CHARACTERISTICS

OF

TRANSIT-FRIENDLY CITIES

by

Sarah Doggett

A thesis submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the Honors Degree in Civil Engineering with Distinction

Spring 2016

CHARACTERISTICS

OF

TRANSIT-FRIENDLY CITIES

by

Sarah Doggett

Approved: ____
Sue McNeil, Ph.D.
Professor in charge of thesis on behalf of the Advisory Committee

Approved: ____
Earl Lee II, Ph.D.
Committee member from the Department of Civil Engineering

Approved: ____
April Kloxin, Ph.D.
Committee member from the Board of Senior Thesis Readers

Approved: ____
Michael Arnold, Ph.D.
Directory, University Honors Program

ACKNOWLEDGMENTS

I would like to thank Dr. Sue McNeil for mentoring and guiding me throughout the past three years. I also thank the Undergraduate Research Program for providing me with the funds to attend the 95th Transportation Research Board Annual Meeting where I learned of new methods and data sources that augmented my research. I thank Dr. Ameile Davis for explaining her research to me. I thank the Summer Scholars Program for giving me the opportunity to begin researching. I thank my parents for supporting me and pushing me to challenge myself.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

This thesis analyzes the relationships among the built environment, socioeconomics, and transit system attributes with respect to transit ridership in the 96 most populated urban areas in the United States. It concludes that the relative importance of these characteristics varies amongst several different types of city. After gathering extensive data on these characteristics for each of the 96 cities, a factor analysis, a cluster analysis, and a regression analysis were conducted to reveal patterns in the data. The analysis showed that the data could be grouped into underlying input and environment factors. These factors are used to identify like cities using cluster analysis. The thesis is divided into five sections: introduction, literature review, methodology, results, and conclusion.

Chapter 1

INTRODUCTION

While researchers have identified many factors that influence the ridership of public transportation, the relative importance of these factors has not been established. My research attempts to answer this question: To what extent do the built environment, socioeconomics, and transit system characteristics determine transit ridership? My hypothesis was that cities with high transit quality, high land use mix, high population and employment density, and low parking coverage will have high rates of transit ridership.

This research is valuable to the field of transportation as it can help planners better understand how to build an efficient, well utilized transit system. The work builds upon the work of past studies that have explored the relationships between transit ridership and variables of transit infrastructure, built environment, and socioeconomics only in isolation. Few studies have analyzed the effect of more than one characteristic of a city or have compared cities on a macro scale. Robert Cervero's Transit Metropolis is one study that has done both, but while it provided much insight and inspiration for my project, its analysis is qualitative (Cervero, 1998). By analyzing all of these variables together, new patterns and relationships could be identified. Future transportation and planning policies could incorporate these new relationships when trying to improve transit ridership. Attention and effort could be redirected from variables that are revealed to have less influence to the variables deemed most influential. Increasing the use of public transit is incredibly important in order to create a more sustainable future. Transit reduces the number of trips made by inefficient gas-guzzling automobiles, strengthens local economies, and increases equity by giving everyone access to opportunities.

To explore the relationship between cities and their transit, I created a database containing variables relating to the transit ridership, transit system, demographic, built environment, socioeconomic, and land use characteristics for the 96 most populated urban areas in the United States. By analyzing this database, I attempted to determine the relative importance of new technologies, including real-time information and bike sharing, in determining transit ridership when compared to traditional factors, such as demographics, transit infrastructure, and the built environment of the city. After normalizing the data and eliminating highly correlated variables, I conducted a factor analysis on the remaining variables to detect any underlying relationships between the variables. This left me with a manageable number of factors to analyze for each city. Through a cluster analysis, I grouped cities that shared similar attributes together and determined that each group of cities had a different set of attributes that influenced transit ridership. Finally, I performed a regression analysis on the entire dataset.

The following chapter contains a detailed discussion of my literature review and the background of my variable selection. Chapter 3 contains my methodology, with subsections dedicated to each step of the research. Chapter 4 reports my analysis of the results. Chapter 5 is the conclusion which summarizes the results and policy implications.

Chapter 2

LITERATURE REVIEW

There has been a lot of research into the effects of the built environment on transit patterns. In this chapter, I present evidence from the literature on variables that influence transit ridership.

2.1 Built Environment

When analyzing its effect on transit ridership, the built environment is often defined in terms of the 3Ds: density, diversity, and design. Although researchers debate the relative importance of each of the 3Ds, Cervero and Kockelman have established that "density, land-use diversity, and pedestrian-oriented designs generally reduce trip rates and encourage non-auto travel in statistically significant ways" (Cervero & Kockelman, 1997).

According to Frank and Engelke, high densities reduce the cost of providing transit (Frank & Engelke, 2001). Additionally, Maat and Harts state that transit works well in high densities because car usage is discouraged (Maat & Harts, 2001). There are many other reasons for density's effect on transit – there are more destinations within shorter distances, more options for travel, more people to share the cost of amenities such as sidewalks, and it is associated with higher land use mixes (Litman & Steele, 2016). Studies done by Ewing and others have shown that increases in density have a strong, significant effect on transportation; as density increases, so does the number of people using public transit for commuting, while the average vehicle ownership and the vehicle miles per capita decrease (Ewing et al, 2003). According to Ewing et al, “density alone is associated with a 10 percentage point increase in public transportation use between high-density and low-density areas” (Ewing et al, 2003). Both population and employment density have been shown to have an effect on transit usage. Litman and Steele state that “automobile commuting declines significantly when workplace densities reach 50-75 employees per gross acre” (Litman & Steele, 2016).

Despite these studies, the link between land use and transit is still controversial. Many scholars admit a failure to account for other important variables, as “while there is a consensus that density is correlated with travel behavior, there is also a general understanding that density may represent only one of a combination of influences on travel behavior” or it might only indicate the presence of other factors, such as land diversity, which are the real source of the relationship (Frank & Engelke, 2001).

An abundance of parking increases sprawl and automobile use, while reducing the viability of alternative modes including public transit (Litman & Doherty, 2011, 4). Because, according to Shoup, “parking is free for 99% of automobile trips in the United States,” public transportation is given a huge cost disadvantage (Shoup, 1999, 549). In their article “Parking, People, and Cities,” Manville and Shoup argue that parking lots “alienate pedestrians and sterilize street life” (242). Although parking has a huge effect on city life and travel patterns, as of 2005, Manville and Shoup stated that “no city we are aware of keeps careful track of its total number of parking spaces” (240). Over ten years later, that is still the case with few exceptions. Because of this lack, I turned to the regression model provided in the article “Estimating parking lot footprints in the Upper Great Lakes Region of the USA" (Davis, Pijanowski, Robinson, & Kidwell, 2010). The model calculates the amount of parking in a zip code based on the amount of urbanized land in that zip code. While the regression model was developed only using data from the Great Lakes region, it serves as an approximation for the entire USA.

Other attributes of the built environment that affect public transportation include land use mix, also known as diversity, network connectivity and design, and walkability. Land use mix is a measure of the extent that different land use types are located near each other. According to Litman, high land use mix "allows more walking and cycling trips … can reduce commute distances," and increases the likelihood of residents using alternative modes (Litman & Steele, 2016, 20). Network connectivity describes the "directness of travel between destinations", often measured by intersection density, and well-connected networks experience reduced automobile travel along with increased public transportation usage (Litman & Steele, 2016, 21-2). Network design describes the actual design of the roadway in terms such as design speed and lane number. Traffic calming designs that slow speeds and narrow streets "tend to reduce automobile traffic and encourage use of alternative modes" (Litman & Steele, 2016, 23).

Walkability describes how easily pedestrians in an area can access important destinations, such as grocery stores and doctor's offices. According to Litman, improvements in walking conditions usually "increase transit travel and reduce automobile travel" (Litman & Steele, 2016, 24). It is important to be able to measure an area's walkability, but many of the more detailed and sophisticated methods require "specialized expertise…can be time-intensive," and require GIS datasets that can be expensive to obtain (Duncan et al, 2011, 4161). Fortunately, simplistic measures of an area's walkability can be obtained freely and easily through WalkScore, which calculates "walkability based on distance to various categories of amenities," and then "produces a score of 0 to 100" (Duncan et al, 2011, 4161). The usefulness of WalkScore for research has been validated by many different researchers (Duncan et al, 2011; Gilderbloom, Riggs, & Meares 2015). However, WalkScore does have problems that should be noted. WalkScore uses Google Maps and user-contributions to locate and classify amenities, which may lead to potential flaws in the data. WalkScore also neglects to include several important qualities of pedestrian-friendliness such as the speed of nearby vehicles, the width and ADA compliance of the sidewalks, and the perceived safety of the area (Gilderbloom, Riggs, & Meares, 2015, 16). Despite these flaws, WalkScore is the best tool currently available to quantify walkability on a large scale.

2.2 Socioeconomic Factors

In addition to attributes of the built environment, socioeconomic factors and transit system characteristics are also believed to affect ridership. The majority of the United States' transit ridership comes from "minorities and low-income households" (Socioeconomics of Urban Travel: Evidence from the 2001 NHTS, 49). Car ownership is another important factor; transit ridership decreases dramatically when vehicle ownership rises.

2.3 Transit System Characteristics

The impact of transit system characteristics on ridership is less clear, but still there. The infrastructure present at a transit station or stop plays a key part in the quality of service experienced by a rider. If people are unable to easily reach the station or stop, whether by walking, biking, or driving, they are unlikely to choose transit. Additionally, riders must be able to access their destinations from a stop. This accessibility can be increased by the presence of bike share stations. Transit riders also place importance on the safety, security, and cleanliness of stations (Eboli & Mazzulla, 2013). In their article, "When the Wait seems Longer: Assessing Quality Attributes of Bus Stops," Woldeamanuel and Somers explain the importance of lighting, benches, and shelters at bus stops. They claim that sufficient lighting "allows for persons such as public transit users to have a sense of safety and comfort as they walk towards a bus stop or any other destination" (Woldeamanuel & Somers, 2016, 4). Benches provide places for passengers to rest, decreasing their perceived waiting time (Woldeamanuel & Somers, 2016, 4). They also argue that shelters are the most important bus stop amenity because they provide protection from the elements and also increase riders' sense of security; therefore shelters factor in "with higher ridership" (Woldeamanuel & Somers, 2016, 4). In addition to improving the quality of service experienced by riders, these amenities increase ridership and "the attractiveness of transit as an alternative form of transportation" (Woldeamanuel & Somers, 2016, 4).

Another transit system characteristic that affects ridership is the presence of real time information. With the proliferation of smartphones, transit agencies are beginning to display information such as arrival times, schedules, and unexpected delays to riders through web browsers, smartphone applications, and text message alerts. These real time information systems "can help passengers adapt to the unreliability of transit" by allowing them to time their departure for a transit stop to minimize their wait (Brakewood, Macfarlane, & Watkins, 2015, 60). Real time information systems also reduce perceived wait time, increase perceived security, and increase passenger satisfaction (Brakewood, Macfarlane, & Watkins, 2015, 60).

2.4 Summary

This literature review furthered my understanding of the variables that influenced transit ridership. The data collection effort described in the following section was guided by the literature review.

Chapter 3

METHODOLOGY

3.1 Data Gathering

Using data from the 2010 United States Census, I selected the top 100 most populated Urban Areas in the United States, as shown in Figure 1, for my analysis. However, several Urban Areas were either merged together or discarded due to a lack of data, so only 96 Urban Areas were actually included in the final analysis. The urban areas of Provo-Orem, Salt Lake City-West Valley City, and Ogden-Layton, Utah were merged into one large urban area for the purposes of this study because all three areas are served by the same transit agency, the Utah Transit Authority. The urban areas of Murrieta-Temecula-Menifee, California and San Juan, Puerto Rico were discarded because of missing data.

