.      95% Confidence
     XX1|  Coefficient       Error       z    |z|>Z*         Interval
--------+--------------------------------------------------------------------
        |Index function for probability......................................
Constant|   .46848***       0.08562     5.47    0          0.30068    0.63628
     XX5|  -.27559***       0.08363    -3.3    0.001     -0.43949   -0.11169
     XD1|   .14142*         0.08109     1.74   0.0812    -0.01753    0.30036
     X15|   .00045*         0.00026     1.71   0.0881    -0.00007    0.00096
   XX341|  -.22171***       0.08296    -2.67   0.0075    -0.38432   -0.0591
    XX36|   .39194***       0.11398     3.44   0.0006     0.16854    0.61534
     337|   .00044*         0.00026     1.71   0.0876    -0.00006    0.00095
   XX211|  -.19881**        0.09767    -2.04   0.0418    -0.39023   -0.00738
    X333|   .00044*         0.00027     1.66   0.0976    -0.00008    0.00096
    X326|  -.00086***       0.00028    -3.05   0.0023    -0.00141   -0.00031
   XX201|  -.18748*         0.11381    -1.65   0.0995    -0.41054    0.03557
    X321|  0.00021          0.00027     0.77   0.4442    -0.00032    0.00074
        |Threshold parameters for index......................................
  Mu(01)|  1.31957***       0.05668    23.28    0          1.20849    1.43066
--------+--------------------------------------------------------------------
***, **, * ==>  Significance at 1%, 5%, 10% level.
Model was estimated on Jul 28, 2017 at 00:40:29 AM
```

Marginal Effects Output

Marginal effects for ordered probability model
M.E.s for dummy variables are Pr[y|x=1]-Pr[y|x=0]
Names for dummy variables are marked by *.

XX1	Partial Effect	Elasticity	z	Prob. |z|>Z*	95% Confidence Interval	
	[Partial effects on Prob[Y=00] at means]					
*X15	-.07221***	-.19345	-2.85	.0044	-.12192	-.02249
*X18	.07524***	.20156	2.72	.0065	.02108	.12939
*XX36	-.13719***	-.36754	-3.69	.0002	-.21003	-.06435
XX156	.00012*	.7839D-04	1.76	.0784	-.00001	.00025
*XX204	-.08219**	-.22020	-2.48	.0131	-.14709	-.01730
*XX211	.08480**	.22720	2.24	.0250	.01062	.15898
*X326	.00020**	.00054	2.42	.0156	.00004	.00036
*X333	-.00027***	-.00073	-3.32	.0009	-.00043	-.00011
	[Partial effects on Prob[Y=01] at means]					
*X15	.57913D-04*	0.000121	.70	0.089	-.88251D-05	.12465D-03
*XX5	-.03864***	-0.08227	-2.88	0.004	-0.06493	-0.01234
*XD1	.01801*	0.038351	.72	0.0851	-0.00249	0.03851
*X321	.26850D-04	.5717D-040	.76	0.4494	-.42715D-04	.96415D-04
*X326	-.00011***	-0.00024	-2.77	0.0057	-0.00019	-0.00003
*X333	.57014D-04	0.000121	.64	0.1013	-.11186D-04	.12521D-03
*X337	.56814D-04*	0.000121	.70	0.0892	-.87057D-05	.12233D-03
*XX36	.02978***	0.063423	.72	0.00020	.01409	0.04547
*XX341	-.03065**	-0.06526	-2.40	0.0164	-0.05568	-0.00562
*XX201	-0.02866	-0.06103	-1.42	0.1552	-0.06817	0.01085
*XX211	-.02960*	-0.06304	-1.77	0.0763	-0.06234	0.00313
	[Partial effects on Prob[Y=02] at means]					
*X15	.00011*	0.00068	1.68	0.0921	-0.00002	0.00024
*XX5	-.06592***	-0.40186	-3.36	0.0008	-0.10438	-0.02745
*XD1	.03510*	0.21399	1.73	0.0837	-0.00468	0.07488
*X321	.51444D-04	0.00031	0.77	0.4436	-.80175D-04	.18306D-03
*X326	-.00021***	-0.00129	-3.03	0.0024	-0.00035	-0.00007
*X333	0.00011	0.00067	1.64	0.1005	-0.00002	0.00024
*X337	.00011*	0.00066	1.69	0.0913	-0.00002	0.00024
*XX36	.10923***	0.66592	3.07	0.0021	0.03949	0.17896
*XX341	-.05338***	-0.32542	-2.72	0.0065	-0.09183	-0.01492
*XX201	-.04329*	-0.26392	-1.76	0.0779	-0.09141	0.00483
*XX211	-.04644**	-0.28314	-2.15	0.0318	-0.08883	0.00405

z, prob values and confidence intervals are given for the partial effect
nnnnn.D-xx or D+xx => multiply by 10 to -xx or +xx.
***, **, * ==> Significance at 1%, 5%, 10% level.
Model was estimated on Jul 28, 2017 at 00:40:29 AM

Cross Tabulation Output

```
--------------------------------------------------------------
Counts of Correct Predictions Out of         637 Observations
------- Method ---------------------------- --- Hits --- Rate
1. Index of the most probable outcome          293      .46%
2. Interval (-inf,0],(0,mu1]... containing x   287      .45%
3. Probability weighted average of outcomes    290      .46%
--------------------------------------------------------------
--------------------------------------------------------------
Cross tabulation of predictions and actual outcomes
Prediction is number of the most probable cell.
Row = actual, Column = Prediction, # = Correct, Model = Probit
+--------+-----+-----+-----*-----+
|     XX1|    0|    1|    2*Total|
+--------+-----+-----+-----*-----+
|    0   |#  74|  167|    0*  241|
+--------+-----+-----+-----*-----+
|    1   |   62|# 219|    0*  281|
+--------+-----+-----+-----*-----+
|    2   |   12|  103|#   0*  115|
*********************************
|  Total |  148|  489|    0*  637|
+--------+-----+-----+-----*-----+
--------------------------------------------------------------
Cross tabulation of outcome and interval containing xb
Row = actual, Column = Prediction, # = Correct, Model = Probit
+--------+-----+-----+-----*-----+
|     XX1|    0|    1|    2*Total|
+--------+-----+-----+-----*-----+
|    0   |#  42|  199|    0*  241|
+--------+-----+-----+-----*-----+
|    1   |   36|# 245|    0*  281|
+--------+-----+-----+-----*-----+
|    2   |    8|  107|#   0*  115|
*********************************
|  Total |   86|  551|    0*  637|
+--------+-----+-----+-----*-----+
--------------------------------------------------------------
Cross tabulation of outcomes and predicted probabilities.
Value(j,m)=Sum(i=1,N)y(i,j)*p(i,m).
Totals may not match cell sums because of rounding error.
Row = actual, Column = Prediction, # = Correct, Model = Probit
+--------+-----+-----+-----*-----+
|     XX1|    0|    1|    2*Total|
+--------+-----+-----+-----*-----+
|    0   |#  96|  108|   38*  241|
+--------+-----+-----+-----*-----+
|    1   |  103|# 129|   49*  281|
+--------+-----+-----+-----*-----+
|    2   |   38|   54|#  23*  115|
*********************************
|  Total |  236|  290|  110*  637|
+--------+-----+-----+-----*-----+
```