Location of the Urban Areas in the Study

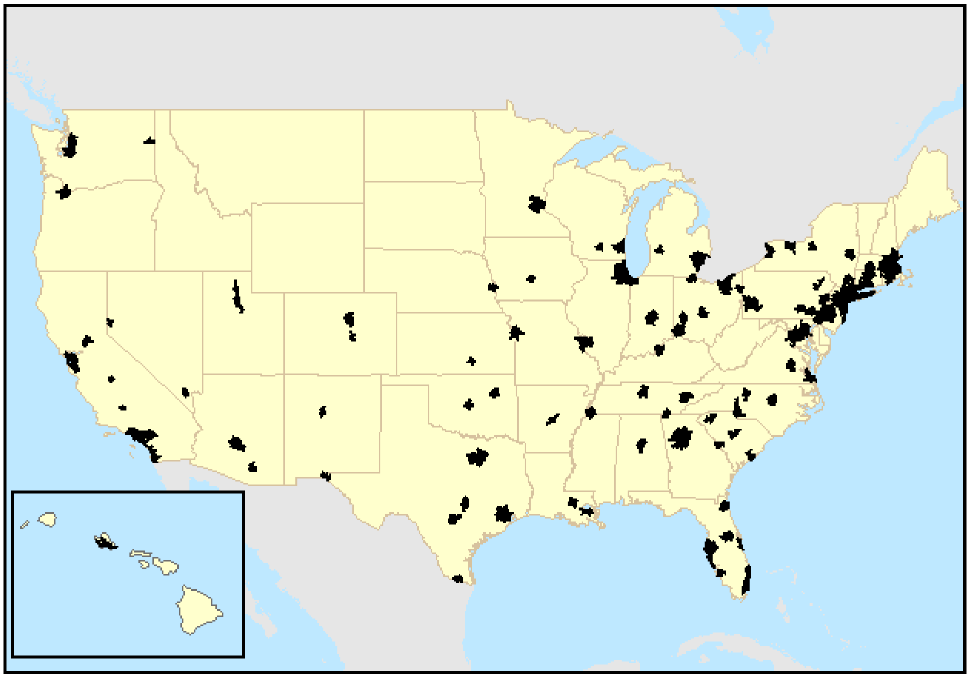

Figure 1: Location of the Urban Areas in the Study

I gathered three types of data described as environment variables, input variables, and output variables. The environment variables describe the physical and socioeconomic characteristics of the Urban Area. They included measures of the area's density, diversity, and design – variables that the literature claims affect transit usage but are not under the direct control of any transit agency. The input variables describe the characteristics of the public transportation system in the area, including the average fare, type of infrastructure available at the transit stations, and the provision of real time information. These variables are within the control of the transit agency. The output variables measure the transit ridership in an urban area by mode.

The data came from many sources, including the US Census, the National Transit Database, the American Public Transportation Association, the EPA's Smart Location Database, WalkScore, the National Land Cover Database, and individual city websites. See Appendix A for a detailed list of the sources of each variable.

Transit ridership in each city was measured by two variables – annual passenger miles and annual unlinked trips. Every transit agency is required to report this data to the National Transit Database. If there was more than one transit agency per mode in an urban area, their data was combined.

The following subsections describe variables that were derived from other measures, specifically land use mix score, data from the smart location database, and parking coverage.

3.1.1 Land Use Data and Mix Calculation

Diversity, or Land Use Mix, is more difficult to measure than the other built environment characteristics because it is characterized by many attributes that are spatially distributed over an urban area. I began by obtaining land use or zoning shapefiles from as many cities as possible. While the land use data was often more detailed than the zoning data, many cities only provided the latter so when both data types were available I chose to use the zoning data. Because land use data for entire Urban Areas was rarely available, I restricted my analysis of the Land Use Mix to the center city of the area. As many cities did not provide this data in a free and/or easily accessible manner, I was only able to obtain data for 51 of the 96 cities.

After gathering this data, I categorized each land use or zoning designation into one of five categories: Residential, Commercial/Institutional, Industrial, Leisure, and Open Space. Each of these land uses generates different travel patterns. Because each city had its own way of labeling their data, I had to make many judgement calls when deciding which of my categories a specific land use fell into. For example, areas labeled as "mixed use residential" were placed in the Residential category because I felt that the residential aspect contributes more to the areas travel patterns than the other uses. Figure 2 illustrates this type of data for Cincinnati.

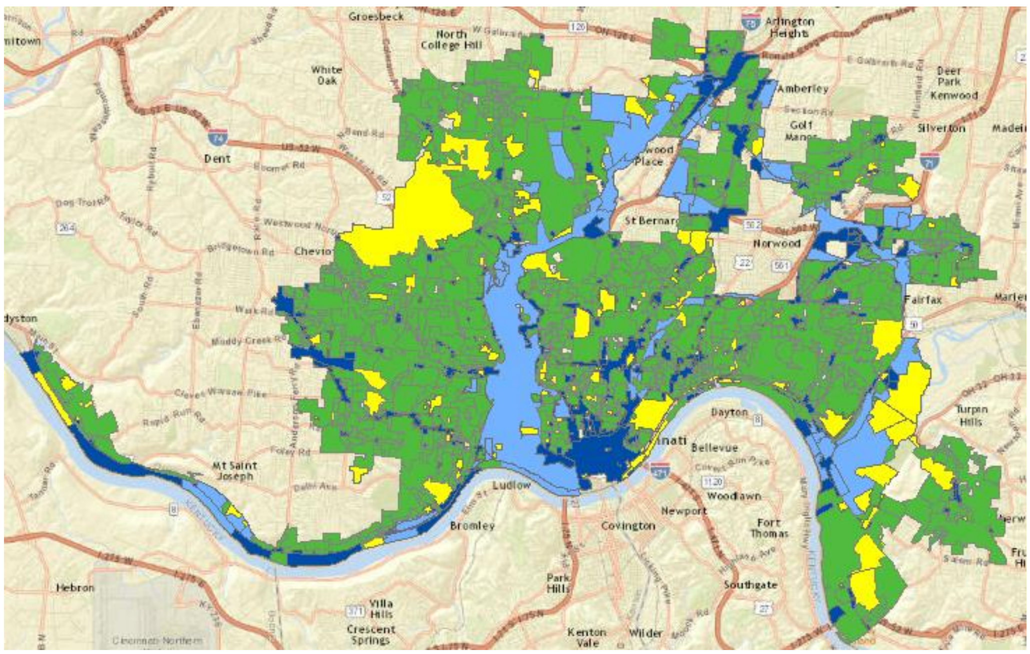

Figure 2: Categorization of Cincinnati's Land Uses (Green = Residential, Dark Blue = Commercial, Light Blue = Industrial, Yellow = Leisure)

The Leisure category was particularly difficult to assign as many cities did not specifically zone for parks, cemeteries, and other leisure areas. Because of this difficulty, I decided to measure the Land Use Mix of each city with two methods. The first would keep Leisure as a separate category and the second would combine the Leisure category with the Commercial/Institutional category.

Next, I calculated the Land Use Mix according to the Entropy method described by Manaugh and Kreider (2013). After calculating the percent each land use category occupied per census tract and the number of land uses present in each census tract, I obtained the entropy index of the tract with the following equation.

$$\text{Entropy Index} = \frac{-\sum A_{ij} \ln A_{ij}}{\ln N_j} \quad \text{where}$$

A_{ij} = Percent of land use i in census tract j
N_j = Number of represented land uses in census tract j

Because the land use data did not match perfectly with the census tract boundaries, there were many census tracts that barely had any land use data in them. To fix this problem, I removed all census tracts that had less than 3 percent of their area filled by the land use data. The final entropy score for each city was the average of all of the census tracts remaining in the city, weighted by the census tract's population. For example, Cincinnati received a score of 0.62, compared to the minimum score of 0.41 in Salt Lake City and the maximum score of 0.79 in Youngstown.

3.1.2 Explanation of Variables from the Smart Location Database

A lot of useful data describing the environment of every census block group in the USA in terms of density, land use diversity, design, destination accessibility, and distance to transit is available from the EPA's Smart Location Database. Created by Ramsey and Bell, this database is a repository of information by block group (Ramsey & Bell, 2014).

To obtain measures for an entire urban area, I averaged the data from each block group, weighting it by the population of that block group. Because the employment density was weighted by census tract population, it effectively measures the employment density only in mixed use areas. This technique allowed me to capture the characteristics of the typical neighborhood in the urban area rather than losing all of the database's detail to aggregation.

While most of the variables I obtained from the Smart Location Database are self-explanatory, a few require further explanation. The first such variable is the *Employment and Housing Entropy Score*, which measures the same characteristic, Diversity, as my two Land Use Mix variables. Instead of using land use data directly, Ramsey and Bell approximate the land use diversity by using "housing unit counts and job counts broken down by employment sector" (Ramsey & Bell, 2014, 17). Although this method is only an approximation of Land Use Mix, it returned results for each of the 96 cities.

Another variable, the *Household Workers per Job Equilibrium Index*, indicates the balance of resident workers to jobs in a census block group. After I aggregated and standardized this variable, a high index score indicated that a city was more balanced, while a low index score indicated that a city was less balanced. The *multimodal network density* variable is the weighted average of the multimodal links per square mile in the census block groups of an Urban Area. Ramsey and Bell define multimodal links as streets or arterials with speeds between 20 and 55 miles per hour where there is both automobile and pedestrian traffic allowed (road with more than three lanes in each direction are excluded from this category) (Ramsey & Bell, 2014, 21). Finally, the *intersection density* variable measures only the intersections that contribute to connectivity – auto-oriented intersections are ignored as they are a barrier to pedestrian mobility, and three way intersections are penalized as they do not provide as much connectivity as four way intersections.

3.1.3 Parking Regression Calculation

In order to estimate the parking coverage of each Urban Area, I used the regression model presented in “Estimating parking lot footprints in the Upper Great Lakes Region of the USA” (Davis, Pijanowski, Robinson, & Kidwell, 2010). After I obtained the 2011 National Land Cover Dataset, I used the ArcGIS *Tabulate Area* tool to count the number of cells of each land cover class that occurred in each zip code in my study area. Three of these land cover classes (codes 22, 23, and 24) represented urbanized land. By summing the area covered by these classes, I obtained the zip code’s URB number (the square meters of urbanized land within the zip code). Using the regression equation developed by Davis (shown below), I estimated the parking lot area in each zip code.

$$Parking\ lot\ area = \left(9.61 * 10^{-6}\mathrm{URB} - 1.08 * 10^{-9}\mathrm{URB}^{\frac{3}{2}} + 14.98\right)^{3}$$

Any zip codes that were only partially inside the study area were excluded from aggregation. Then, the total area of parking in each Urban Area was normalized by the total area of the zip codes in each Urban Area to obtain the percent parking coverage for each city.

While Davis’s model was only developed using data from the Upper Great Lakes Region, it is the best parking estimation model available. Additionally, as the region contains many different scales of cities, it can be applied to the country as a whole.

3.2 Elimination of Variables

At the end of the data gathering process, I had 82 different variables (see Appendix A for sources and descriptions and Appendix B for summary statistics). I then began the process of eliminating highly correlated variables and those variables with too much missing data. Due to a lack of data, I decided to eliminate the ferry and demand responsive modes from my analysis, leaving just bus and rail. Then I calculated the correlations of the remaining data; only two variables were highly correlated, the *average number of attended stations* and the *average number of stations with restrooms*. I decided to remove the restroom variable because it was less descriptive than the attended variable. I then standardized the remaining data and reran the correlations. This time, there were many highly correlated variables and they were eliminated (see Appendix C).

Missing data was imputed using the mean if the variable had data in at least 36 cities. Otherwise, the variable was completely eliminated. Because of this, most of the rail input variables were removed. Almost all of the variables obtained from the APTA fare database were also eliminated. Although these variables had data for over 36 cities for the bus mode, they had almost none for the other modes, so all of the variables were removed. Additionally, ridership for modes that were not reported by an agency to the NTD was assumed to be 0.

3.3 Factor Analysis

Using MATLAB, I conducted a factor analysis to determine what latent factors were present between the remaining variables. During a factor analysis, the variables are analyzed for similar response patterns that may be caused by an underlying factor. Each variable is given a factor loading that describes its relationship to that factor

(Rahn, 2016). Because I believed that the factors would be correlated, I performed an oblique rotation, promax, instead of the default orthogonal rotation (Osborne, 2015).

I did three different versions of the factor analysis – one with just the environment variables, one with just the input variables, and one with both sets of variables. The last version was to determine if there were any significant relationships between the environment and input factors; the analysis revealed that there were none. Each version of the factor analysis was iterated several times to determine the appropriate number of factors. The iteration was stopped when the differences between the factors were no longer interpretable. I ended up with 6 environment factors and 5 input factors.

As some of the variables did not contribute significantly (i.e. with a loading of 0.5 or greater) to any of the factors, I removed them from the final factor analysis. I chose 0.5 as my cutoff value based on my review of the literature. Table 1 lists these variables. Variable descriptions are included in Appendix A. The variables marked with asterisks were still thought to be important, however, so they were still included in the later cluster analysis.

Table 1: Non-Significant Variables

Environment	Input
• CBSA Area • CBSA Employment Density • Center City Area • Center City Employment Density • Choice Riders* • Percent of Households Receiving Food Stamps • Percent of Households where no on over 14 speaks English • Weighted Average Jobs/Households of CBG in Urban Area • Total Road Network Density • Pedestrian Oriented Network Density • Land Use Mix (Leisure merged with Commercial)*	• Average Adult Base Bus Fare* • Average Number of Bus Stations with Concessions • Average Number of Bus Stations with lockers or secure bike facility • Average Number of Bus Stations with Bike Racks • Average Number of All-day Parking Spaces per Bus Station • Average Number of Part-day Parking Spaces per Bus Station • Average Number of Bicycle Rack Spaces per Bus Station • Phone Access to Bus Real Time Information System • Average Adult Rail Base Fare*

The loadings from the final factor analysis were used to compute the factor scores for the individual cities. The factors and their variable loadings are listed below in the results section.

3.4 Cluster Analysis

Based on the cities' scores for Factors A-K and the variables I decided to include (Choice Riders, Land Use Mix, Average Adult Bus and Rail Fares) were put into clusters. A cluster analysis is a statistical technique and discovery tool that reveals relationships within large amounts of data by dividing cases into clusters so that the cases within a cluster are similar to each other and different from the other clusters. To do this, I used Orange, an open source data visualization and analysis tool, to calculate the Euclidean distances between the cities and input that data into the Hierarchical Clustering program (Demsar et al, 2013). The program I used is shown in Figure 3. The average linkage method was used so that each city would be in a cluster where it was similar to all other cities in the cluster, not just to one other (Lab 13 – Cluster Analysis, n.d.). This meant that the cities with similar characteristics would be grouped together.

I did several different versions of the cluster analysis with different subsets of the data. One version only analyzed the environmental factors and variables, while another only the input factors and variables. However, the main cluster analysis analyzed all of the data: all of the factors, the choice ridership, the land use mix, the fares, and the ridership (bus, rail, and total) in terms of annual unlinked trips.

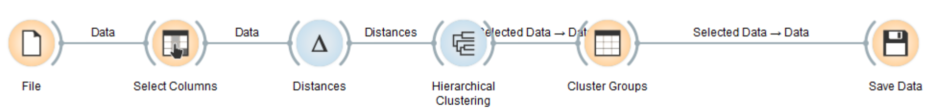

Figure 3: Orange Program for Cluster Analysis

3.5 Regression

Finally, I performed a regression analysis on the data. I did this on three subsets of the data: one with all of the factors and selected variables, one on just the factors and selected variables relevant to bus transit, and one on just the factors and selected variables relevant to rail ridership. For the first regression analysis, the *standardized annual unlinked passenger trips - all modes* was the dependent variable. The independent variables were Factors A-K, choice ridership, land use mix score, average adult bus fare, and average adult rail fare. For the second regression analysis, the dependent variable was the *standardized annual unlinked passenger bus trips* and the independent variables were Factors A-F and H-K, choice ridership, land use mix score, and average adult bus fare. For the third regression analysis, the dependent variable was the *standardized annual unlinked passenger rail trips* and the independent variables were Factors A-G, choice ridership, land use mix score, and average adult rail fare. While I would have liked to analyze each cluster, they were all too small to return meaningful results.

Chapter 4

RESULTS

4.1 Factor Analysis Results

The results of the factor analysis are documented in Appendix D showing the factor scores for each city. The final factor analysis resulted in the identification of 11 underlying factors (six environmental and five input factors). Table 2 identifies these factors by giving their descriptive name, component variables, and variable weights. Only variables with weights greater than 0.5 were included, which is consistent with the practices reported in the literature. The six environmental factors, A through F, and five input factors, G through K, are each discussed in the following subsections.

Table 1: Factors and the Weights of their Component Variables

Factor Name	Component Variables	Variable Weight
Factor A: Area Attributes	Population of Urban Area	0.95
	Land Area	1.06
	CBSA Employment	0.58
	Center City Employment	0.83
	Weighted Average Jobs within a 45 Minute Auto Commute	0.66
Factor B: Need	Percent of Population Below Poverty Level	0.71
	Percent of Population with a Disability	1.02
	Median Household Income	-0.66
	Percent of Population 65 or Older	0.55
Factor C: Density and Walkability	Percent of Households without a Vehicle	0.63
	Weighted Average Residential Density of CBG in Urban Area	0.80
	Weighted Average Employment Density CBG in Urban Area	0.77
	WalkScore	0.45
Factor D: Land Use Diversity	Weighted Average of Employment and Housing Entropy Score	1.00
	Weighted Average of Worker/Jobs Equilibrium Index	0.83
	Multimodal Network Density	0.51
Factor E: Automobile Infrastructure Characteristics	Urban Area Population Density	0.86
	Weighted Average Jobs within a 45 Minute Auto Commute	0.46
	Weighted Average Intersection Density	0.69
	Estimated Percent Parking Coverage	0.74

Factor Name	Component Variables	Variable Weight
Factor F: Education	Percent of Population with a High School Degree or Higher	0.91
	Percent of Population with a Bachelor's Degree or Higher	0.82
	Land Use Mix – Weighted Average Entropy Score	-0.47
Factor G: Rail Real Time Information	Real Time Information System	0.73
	Phone Access	0.63
	Text/SMS Access	1.01
	Website Access	0.86
	Smartphone App Access	0.75
Factor H: Bus Real Time Information	Real Time Information System	0.68
	Text/SMS Access	0.87
	Website Access	0.80
	Smartphone App Access	0.94
Factor I: Extra Bus Station Infrastructure	Average Number of Bus Stations with Bikeshare	1.00
	Average Number of Bus Stations with Emergency Call Buttons	0.78
	Average Number of Bike Spaces in Secure Facilities per Bus Station	0.97
Factor J: Bus Stop Infrastructure	Average Number of Bus Stops with Shelters	0.69
	Average Number of Bus Stops with Benches	0.71
	Average Number of Bus Stops with Lighting	0.92
Factor K: Basic Bus Station Infrastructure	Average Number of ADA Accessible Bus Stations	0.65
	Average Number of Attended Bus Stations	0.69
	Average Number of Bus Stations with Security Cameras	0.77

4.1.1 Factor A: Area Attributes

Factor A is a measure of the demographic and spatial attributes of an urban area. The physical size of the area is the variable with the highest weight, followed closely by the urban area's population.

The other three components of Factor A relate to employment. The weights of these three variables vary inversely with the area being considered, with the center city employment having the highest and the CBSA employment having the lowest weight. This suggests that concentrated areas of employment in a city contribute more to the area's character than the total number of jobs spread over the urbanized region.

With these variables, it is no surprise that New York receives the highest score for this factor. While New York is an outlier with a score of 25 (the next highest is Los Angeles at 13, followed by Chicago at 9), there are a number of cities clustered around the low end of the spectrum (23 cities have scores between -2.0 and -2.7). This suggests that the majority of US cities are similar to each other and very different from New York and other massive metropolitan areas in terms of their area characteristics.

4.1.2 Factor B: Need

Factor B measures the need of the population in an urban area. The variables that contribute positively to this factor indicate the percentages of the population with a disability (1.02), the percentage living below the poverty level (0.71), and the percentage of the population of age 65 or older (0.55). The low weight given to the elderly population in calculating need suggests that a significant number of elderly are better off than the poor and/or disabled. The variable indicating the median household income in the area is weighted negatively (-0.66), which makes intuitive sense as need decreases with wealth.

The city with the highest need score, 12, is Sarasota. Sarasota has the largest percent of population under the poverty level, the largest percentage with a disability, and the largest percentage at or over age 65. It also has the lowest household income of the cities analyzed. While Sarasota is a clear outlier, the other cities with high need scores are much closer together, with Youngstown, Madison, and Cape Coral all scoring between 3 and 4.

Salt Lake City has the smallest need score at -7.8. The Salt Lake City area is a combination of the three official Urban Areas in Utah, which all have relatively low need scores. Although none of the individual cities had the lowest score, when combined the little need in the area was diluted and caused the extremely low score. The cities with the next lowest need, Washington DC, Mission Viejo, and Concord, all scored about -4.7.

4.1.3 Factor C: Density and Walkability

Factor C is composed of density and walkability related variables. This makes sense as dense areas are typically more walkable. The residential and employment density have nearly equal weights, confirming the importance of jobs/housing balance. The percent of households without a vehicle is more heavily weighted than the WalkScore variable. This suggests that the former is a better measure of walkability than the latter, which makes sense when considering WalkScore’s flaws.

Not surprisingly, New York City is again an outlier with the extremely high score of 16. San Francisco trails it with a score of 5. As with Factor A, there were no outliers on the negative end of the spectrum. Generally, the cities with a high score in Factor A also have high scores in Factor C. However there are several exceptions, including Atlanta, Houston and Dallas, where the cities are large and sprawling.

Additionally, some small cities have similar Factor C scores to those of much larger cities. For example, Honolulu's score of 4.8 is comparable to the scores of San Francisco and Boston (5.4 and 4.3) which both scored significantly higher in area attributes.

4.1.4 Factor D: Land Use Diversity

Factor D measures the land use diversity of an urban area. While both diversity measures obtained from the Smart Location Database are included in this factor, the Land Use Mix variables I calculated are not. This could be due to the incompleteness of the Land Use Mix data compared to the Smart Location data. The high weight given to the Employment and Housing Entropy Score is interesting because it suggests that this score is better at capturing city characteristics than the Worker/Job Equilibrium Index. While Jobs/Housing balance is frequently used as an approximate measure of land use diversity, the entropy method used in the Smart Location Database is less common although it seems to be a better measure.

The multimodal network density also contributes to this factor although it has a lower weight. This variable appears to be more of a measure of walkability and connectivity than of diversity. However, its inclusion in this factor may indicate the importance of connectivity in creating a diverse mix of land uses.

Factor D does not have any outliers. The highest scoring city is Boston (6.3), followed closely by Springfield (5.6) and Lancaster (5.1). The lowest scoring city is Las Vegas (-4.2), followed by Honolulu (-4.0) and Reno (-3.5).

4.1.5 Factor E: Automobile Infrastructure Characteristics

Factor E is composed mostly of variables that relate to the automobile infrastructure of the urban area. The only variable that does not fit in this category is the Urban Area Population Density which has a weight of 0.86. I believe this variable is included in this factor because of the way that the parking coverage was estimated. The regression model developed by Davis calculated the parking in a zip code based on the area of urbanized land in the zip code. The urban area population density would be highest in the areas designated as urbanized by the National Land Cover Database.

While the Estimated Percent Parking Coverage and the Jobs within a 45 minute Auto Commute are obvious linked to increased emphasis on the automobile, the intersection density is less clear. As discussed earlier, the Smart Location Database intended the intersection density variable to represent the pedestrian connectivity of the area. Auto-oriented intersections were even ignored in the variable's calculation. However, intersections between multimodal links were still included, and those links are still traveled by automobiles albeit at slower speeds. This means that intersection density is still a measure, although a weaker one, of automobile infrastructure.

Los Angeles, not surprisingly, has the highest score (10.8) in the automobile infrastructure factor. It is followed by San Francisco (7.3) and New York City (6.0). Poughkeepsie has the lowest score (-3.5) followed by Winston (-3.2).

4.1.6 Factor F: Education

Factor F consists mostly of variables measuring education. The variable measuring the percent of the population with at least a high school degree has a weight of 0.91, while the variable measuring the population with at least a bachelor's degree is 0.82. However, the Land Use Mix has a negative weight on this factor. This may be because areas that have more integrated land uses tend to be in poorer inner city areas. Educated people tend to live in richer, more suburban areas.

The cities with the highest Education scores are Washington DC (3.1), McAllen (2.8), and Concord (2.8). Salt Lake City has the lowest Education score by far (-12.3). The next lowest are Madison (-3.8) and Bakersfield (-3.0).

Salt Lake City's low score is mostly due to its high Land Use Mix, not its education performance as 93 percent of its residents have at least a high school degree and 35 percent have at least a bachelor's degree. However, Madison's residents are least likely to have at least a high school degree (at 63 percent) and Bakersfield's residents are least likely to have at least a bachelor's degree (at 17.5%) so these cities scores are mostly due to education performance.

4.1.7 Factor G: Rail Real Time Information

Factor G represents the availability of real time information (RTI) on the rail transit systems in an urban area. Interestingly, the availability of RTI through texting, a website, or a smartphone application is more heavily weighted (1.01, 0.86, and 0.75 respectively) than the mere presence of a RTI system (0.73). This suggests that transit agencies need to provide convenient access to RTI in addition to providing the information itself. Because phone access has the lowest weight, agencies should avoid this access method and instead choose one of the others.

Washington DC, Boston, and Portland have the highest rail RTI scores (5.6). Sixteen cities have no rail RTI at all. New York and Chicago have the lowest scores of the cities that do have some RTI (-3.18 and -3.01 respectively). Unfortunately, 59 of the 96 cities did not report if they had or did not have rail RTI.

4.1.8 Factor H: Bus Real Time Information

Factor H represents the availability of RTI on the bus transit systems in an urban area. As with Factor G, the presence of RTI is contributes less weight than the methods by which the information is accessed. Smartphone access has the most weight, followed by texting and website access. Phone access is not even significant.

Seven cities tie for the highest bus RTI score (5.3): Boston, Austin, San Antonio, Salt Lake City, Syracuse, Milwaukee, and Chicago. Thirty five cities have no bus RTI at all; 13 of these cities also had no rail RTI and the other 22 cities did not report their rail RTI status. Of the cities that had bus RTI, Miami, San Francisco, and Harford had the lowest scores (all about -2.2). Twenty seven cities did not report their bus RTI status.

4.1.9 Factor I: Extra Bus Station Infrastructure

Factor I represents the presence of extra features of bus station infrastructure in an urban area. These extra features include the presence of bikeshare stations, emergency call buttons, and secure bicycle facilities. The bikeshare and secure bicycle facilities variables are the most highly weighted. This makes sense as these features increase the accessibility of the bus station, a quality that is highly valued by riders. The emergency call button variable is also highly weighted, though not as high, because increases the safety and security of the riders.

Boston has a very high score in Factor I (19.4), distantly followed by Minneapolis (5.2) and Pittsburgh (1.4). Boston’s high score is explained by the fact that its bus stations all have emergency call buttons and bikeshare stations, while one third also have secure bus facilities. Fifty one cities did not provide any extra bus station infrastructure. Of the cities that did provide some, San Francisco and Los

Angeles scored lowest at -0.5. Thirty two cities did not report the status of their extra bus station infrastructure.

4.1.10 Factor J: Bus Stop Infrastructure

Factor J indicates the presence of bus stop infrastructure (shelters, benches, lighting) at a typical bus stop in an urban area. While Woldeamanuel and Somers argued that shelters were the most important bus stop amenity, the shelter variable has the lowest weight and the lighting variable has the highest.

Minneapolis has the highest bus stop infrastructure score at 8.9, followed by Jacksonville at 7.2 and Seattle at 4.7. Birmingham has the lowest score at -1.4. Thirty two cities did not report the status of their bus stop infrastructure.

4.1.11 Factor K: Basic Bus Station Infrastructure

Factor K represents the presence of basic traditional bus station features, including ADA accessibility, staff attendance, and security cameras. Bus stations are more likely to have these features than those in Factor I. The security camera variable has the highest weight, confirming the importance of security to riders.

Six cities tie for the highest score for this factor: Oklahoma City, Boston, Winston-Salem, Louisville/Jefferson County, Little Rock, and Allentown all have a score of 2.6. Eight cities have none of these features in any bus station. Of the cities that have these features, Chicago and Los Angeles score the lowest (-2.9 and -2.6 respectively). Thirty two cities did not report the status of their basic bus station infrastructure.

4.2 Cluster Analysis Results

The detailed results of the cluster analysis are shown in Appendix E. The cluster analysis separated the cities into 11 distinct clusters containing multiple cities and twelve clusters of one (see Figure 4). The single city clusters reflect that the cities are too different from any of the other cities to be grouped. The unique cities include Boston, Chattanooga, Chicago, Jacksonville, Los Angeles, Minneapolis, New York City, Salt Lake City, San Francisco, Sarasota, Urban Honolulu, and Washington DC. These cities are unique because of each city's size, population, and transit system. Boston, Chicago, Los Angeles, New York City, San Francisco, and Washington DC have the largest transit systems in country but each city is different. Urban Honolulu's uniqueness could be explained by it being an island with a large tourist economy. Also, Urban Honolulu was missing some of the data that was provided for the mainland cities which may have caused its differences. Further research is required to determine why Chattanooga, Jacksonville, Minneapolis, Salt Lake City, and Sarasota are unique.

The 11 distinct clusters, shown in Figure 4 and 5, were each given a descriptive name to reflect their common characteristics:

- Modern Large Cities
- Bus-Dependent Midsize Cities
- Auto-Dependent Newer Cities
- Non-Transit Focused Cities
- Auto Focused Cities with a Rail Transit Culture
- Older Cities with Established Transit
- Changing Large Cities
- Cities with High Need and Basic Transit Service

- Cities with High Need and Choice Riders
- Cities with Potential
- Midsize Cities with Little Transit Culture

Each cluster is described in the following subsections.

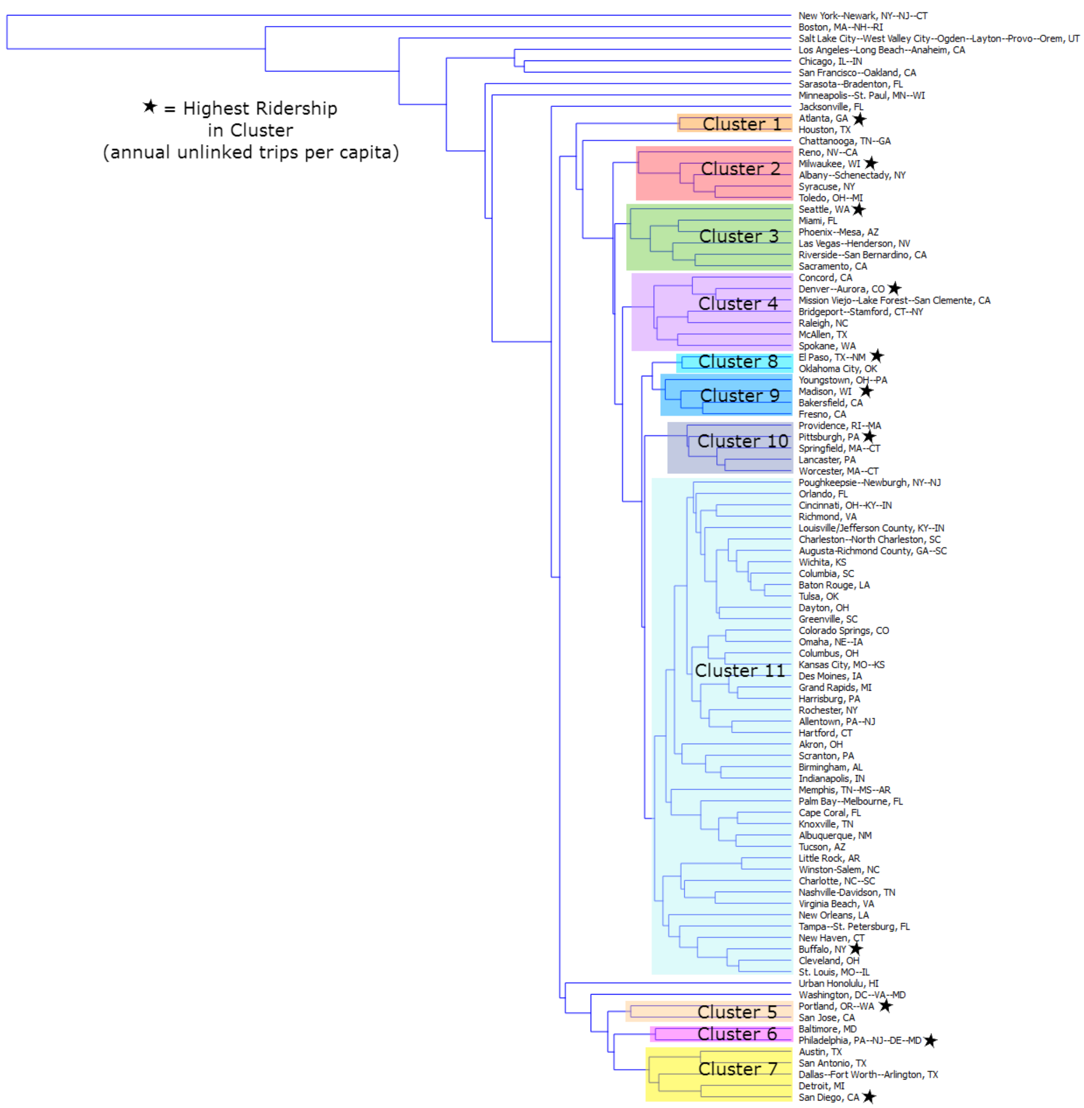

Figure 4: Dendogram Showing Cluster Analysis Results

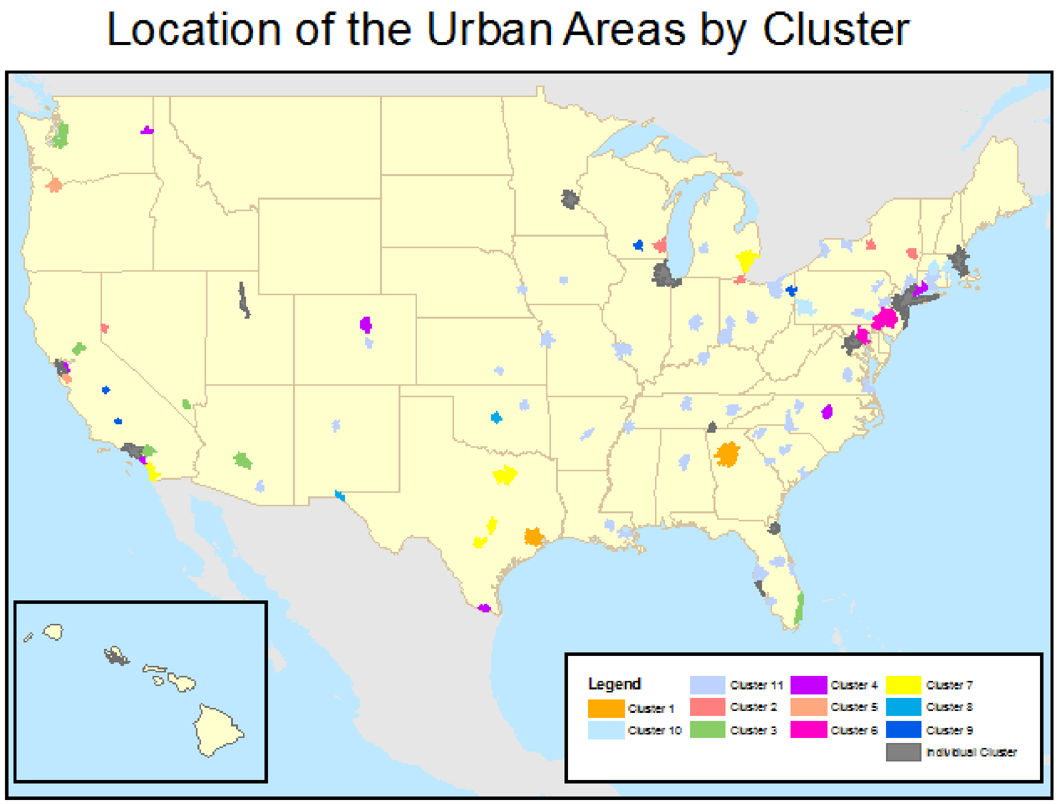

Figure 5: Location of Urban Areas by Cluster

4.2.1 Cluster 1: Modern Large Cities

This cluster contains two cities, Atlanta and Houston. Because these cities both experienced a huge population boom during the past few decades, this cluster was named “Modern Large Cities.” Both cities have a high *area attribute* factor score, indicating that they are large in population, employment, and physical size. For the *need* factor, the cities scored around -2; this means that the cities have below average need but are still worse off than cities like Salt Lake City. While both cities have about 16 percent of their population below the poverty level (about the national average), they experience below average proportions of disabled and elderly populations and above average median household income. This helps to explain the above average proportion of choice riders in both cities. In terms of the *density/walkability* factor and the *land use diversity* factor, the cities score below average.

Neither city has a real time information system for bus or rail transit. They also lack extra bus station infrastructure, including bikeshare, bike lockers, and emergency call buttons. The bus stop and basic bus station infrastructure in both cities is below average, with Atlanta being worse than Houston. Each city has about 69,000,000 annual unlinked passenger trips by bus.

Atlanta has more trips per capita than Houston, even though Atlanta scores worse in *density/walkability* and *bus station/stop infrastructure* than Houston. However, Atlanta has six times as much rail transit as Houston has. The people of these cities may prefer rail to bus – this may explain why Atlanta has more ridership despite its inferior bus system. If this is true, Houston could improve its ridership by improving its rail system.

4.2.2 Cluster 2: Bus-Dependent Midsize Cities

This four city cluster contains Reno, Milwaukee, Albany, Syracuse, and Toledo. Their *area attribute* scores are all around -2 (except Milwaukee which is -0.5), which means that they are smaller than the average city in this study. These cities exhibit a diverse environmental landscape. These cities do not have any rail transit. While their bus infrastructure scores vary, they provide real time information for their bus systems. Except for Milwaukee and Toledo, the cities experience bus ridership of about 10 million unlinked trips per year.

Milwaukee has much more ridership than the others – its annual unlinked bus trips are about 127 million and there are 93 annual unlinked trips per capita. There are several differences between Milwaukee and the other three cities that could explain this discrepancy. Milwaukee is about twice as large as the others in both population and physical size. It has more center city employment and jobs with an automobile commute than the other cities, which indicates that it has a strong central city. In terms of residential, employment, urban area population, and intersection density, Milwaukee scores highest for the cluster.

In this cluster, Toledo has the lowest ridership per capita (6) and the lowest overall ridership, 3 million annual unlinked trips. Toledo is different from the other cities in several ways. In the cluster, Toledo has the highest percentage of population below the poverty level, the highest percentage of population with a disability, and the least median household income. Toledo has the lowest residential density, with only 2.63 households per acre, and the lowest employment density, with only 1.65 jobs per acre. Toledo has the smallest percentage of bachelors' degrees of these four cities. In terms of bus stop infrastructure, it has the least shelters per stop and has no stops with benches or lighting. Toledo also has the lowest base bus fare, which may cause it to have less money to spend on amenities.

4.2.3 Cluster 3: Auto-Dependent Newer Cities

This cluster contains Seattle, Miami, Phoenix, Las Vegas, Riverside, and Sacramento. These cities all have negative scores for the *land use diversity* factor, indicating that they typically have segregated land uses. Las Vegas has the lowest *land use diversity* score of the group, at -4.2. Sacramento has the highest of the group, at -0.9. The cities in this cluster have *automobile infrastructure characteristics* scores of around 2, indicating that they have similar systems that are above average. However, they do not experience the outlier status of NYC and San Francisco. Three of the cities in this cluster (Riverside, Sacramento, and Seattle) have high percentages of choice riders (all are around 85%). The other half of the cluster is closer to the average percentage. The majority of transit in these cities is done by bus, though there is some rail.

Within this cluster, Seattle has the best ridership with 65 annual unlinked trips per capita and Riverside has the worst with only 13 annual unlinked trips per capita. The other cities in the cluster have a ridership around 30 unlinked trips per capita.

Seattle is different from the other cities in several ways. It has the highest percent of population with a disability, percent of population without a vehicle, percent of population with a high school degree, employment density, and intersection density. It is the only city in the cluster with an above average score for multimodal network density. Although Seattle has the cheapest bus fare (the average base adult bus fare is only $0.47), it has good station and stop infrastructure characteristics. A few of its bus stations have emergency call buttons, half of its bus stops have shelters, and more than two thirds have lighting.

Riverside has the most CBSA employment by far – it is about three times the national average while the other cities in the cluster are mostly below average. At the same time, Riverside has the lowest center city employment at about a third of the national average. This implies that the city lacks a strong core. Additionally, compared to the other cities in the cluster, more of Riverside's population is below the poverty level and less of the population has a disability, is elderly, or is educated. Only 5.5 percent of Riverside households lack a vehicle, the lowest percentage in the cluster. Riverside is the only city in the cluster with a residential density below the national average. Riverside has the largest benches-to-stop ratio, which suggests that benches are not an important bus stop feature for these cities. Finally, Riverside is the only city in the cluster that completely lacks rail transit.

4.2.4 Cluster 4: Non-Transit Focused Cities

This cluster contains Concord, Denver, Mission Viejo, Bridgeport, Raleigh, McAllen, and Spokane.

These cities all have a *need* score between -1 and -5. They are relatively well off and the populace does not need transit as much as in other cities. Unsurprisingly, these cities all had high scores in education (around 2). Except in Raleigh, where choice ridership was below average, the cities in this cluster had choice ridership scores around 2, indicating their ridership patterns were close to those exhibited in Cluster 3. Although I was unable to get Land Use Mix data for 4 of the cities in the cluster, the three I did get scored below average.

For the most part, transit agencies in these cities did not report the status of their infrastructure or real time information systems. When they did report, the score was typically very low. Except for Denver, which had a little rail, none of these cities reported any rail transit and these cities all had fewer than 12 million annual unlinked bus trips.

Denver had the highest ridership, with 42 annual unlinked trips per capita. Of the cities in this cluster, Denver was the only one with an urban area population above the national average. It also had the most center city employment and the highest employment density. In terms of education, Denver was the worst in the cluster with only 89.2 percent of the population with at least a high school degree and only 40.4 percent with at least a bachelor’s degree.

Mission Viejo had the worst ridership at 1.2 annual unlinked trips per capita. Similar to Riverside in cluster 3, Mission Viejo had the highest CBSA employment, the lowest center city employment, and the smallest percentage of households without a vehicle (3.8 percent).

4.2.5 Cluster 5: Auto Focused Cities with a Rail Transit Culture

This cluster contains Portland and San Jose. Both cities experience below average need, above average density/walkability, high scores in automobile infrastructure, and above average education. Both cities have high scores in their provision of rail real time information. Their bus station infrastructure scores below average in both basic and extra features. They have above average choice ridership, but generally score lower in that than the cities in Clusters 3 and 4.

Portland has a ridership of 60 annual unlinked trips per capita, which is more than twice as San Jose's 26 trips per capita. In terms of employment, Portland has slightly more center city jobs than San Jose and much less CBSA employment. Portland scores higher in the *need* factor than San Jose. Despite having a lower residential density, Portland has more vehicle-less households, a higher WalkScore, and a higher multimodal network density. It also has a slightly lower parking coverage. While Portland has a lower score for the *bus stop infrastructure* factor, it's bus stations are better than San Jose's. Portland has a higher bus fare than San Jose but a lower rail fare. Portland has a much larger rail system than San Jose. By imitating Portland's rail system, San Jose could increase its overall transit ridership.

4.2.6 Cluster 6: Older Cities with Established Transit

This cluster contains Baltimore and Philadelphia. Because these cities were established well before the age of private automobile, they are have a higher land use mix score and are more walkable and dense than the average city. This may explain why both cities have a below average parking coverage. These cities have above average scores for *area attributes* and below average scores for *need*. These cities provide real time information for both their rail and bus transit systems. However, the features of their bus stations and stops score below average.

While the ridership in both cities is far above the national average of 25 annual unlinked trips per capita, Philadelphia's ridership of 70 annual unlinked trips per capita is significantly higher than Baltimore's ridership of 48 annual unlinked trips per capita. Philadelphia is more than twice the size and population of Baltimore. Philadelphia has much more employment within its CBSA (251 million compared to 65 million) and nearly twice as much center city employment. Philadelphia has poorer, more disabled, and more elderly people than Baltimore. Philadelphia has higher residential and employment densities, has more vehicle-less households, and has a higher WalkScore. Philadelphia has a higher ridership because its built environment and demographics are more suited to transit than Baltimore's.

4.2.7 Cluster 7: Changing Large Cities

This cluster contains Austin, San Antonio, Dallas, Detroit, and San Diego. These cities have above average area attribute scores, indicating that they have more population, employment, and/or land area than the average American city. In terms of land use diversity, these cities are at or slightly below the national average. They provide better rail and bus real time information than the cities in cluster 6. The basic bus station infrastructure in each city scores above average (except in Austin which did not report bus station data). Generally, each city also provided above average bus stop infrastructure (the exceptions being Austin and Detroit).

San Diego has the highest ridership in the cluster (40 annual unlinked trips per capita) and Detroit has the lowest (12 trips per capita). While San Diego has the lowest CBSA employment of the cluster, its population, residential, and employment density are the highest in the cluster. Compared to the other cities in the cluster, San Diego has the largest proportion of benches per bus stop, the highest percentage of choice riders, the cheapest bus fare, and the highest rail ridership.

While Detroit has the most elderly, vehicle-less households, and the largest WalkScore, San Diego is second in all of these variables. This suggests that these variables do not have a strong influence on ridership in this cluster. Detroit has the smallest employment density and is the only city in the cluster with below average center city employment. Detroit is the only city in the cluster that does not provide website access to its real time information system. Also, Detroit has the smallest proportion of shelters and benches to bus stops in the cluster. These variables tend to increase with ridership in the cluster so they may have some explanatory power.

To improve ridership, Detroit should increase employment density, especially in the center city, and improve its bus stop infrastructure. Demographically, the city is very similar to San Diego, so by emulating that city, its ridership should improve.

4.2.8 Cluster 8: Cities with High Need and Basic Transit Service

This cluster contains El Paso and Oklahoma City, which are below average in *area attributes*, *density/walkability*, and *education* but are above average in *need*. Both cities lack a rail system and only provide basic bus transit. They lack real time information, emergency call buttons, bikeshare stations, and bike lockers. However, they do have above average scores in bus stop and basic bus station infrastructure. In both cities, the choice ridership is about 63% which is only slightly below the national average.

El Paso has a higher ridership than Oklahoma City (16 annual trips per capita vs 3 annual trips per capita). While El Paso has lower employment in both the CBSA and the center city, the ratio between center city and CBSA employment is 0.4 in El Paso and 0.06 in Oklahoma City. El Paso residents are much more likely to live below the poverty level and have a lower income than Oklahoma City residents. They are also less likely to have a high school or bachelor's degree. El Paso has a higher employment and residential density, a higher percent of households without vehicles, and a higher urban area population density. El Paso has a dramatically different street network than Oklahoma City. For every type of road, including multimodal and pedestrian orientated links, its network density is much lower, and correspondingly El Paso has about half the intersection density of Oklahoma City. Despite this, El Paso has a slightly higher WalkScore, perhaps due to its higher residential and employment densities. While every bus stop in Oklahoma City has a shelter, only 17 percent of El Paso's stops have one. Also, stops in Oklahoma City are more likely to have lighting than in El Paso. However, over half of El Paso's stops have benches and none of Oklahoma City's stops do. While it is possible that Oklahoma City's transit ridership could be improved by adding benches to its bus stops, El Paso's higher ridership is almost entirely explained by the higher need of its population and the higher density of the city.

4.2.9 Cluster 9: Midsize Cities with High Need and Choice Riders

This cluster includes Bakersfield, Fresno, Madison, and Youngstown. They are all similar in size to each other and to the cities in Cluster 2, with an *area attribute* score of around -2. This cluster is most similar to Cluster 8, but it has several differences. In addition to being smaller, the Cluster 9 cities generally have higher *need* (except for Bakersfield), lower *education*, lower *bus stop infrastructure,* lower *basic bus station infrastructure*, higher *choice riders*, and higher *land use mix* scores.

Madison has the highest ridership in the cluster with 37 annual unlinked trips per capita and Youngstown has the lowest with only 0.5 annual unlinked trips per capita. Madison differs from the other cities in the cluster in several ways. It has the largest percentage of the population living below the poverty level, highest residential and employment density, highest WalkScore, lowest percentage of the population with at least a high school degree, and the lowest land use mix score. Madison is the only city in the cluster that provides a real time information system. It also has security cameras in all of its bus stations. Despite its poverty, Madison has the largest percentage of choice riders, 73 percent, and the most expensive base bus fare, 2 dollars.

Compared to the rest of the cluster, Youngstown has the highest CBSA employment, the lowest center city employment, and the lowest jobs within a 45 minute auto commute. This indicates that the jobs in Youngstown tend to be far away from both the city core and the residential areas. However, the city has a *Worker/Jobs Equilibrium Index* score of 0.34 and a land use mix score of 0.79, the highest in the cluster, which seems to indicate the opposite. Youngstown also has the lowest percentage of residents living below the poverty level in the cluster. However, residents of the city are more likely to be disabled and/or elderly than in the other cities in the cluster. Youngstown has the cheapest base bus fare, 1.25 dollars. Its ridership may improve if that fare is raised so that it can provide the station and stop amenities that are needed by its elderly and disabled residents.

4.2.10 Cluster 10: Cities with Potential

This cluster includes Providence, Pittsburgh, Springfield, Lancaster, and Worchester. Except for Pittsburgh, these cities are below average in *area attributes* but generally score higher than the cities in Clusters 2 and 9. These cities have high scores in *land use diversity* and below average scores in *automobile infrastructure*.

Pittsburgh has the highest ridership in this cluster with 37 annual trips per capita and Lancaster has the lowest with 5 annual trips per capita. Pittsburgh is the largest city in the cluster, by both population and size. It has the highest percent of vehicle-less households, the highest employment density, and the highest percentage of high school and college graduates. Pittsburgh is the only city in the cluster that provides real time information though it is only accessible by website. It has better bus station and stop infrastructure than the others in the cluster. It is also the only city in the cluster with a rail transit system.

Lancaster is the smallest city in the cluster by population and size. It has the fewest vehicle-less households, the lowest employment density, the lowest percentage of choice riders, and the lowest intersection density. However, it has the highest WalkScore at 80. This suggests that walkability is not a major factor in determining transit ridership for these cities. To improve ridership Lancaster should improve its transit infrastructure and increase employment density.

4.2.11 Cluster 11: Cities with Little Transit Culture

This cluster contains 44 cities. It includes Poughkeepsie, Orlando, Cincinnati, Richmond, Louisville/Jefferson County, Charleston, Augusta-Richmond County, Wichita, Columbia, Baton Rouge, Tulsa, Dayton, Greenville, Colorado Springs, Omaha, Columbus, Kansas City, Des Moines, Grand Rapids, Harrisburg, Rochester, Allentown, Hartford, Akron, Scranton, Birmingham, Indianapolis, Memphis, Palm Bay, Cape Coral, Knoxville, Albuquerque, Tucson, Little Rock, Winston-Salem, Charlotte, Nashville-Davidson, Virginia Beach, New Orleans, Tampa, Buffalo, Cleveland, and St. Louis. Most of these cities have below average *area attributes*, *density/walkability*, and *choice riders.* Almost all of the cities lack a real time information system and any extra bus station infrastructure. Most have a bus stop infrastructure score that is below the national average. Only 15 of the 44 cities have a rail transit system.

In this cluster, Buffalo has the highest ridership with 31 annual trips per capita and Augusta-Richmond County has the lowest with 2 annual trips per capita. Compared to the rest of the cluster, Buffalo has the largest percent of vehicle-less households, a high residential density, a high multimodal network density, and a low percentage of choice riders. Augusta-Richmond County has a low residential and employment density, a high percentage of choice riders, and the lowest intersection density of the cluster. To increase ridership, the cities in Cluster 11 should focus on increasing density and walkability.

4.3 Regression Results

The results of the regression analyses are reported in Appendix F. As shown in Table 8, the regression on the entire dataset returned, with an R square value of 0.87, five factors with a significant p-value (less than 0.05): Factors A (Area Attributes), C (Density/Walkability), D (Land Use Diversity), E (Automobile Infrastructure Characteristics), and H (Bus Real Time Information). None of these factors had a very strong coefficient, the highest being Factor C at 0.23. Factors A and C had positive coefficients, while Factors D, E, and H had negative ones. These results mean that cities that were larger and more dense and walkable had higher ridership. Ridership was lower in cities with higher scores in Land Use Diversity, Automobile Infrastructure Characteristics, and Bus Real Time Information.

As shown in Table 9, the regression on the bus-relevant subset of the data returned, with an R square value of 0.92, two factors with a significant p-value: Factors A and C. Their coefficients were still positive but were smaller than they were for the entire dataset.

As shown in Table 10, the regression on the rail-relevant subset of the data returned, with an R square value of 0.80, four factors with a significant p-value: Factors A, C, D, and E. As in the first regression, A and C were positive while D and E were negative.

Only a few significant factors were identified at national scale, which supports the idea that the factors vary in relative importance between clusters. Unfortunately, the clusters were all too small to successfully conduct a regression analysis. Future research is necessary to confirm this theory.

It is also surprising the Land Use Diversity negatively affects ridership in both the entire dataset and rail-related dataset regression analyses. This runs counter to the literature and the results from the cluster analysis. Further research is required.

Chapter 5

CONCLUSION

My initial hypothesis was proven wrong, as it made the assumption that all cities with high ridership would share the same built environment, socioeconomics, and transit system characteristics. The research has shown that the relationships are much more nuanced as the importance of various socioeconomic, built environment and transit system characteristics varies between different types of city. There is no one size fits all method for improving transit ridership. Instead of trying to copy New York City's model, cities should emulate the most successful member of their cluster. For example, in Modern Large Cities, like Atlanta and Houston, ridership increased with the quality of rail transit while in Bus-Dependent Midsize Cities, like Albany and Toledo, ridership depends on the provision of basic bus stop amenities and the presence of a strong central city. However, in nearly every cluster the cities with the strongest city core, as represented by employment and population density, had the highest ridership per capita.

In the future, this research could be improved in several ways. Adding more variables into the analysis could shed light on areas that have been overlooked. The large amount of missing data was a major problem, as it prevented me from analyzing nuanced fare data and rail station data. By obtaining accurate and complete data the analysis would be greatly improved. Another avenue for future research is to apply the methodology to a smaller scale and analyze the neighborhoods inside one large city. This would allow more detail to be captured in the analysis.

REFERENCES

Brakewood, C., Macfarlane, G., & Watkins, K. (2015). The impact of real-time information on bus ridership in New York City. *Transportation Research Part C* (53), 59-75.

Cervero, R. (1998). *The transit metropolis: a global inquiry*. Washington, D.C.: Island Press.

Cervero, R., Kockelman, K. (1997). Travel Demand and The 3Ds: Density, Diversity, and Design. *Transportation Research Part D: Transport and Environment* (Volume 2, Issue 3), p. 199-219.

Davis, A., Pijanowski, B., Robinson, K., & Kidwell, P. (2010). Estimating parking lot footprints in the Upper Great Lakes Region of the USA. *Landscape and Urban Planning*, 68-77.

Demsar J, Curk T, Erjavec A, Gorup C, Hocevar T, Milutinovic M, Mozina M, Polajnar M, Toplak M, Staric A, Stajdohar M, Umek L, Zagar L, Zbontar J, Zitnik M, Zupan B (2013) Orange: Data Mining Toolbox in Python. *Journal of Machine Learning Research* 14(Aug):2349−2353.

Duncan, D., Aldstadt, J., Whalen, J., Melly, S., & Gortmaker, S. (2011). Validation of Walk Score® for Estimating Neighborhood Walkability: An Analysis of Four US Metropolitan Areas. *International Journal of Environmental Research and Public Health* (8), 4160-4179.

Dickens, M. (2015). 2014 Public Transportation Fare Database. Washington, DC: American Public Transportation Association.

Dickens, M. (2015). 2014 Public Transportation Infrastructure Database. Washington, DC: American Public Transportation Association.

Eboli, L., & Mazzulla, G. (2013). A Multicriteria Approach for Analyzing Railway Service Quality. *Transportation Research Board 92nd Annual Meeting*.

Ewing, R., Pendall, R., & Chen, D. (2003). Measuring Sprawl and Its Transportation Impacts. *Transportation Research Record* (1831), 175-183.

Frank, L., & Engelke, P. (2001). *How Land Use and Transportation Systems Impact Public Health: A Literature Review of the Relationship Between Physical Activity and Built Form.* Retrieved from http://www.cdc.gov/nccdphp/dnpa/pdf/aces-workingpaper1.pdf

Gilderbloom, J., Riggs, W., & Meares, W. (2015). Does walkability matter? An examination of walkability's impact on housing values, foreclosures and crime. Cities (42), 13-24.

Lab 13 — Cluster Analysis. (n.d.). Retrieved from: http://ecology.msu.montana.edu/labdsv/R/labs/lab13/lab13.html

Litman, T., & Steele, R. (2016). *Land Use Impacts on Transport*. Retrieved from Victoria Transport Policy Institute: http://www.vtpi.org/landtravel.pdf

Litman, T., & Doherty, E. (2011). *Transportation Cost and Benefit Analysis II – Parking Costs*. Retrieved from Victoria Transport Policy Institute: http://www.vtpi.org/tca/tca0504.pdf

Maat, K., & Harts, J. J. (2001). Implications of Urban Development for Travel Demand in the Netherlands. *Transportation Research Record* (1780), 9-16.

Manaugh, K., & Kreider, T. (2013). What is mixed use? Presenting an interaction method for measuring land use mix. *Journal of Transport and Land Use*, 6(1), 63-72. doi:10.5198/jtlu.v6i1.291

Manville, M., & Shoup, D. (2005). Parking, People, and Cities. *Journal of Urban Planning and Development*, 233-245.

Osborne, J. (2015). What is Rotating in Exploratory Factor Analysis? *Practical Assessment, Research & Evaluation* (20), 1-7.

Pucher, J., & Renne, J. (2003). Socioeconomics of Urban Travel: Evidence from the 2001 NHTS. *Transportation Quarterly* (53), 49-77.

Rahn, M. (n.d.) Factor Analysis: A Short Introduction, Part 1. Retrieved from: http://www.theanalysisfactor.com/factor-analysis-1-introduction/

Ramsey, K., & Bell, A. (2014). Smart Location Database. Retrieved from: https://www.epa.gov/smartgrowth/smart-location-database-technical-documentation-and-user-guide

Shoup, D. (1999). The trouble with minimum parking requirements. *Transportation Research Part A*, 549-574.

U.S. Census Bureau. (n.d.). American Community Survey. Retrieved from American Fact Finder: http://factfinder2.census.gov

Woldeamanuel, M., & Somers, A. (2016). When the Wait seems Longer: Assessing Quality Attributes of Bus Stops. *Transportation Research Board 2016 Annual Meeting*.

Appendix A

COMPLETE LIST OF DATA SOURCES AND DATA DESCRIPTION

Table 3: List of Data Sources and Descriptions

Variable	Source	Description
Population of Urban Area (2010)	Census Urban Area List Layout	Number of people living in the Urban Area
Housing Unit Count for Urban Area (2010)	Census Urban Area List Layout	Number of houses in the Urban Area
Land Area (square meters)	Census Urban Area List Layout	Square meters of land within the Urban Area
Land Area (square miles)	Census Urban Area List Layout	Square miles of land within the Urban Area
Urban Area Population Density (pop/sqmi)	Census Urban Area List Layout	Population per square mile in the urban area
CBSA Employment	Smart Location Database	Employment in the CBSA (from Census LEHD, 2010)
CBSA Population	Smart Location Database	Population in the CBSA (from Census, 2010)
CBSA Workers	Smart Location Database	Number of workers living in the CBSA (from Census LEHD, 2010)
CBSA Area (sqmi)	On The Map	Square miles of the Core Based Statistical Area that the Urban Area is in
CBSA Employment Density (jobs/sqmi)	Derived from other variables	Total Employment per square mile of CBSA
Urban Area Employment	Smart Location Database	Total Employment of all Census Block Groups in the Urban Area (Census 2010, InfoUSA 2011 for MA)
Center City Employment (2013)	On The Map	The employment for the center city of the urban area, as identified by On The Map (categorized as a Place)
Center City Area (sqmi)	On The Map	The area of the center city of the urban area, as identified by On The Map (categorized as a Place)
Center City Employment Density	Derived from other variables	Center city Employment per Area
Choice Riders	Census Data, B08141	Percentage of those who use transit to get to work that have at least one vehicle available
Percent of Households without a vehicle	Census Data, B08201,	Percent of Households without a vehicle in the Urban Area

Variable	Source	Description
Percent of Population below Poverty Level	Census Data, S1701	Percent of Population below Poverty Level in the Urban Area
Percent of Households receiving food stamps	Census Data, S2201	Percent of Households receiving food stamps in the Urban Area
Percent of Households where no one over 14 speaks English	Census Data, S1602	Percent of Households where no one over 14 speaks English in the Urban Area
Percent of Population with a disability	Census Data, S1810	Percent of Population with a disability in the Urban Area
Mean Household Income	Census Data, S1901	Mean Household Income in the Urban Area
Median Household Income	Census Data, S1901	Median Household Income in the Urban Area
Percent of Population 65 or older	Census Data, S0103	Percent of Population 65 or older in the Urban Area
Percent with High School Degree or Higher	Census Data, S1501	Percent with High School Degree or Higher in the Urban Area
Percent with Bachelor's Degree or Higher	Census Data, S1501	Percent with Bachelor's Degree or Higher in the Urban Area
Weighted Average Residential Density of CBG in Urban Area (HH/acre)	Smart Location Database	The average residential density (households/acre on unprotected land) of census block groups in the urban area, weighted by population of the census block group
Weighted Average Population Density of CBG in Urban Area (ppl/acre)	Smart Location Database	The average population density (people/acre on unprotected land) of census block groups in the urban area, weighted by population of the census block group
Weighted Average Employment Density of CBG in Urban Area (jobs/acre)	Smart Location Database	The average Employment density (jobs/acre on unprotected land) of census block groups in the urban area, weighted by population of the census block group
Weighted Avg Jobs/HH of CBG in UZA	Smart Location Database	The average Jobs per Household of census block groups in the urban area, weighted by population of the census block group
Weighted Avg of Employment and Housing Entropy Score	Smart Location Database	The average Employment and Housing entropy of census block groups in the urban area, weighted by population of the census block group.
Weighted Avg of Worker/Jobs Equilibrium Index	Smart Location Database	The average Household Workers per Job Equilibrium Index score of census block groups in the urban area, weighted by population of the census block group.
# of workers earning $1250/month or less (work location), 2010	Smart Location Database	# of workers earning $1250/month or less (work location) in all Census Block Groups in the Urban Area
# of workers earning $1250/month or less (home location), 2010	Smart Location Database	# of workers earning $1250/month or less (home location) in all Census Block Groups in the Urban Area
Weighted Avg Jobs within 45 min auto commute	Smart Location Database	The average number of jobs within a 45 minute auto commute from a census block group in the urban area, weighted by population of the census block group
Weighted Avg Jobs within 45 min transit commute	Smart Location Database	The average number of jobs within a 45 minute transit commute from a census block group in the urban area, weighted by population of the census block group

Variable	Source	Description
Proportion of Unprotected Land within 1/2 mile of a fixed guideway transit stop	Smart Location Database	Acres of Unprotected Land within 1/2 mile buffer of a fixed guideway transit stop out of total arces of unprotected land. As unprotected land is where housing and employment is, this "value roughly approximates the proportion of the CBG's activity (housing units and employment) that are proximate to rapid transit"
Proportion of Unprotected Land within 1/4 mile of a fixed guideway transit stop	Smart Location Database	Acres of Unprotected Land within 1/4 mile buffer of a fixed guideway transit stop out of total arces of unprotected land. As unprotected land is where housing and employment is, this "value roughly approximates the proportion of the CBG's activity (housing units and employment) that are proximate to rapid transit"
Weighted Avg Distance of Population Centroid of CBG to nearest transit stop	Smart Location Database	Average Distance to the Nearest Transit Stop from the population centroid of census block groups in the Urban Area, weighted by the population of the Census Block Group
Total road network density	Smart Location Database	Road network density of entire Urban Area
Auto Oriented Network Density	Smart Location Database	Network density of Urban Area in terms of facility miles of auto-oriented links per square mile
Multi-modal Network Density	Smart Location Database	Network density of Urban Area in terms of facility miles of multi-modal links per square mile
Ped Oriented Network Density	Smart Location Database	Network density of Urban Area in terms of facility miles of ped-oriented links per square mile
Weighted Avg Intersection Density	Smart Location Database	Average Intersection Density of CBGs in Urban Area, weighted by population of CBG. The intersection density was calculated by a "weighted sum of component intersection metrics" designed so the measure would reflect the street connectivity as experienced by pedestrians and bicyclists.
Estimated Percent Parking Coverage (Parking Area/Total Area in UZA)	National Land Cover Database, Davis et al	This value was calculated based on the work done by Davis et al. in "Estimating parking lot footprints in the Upper Great Lakes Region of the USA". Davis developed a regression equation to model the parking lot area in the Great Lakes Region, based on data from the National Land Cover Database. I applied this formula to the entire USA, to get the parking area for each urban area, then normalized the data by dividing by the total urban area.

Variable	Source	Description
Land Use Mix - Weighted Average Entropy Score (including Leisure Category	Land Use or Zoning Shapefiles from individual city websites	To calculate the land use mix, I first gathered all of the land use and zoning shapefiles that I could find for the cities I was analyzing. I only collected this data for the main city in the urban area, as it was difficult to find data that included the entire urban area. While I believed that land use data was more accurate and detailed, most cities only provided zoning data. For the cities that provided both, I analyzed the zoning data for consistency. I categorized the data into residential, commercial/institutional, leisure, and industrial uses. I calculated entropy scores for each census tract in the urban area that I had data for. These scores were then averaged, weighted by population of the census tract.
Land Use Mix - Weighted Average Entropy Score (Leisure merged with Commercial Category)	Land Use or Zoning Shapefiles from individual city websites	Same as above, except the leisure uses were merged in with the commercial/instituational uses, so only there were only three categories of land use. This was done to account for the lack of consistency and quantity in the lesiure category
WalkScore	WalkScore	The WalkScore of the main city in the urban area.
Annual Unlinked Passenger Trips	April 2015 Raw Data NTD	The annual total number of passengers who boarded public transit as reported by the transit agenices in the urban areas. (Individuals are counted each time they board a vehicle) (Data is from Fiscal Year 2013)
Annual Passenger Miles Traveled	April 2015 Raw Data NTD	The annual cumulative sum of the distances ridden by each passenger. Small systems did not have to report this which is why there are some zeros for miles traveled even when unlinked trips is non-zero. This means the true value of passenger miles traveled may be higher. (Data is from Fiscal Year 2013)
Average Adult Base Fare	2014-APTA-Fare-Database	Average Adult Base Fare (weighted by unlinked trips of transit agency)

Variable	Source	Description
Transfer Surcharge Score	2014-APTA-Fare-Database	The average Transfer Surcharge score represents the average of the scores (0 = no, 1 = yes) for each transit agency in the urban area, weighted by the unlinked trips of the transit agency. A transfer surcharge is an extra fare required to transfer to anothre vehicle or mode within the same transit agency.
Zone Surcharge Score	2014-APTA-Fare-Database	The average Zone Surcharge score represents the average of the scores (0 = no, 1 = yes) for each transit agency in the urban area, weighted by the unlinked trips of the transit agency. A zone surcharge is an extra fare based on miles traveled.
Peak-Hour Surcharge Score	2014-APTA-Fare-Database	The average Peak-Hour Surcharge score represents the average of the scores (0 = no, 1 = yes) for each transit agency in the urban area, weighted by the unlinked trips of the transit agency. A peak hour surcharge is an extra fare for trips during peak hours.
Cash Surcharge Score	2014-APTA-Fare-Database	The average cash surcharge score represents the average of the cash surcharge scores (0 = no cash surcharge, 1 = cash surcharge) for each transit agency in the urban area, weighted by the unlinked trips of the transit agency. A cash surcharge is an extra fare for using cash or a non-prefered fare media.
Multi-agency regional pass score	2014-APTA-Fare-Database	The average Multi-agency Regional Pass score represents the average of the scores (0 = no, 1 = yes) for each transit agency in the urban area, weighted by the unlinked trips of the transit agency. A regional pass is a pass designed for use on all transit agencies in a region.

Variable	Source	Description
Stored Value Media Score	2014-APTA-Fare-Database	The average Stored Value Media score represents the average of the scores (0 = no, 1 = yes) for each transit agency in the urban area, weighted by the unlinked trips of the transit agency. If an agency has stored value media, that means that it has a deal with another agency for a shared stored value media such as a farecard.
Regional Smart Card Score	2014-APTA-Fare-Database	The average Regional Smart Card score represents the average of the scores (0 = no, 1 = yes) for each transit agency in the urban area, weighted by the unlinked trips of the transit agency. A regional smart card is a contactless card used by multiple transit agencies in a region.
SmartCard Use Score	2014-APTA-Fare-Database	The average SmartCard Use score represents the average of the scores (0 = no, 1 = yes) for each transit agency in the urban area, weighted by the unlinked trips of the transit agency. According to the APTA, a smart card is "a single piece of material without a magnetic strip but with a small computer chip good for one or more trips that is usually not surrendered but altered by machine removal of some or all of the stored value as each trip is taken."
Guaranteed Ride Home Plan Score	2014-APTA-Fare-Database	The average Guaranteed Ride Home plan score represents the average of the scores (0 = no, 1 = yes) for each transit agency in the urban area, weighted by the unlinked trips of the transit agency
Average Number of ADA Accessible Stations per Total Stations	2014-APTA-infrastructure-Database	Average Number of ADA Accessible Stations per Total Stations (Weighted by unlinked trips of transit agency)

Variable	Source	Description
Average Number of Attended stations per Total Stations	2014-APTA-infrastructure-Database	Average Number of Attended stations (stations with a person on duty to answer question and monitor equipment) per Total Stations (weighted by unlinked trips of transit agency)
Average Number of Stations with Security Cameras per Total Stations	2014-APTA-infrastructure-Database	Average Number of Stations with Security Cameras per Total Stations (weighted by unlinked trips of transit agency)
Average Number of stations with emergency call buttons per total stations	2014-APTA-infrastructure-Database	Average Number of stations with emergency call buttons per total stations (weighted by unlinked trips of transit agency)
Average Number of Stations with Concessions per Total Stations	2014-APTA-infrastructure-Database	Average Number of Stations with Concessions per Total Stations (weighted by unlinked trips of transit agency)
Average Number of Stations with Restrooms per Total Stations	2014-APTA-infrastructure-Database	Average Number of Stations with Restrooms per Total Stations (weighted by unlinked trips of transit agency)
Average Number of stations with all-day auto parking per total stations	2014-APTA-infrastructure-Database	Average Number of stations with all-day auto parking per total stations (weighted by unlinked trips of transit agency)

Variable	Source	Description
Average Number of stations with lockers or secure bike facility per Total Station	2014-APTA-infrastructure-Database	Average Number of stations with lockers or secure bike facility per Total Station (weighted by unlinked trips of transit agency)
Average Number of stations with bike racks per total stations	2014-APTA-infrastructure-Database	Average Number of stations with bike racks per total stations (weighted by unlinked trips of transit agency)
Average Number of stations with Bikeshare station per total stations	2014-APTA-infrastructure-Database	Average Number of stations with Bikeshare station per total stations (weighted by unlinked trips of transit agency)
Average Number of Bus/Rail Stops with Shelters per total Stops	2014-APTA-infrastructure-Database	Average Number of Bus/Rail Stops with Shelters per total Stops (weighted by unlinked trips agency)
Average Number of Bus/Rail Stops with Benches per total stops	2014-APTA-infrastructure-Database	Average Number of Bus/Rail Stops with Benches per total stops (weighted by unlinked trips of transit agency)

Variable	Source	Description
Average Number of Bus/Rail Stops with Lighting per total stops	2014-APTA-infrastructure-Database	Average Number of Bus/Rail Stops with Lighting per total stops (weighted by unlinked trips of transit agency)
Average Number of bus/rail stops with auto lot per total stops	2014-APTA-infrastructure-Database	Average Number of bus/rail stops with auto lot per total stops (weighted by unlinked trips of transit agency)
Average Number of All-day Parking Spaces per Station	2014-APTA-infrastructure-Database	Average Number of All-day Parking Spaces per Station (weighted by unlinked trips of transit agency)
Average Number of Part-day Parking Spaces per Station	2014-APTA-infrastructure-Database	Average Number of Part-day Parking Spaces per Station (weighted by unlinked trips of transit agency)
Average Number of Bicycle Rack Spaces per Station	2014-APTA-infrastructure-Database	Average Number of Bicycle Rack Spaces per Station (weighted by unlinked trips of transit agency)
Average Number of bicycle spaces in secure facilities - locker/facility per Station	2014-APTA-infrastructure-Database	Average Number of bicycle spaces in secure facilities - locker/facility per Station (weighted by unlinked trips of transit agency)

Variable	Source	Description
Real Time Information System Score	2014-APTA-infrastructure-Database	The average Real Time Info System score represents the average of the scores (0 = no, 1 = yes) for each transit agency in the urban area, weighted by the unlinked trips of the transit agency
Phone Access to Real Time Information System Score	2014-APTA-infrastructure-Database	The average Phone Access to Real Time Info System score represents the average of the scores (0 = no, 1 = yes) for each transit agency in the urban area, weighted by the unlinked trips of the transit agency
Smartphone App Access to Real Time Information System Score	2014-APTA-infrastructure-Database	The average Smartphone App Access to Real Time Info System score represents the average of the scores (0 = no, 1 = yes) for each transit agency in the urban area, weighted by the unlinked trips of the transit agency
Text/SMS Access to Real Time Information System Score	2014-APTA-infrastructure-Database	The average Text/SMS Access to Real Time Info System score represents the average of the scores (0 = no, 1 = yes) for each transit agency in the urban area, weighted by the unlinked trips of the transit agency
Website Access to Real Time Information System Score	2014-APTA-infrastructure-Database	The average Website Access to Real Time Info System score represents the average of the scores (0 = no, 1 = yes) for each transit agency in the urban area, weighted by the unlinked trips of the transit agency

Appendix B

COMPLETE LIST OF VARIABLES AND SUMMARY STATISTICS

Table 4: Summary Statistics of Variables

Variable	Mean	Std Dev	Max	Min	Count
Population of Urban Area (2010)	1708683	2443768	18351295	381112	100
Housing Unit Count for Urban Area (2010)	701279	952680	7263095	138258	100
Land Area (square meters)	1535549895	1463810961	8935981360	358555119	100
Land Area (square miles)	593	565	3450	138	100
Urban Area Population Density (pop/sqmi)	2583	1058	6999	1212	100
CBSA Employment	40061861	105402540	894559251	184272	99
CBSA Population	97445044	249144141	2106508731	425417	100
CBSA Workers	39488970	104227226	889748309	179685	99
CBSA Area (sqmi)	5035	3567	27408	837	95
CBSA Employment Density (jobs/sqmi)	215	192	906	22	95
Urban Area Employment	806039	1100700	7976951	91704	99

Variable	Mean	Std Dev	Max	Min	Count
Center City Employment (2013)	298029	441774	3648146	10278	95
Center City Area (sqmi)	162	163	875	6	95
Center City Employment Density	2229.05	1588.16	8970.88	20.65	95
Choice Riders	0.66	0.11	0.97	0.41	100
Percent of Households without a vehicle	0.09	0.04	0.32	0.03	100
Percent of Population below Poverty Level	0.15	0.05	0.41	0.06	100
Percent of Households receiving food stamps	0.13	0.05	0.34	0.02	100
Percent of Households where no one over 14 speaks English	0.05	0.07	0.66	0.01	100
Percent of Population with a disability	0.12	0.02	0.22	0.07	100
Mean Household Income	76685.76	16882.85	141775.00	34248.00	100
Median Household Income	56339.80	12416.82	103068.00	21806.00	100
Percent of Population 65 or older	0.13	0.03	0.30	0.07	99
Percent with High School Degree or Higher	0.88	0.05	0.95	0.63	100
Percent with Bachelor's Degree or Higher	0.33	0.07	0.54	0.18	100
Weighted Average Residential Density of CBG in Urban Area (HH/acre)	3.51	2.81	24.78	1.05	100

Variable	Mean	Std Dev	Max	Min	Count
Weighted Average Population Density of CBG in Urban Area (ppl/acre)	8.33	6.68	58.45	2.48	100
Weighted Average Employment Density of CBG in Urban Area (jobs/acre)	2.70	1.96	15.96	0.73	99
Weighted Avg Jobs/HH of CBG in UZA	6.68	15.81	110.54	0.62	99
Weighted Avg of Employment and Housing Entropy Score	0.47	0.04	0.56	0.37	99
Weighted Avg of Worker/Jobs Equilibrium Index	0.29	0.05	0.48	0.21	99
# of workers earning $1250/month or less (work location), 2010	187397.96	250906.48	1753910.00	12348.00	95
# of workers earning $1250/month or less (home location), 2010	170383.33	236556.32	1721194.00	7320.00	99
Weighted Avg Jobs within 45 min auto commute	99709.95	73109.66	488469.40	23729.56	99
Weighted Avg Jobs within 45 min transit commute	5259.53	6589.70	46585.50	0.75	65
Proportion of Unprotected Land within 1/2 mile of a fixed guideway transit stop	0.02	0.02	0.08	0.00	52
Proportion of Unprotected Land within 1/4 mile of a fixed guideway transit stop	0.01	0.01	0.03	0.00	52
Weighted Avg Distance of Population Centroid of CBG to nearest transit stop	229.25	120.95	466.56	7.71	73
Total road network density	12.71	52.27	522.16	2.11	100
Auto Oriented Network Density	7.33	52.37	517.36	0.17	100
Multi-modal Network Density	1.15	0.53	2.86	0.23	100

Variable		Mean	Std Dev	Max	Min	Count
Ped Oriented Network Density		4.34	1.42	8.38	1.58	100
Weighted Avg Intersection Density		127.55	109.76	688.70	28.99	100
Estimated Percent Parking Coverage (Parking Area/Total Area in UZA)		0.03	0.01	0.04	0.01	98
Land Use Mix - Weighted Average Entropy Score (including Leisure Category		0.56	0.06	0.69	0.42	51
Land Use Mix - Weighted Average Entropy Score (Leisure merged with Commercial Category)		0.62	0.09	0.83	0.42	51
WalkScore		47	17	88	10	99
Annual Unlinked Passenger Trips	Bus	209458615	544962934	4704081975	0	96
	Demand Response	18620095	37959857	280078470	0	95
	Ferry	31794240	55008442	177588951	0	12
	Rail	663934860	2607583177	18187543140	0	49
	Total	565131096	2410813351	23113743835	0	97
Annual Passenger Miles Traveled	Bus	51413488	140628027	1209557773	101446	96
	Demand Response	1161699	1987530	12610014	9494	95
	Ferry	4562992	8285296	23023117	0	12
	Rail	100953624	428401984	3017721361	0	49
	Total	103601622	442538722	4260020903	196386	97
Average Adult Base Fare	Bus	1.78	0.91	5.97	0	89
	Demand Response	2.27	3.21	15	0	84
	Ferry	1.70	2.27	6	0	9
	Rail	2.14	2.30	15	0	44
Transfer Surcharge Score	Bus	0.37	0.47	1	0	50
	Demand Response	0.00	0.00	0	0	1
	Ferry	0.00	0.00	0	0	3
	Rail	0.19	0.38	1	0	29
Zone Surcharge Score	Bus	0.12	0.31	1	0	50
	Demand Response	0.00	0.00	0	0	1

Variable		Mean	Std Dev	Max	Min	Count
	Ferry	0.00	0.00	0	0	3
	Rail	0.35	0.45	1	0	29
Peak-Hour Surcharge Score	Bus	0.02	0.14	1	0	50
	Demand Response	0.00	0.00	0	0	1
	Ferry	0.00	0.00	0	0	3
	Rail	0.14	0.33	1	0	29
Cash Surcharge Score	Bus	0.04	0.20	1	0	50
	Demand Response	0.00	0.00	0	0	1
	Ferry	0.00	0.00	0	0	3
	Rail	0.07	0.25	1	0	29
Multi-agency regional pass score	Bus	0.28	0.44	1	0	50
	Demand Response	0.00	0.00	0	0	1
	Ferry	0.00	0.00	0	0	3
	Rail	0.34	0.47	1	0	29
Stored Value Media Score	Bus	0.17	0.35	1	0	50
	Demand Response	0.00	0.00	0	0	1
	Ferry	0.00	0.00	0	0	3
	Rail	0.33	0.46	1	0	29
Regional Smart Card Score	Bus	0.15	0.35	1	0	50
	Demand Response	0.00	0.00	0	0	1
	Ferry	0.00	0.00	0	0	3
	Rail	0.28	0.45	1	0	29
SmartCard Use Score	Bus	0.32	0.46	1	0	50
	Demand Response	0.00	0.00	0	0	1
	Ferry	0.00	0.00	0	0	3
	Rail	0.43	0.48	1	0	29
Guaranteed Ride Home Plan Score	Bus	0.31	0.45	1	0	50
	Demand Response	1.00	0.00	1	1	1
	Ferry	0.60	0.43	1	0	3
	Rail	0.24	0.43	1	0	29
Average Number of ADA Accessible Stations per Total Stations	Bus	0.81	0.36	1	0	65
	Demand Response	N/A	N/A	0	0	0
	Ferry	0.50	0.50	1	0	4

Variable		Mean	Std Dev	Max	Min	Count
	Rail	0.84	0.29	1	0	29
Average Number of Attended stations per Total Stations	Bus	0.36	0.37	1	0	65
	Demand Response	N/A	N/A	0	0	0
	Ferry	0.49	0.49	1	0	4
	Rail	0.21	0.32	1	0	29
Average Number of Stations with Security Cameras per Total Stations	Bus	0.48	0.41	1	0	65
	Demand Response	N/A	N/A	0	0	0
	Ferry	0.25	0.43	1	0	4
	Rail	0.63	0.41	1	0	29
Average Number of stations with emergency call buttons per total stations	Bus	0.04	0.18	1	0	65
	Demand Response	N/A	N/A	0	0	0
	Ferry	0.00	0.00	0	0	4
	Rail	0.17	0.32	1	0	29
Average Number of Stations with Concessions per Total Stations	Bus	0.19	0.31	1	0	65
	Demand Response	N/A	N/A	0	0	0
	Ferry	0.03	0.04	0.1	0	4
	Rail	0.12	0.22	1	0	29
Average Number of Stations with Restrooms per Total Stations	Bus	0.39	0.41	1	0	65
	Demand Response	N/A	N/A	0	0	0
	Ferry	0.38	0.41	1	0	4
	Rail	0.22	0.33	1	0	29
Average Number of stations with all-day auto parking per total stations	Bus	4.74	36.37	296	0	65
	Demand Response	N/A	N/A	0	0	0
	Ferry	0.04	0.06	0.15	0	4
	Rail	0.77	2.62	15	0	29
Average Number of stations with lockers or secure bike facility per Total Station	Bus	0.05	0.14	0.73	0	65
	Demand Response	N/A	N/A	0	0	0
	Ferry	0.01	0.02	0.05	0	4
	Rail	0.19	0.28	0.85	0	29
Average Number of stations with bike racks per total stations	Bus	0.20	0.33	1	0	65
	Demand Response	N/A	N/A	0	0	0
	Ferry	0.16	0.28	0.65	0	4
	Rail	0.28	0.37	1	0	29

Variable		Mean	Std Dev	Max	Min	Count
Average Number of stations with Bikeshare station per total stations	Bus	0.02	0.13	1	0	65
	Demand Response	N/A	N/A	0	0	0
	Ferry	0.03	0.04	0.1	0	4
	Rail	0.01	0.05	0.2	0	29
Average Number of Bus/Rail Stops with Shelters per total Stops	Bus	0.15	0.21	1	0	65
	Demand Response	N/A	N/A	0	0	0
	Ferry	0.25	0.43	1	0	4
	Rail	0.12	0.29	1	0	29
Average Number of Bus/Rail Stops with Benches per total stops	Bus	0.15	0.21	0.98	0	65
	Demand Response	N/A	N/A	0	0	0
	Ferry	0.25	0.43	1	0	4
	Rail	0.03	0.16	0.85	0	29
Average Number of Bus/Rail Stops with Lighting per total stops	Bus	0.09	0.22	1	0	65
	Demand Response	N/A	N/A	0	0	0
	Ferry	0.25	0.43	1	0	4
	Rail	0.03	0.18	1	0	29
Average Number of bus/rail stops with auto lot per total stops	Bus	0.01	0.11	0.91	0	65
	Demand Response	N/A	N/A	0	0	0
	Ferry	0.00	0.00	0	0	4
	Rail	0.01	0.07	0.38	0	29
Average Number of All-day Parking Spaces per Station	Bus	132.38	182.32	688	0	65
	Demand Response	N/A	N/A	0	0	0
	Ferry	24.55	42.52	98	0	4
	Rail	236.72	215.41	965	0	29
Average Number of Part-day Parking Spaces per Station	Bus	10.59	54.22	330	0	65
	Demand Response	N/A	N/A	0	0	0
	Ferry	24.55	42.52	98	0	4
	Rail	2.16	10.12	56	0	29
Average Number of Bicycle Rack Spaces per Station	Bus	6.00	15.96	123	0	65
	Demand Response	N/A	N/A	0	0	0
	Ferry	0.16	0.28	0.65	0	4
	Rail	7.19	11.46	48.46	0	29
Average Number of bicycle spaces	Bus	0.75	4.16	33	0	65

Variable		Mean	Std Dev	Max	Min	Count
in secure facilities - locker/facility per Station	Demand Response	N/A	N/A	0	0	0
	Ferry	0.01	0.02	0.05	0	4
	Rail	3.06	7.15	34.07	0	29
Real Time Information System Score	Bus	0.33	0.45	1	0	69
	Demand Response	0.37	0.47	1	0	65
	Ferry	0.40	0.49	1	0	5
	Rail	0.42	0.47	1	0	37
Phone Access to Real Time Information System Score	Bus	0.13	0.30	1	0	69
	Demand Response	0.24	0.42	1	0	65
	Ferry	0.40	0.49	1	0	5
	Rail	0.21	0.39	1	0	37
Smartphone App Access to Real Time Information System Score	Bus	0.27	0.42	1	0	69
	Demand Response	0.25	0.43	1	0	65
	Ferry	0.20	0.40	1	0	5
	Rail	0.34	0.46	1	0	37
Text/SMS Access to Real Time Information System Score	Bus	0.26	0.42	1	0	69
	Demand Response	0.30	0.45	1	0	65
	Ferry	0.20	0.40	1	0	5
	Rail	0.43	0.47	1	0	37
Website Access to Real Time Information System Score	Bus	0.34	0.45	1	0	69
	Demand Response	0.37	0.47	1	0	65
	Ferry	0.20	0.40	1	0	5
	Rail	0.41	0.47	1	0	37

Appendix C

ELIMINATED DATA

Table 5: List of Removed Variables

Variable Removed	Reason for Removal
Ferry Mode	Lack of Data
Demand Response Mode	Lack of Data
Average Number of Stations with Restrooms [Bus and Rail]	Highly correlated with Average Number of Attended Stations
Housing Unit Count for Urban Area	Highly correlated with Urban Area Population
CBSA Population	Highly correlated with Urban Area Population, Housing Unit Count, and CBSA Employment
CBSA Workers	Highly correlated with Urban Area Population, CBSA Employment, CBSA Population
Urban Area Employment	Highly correlated with Urban Area Population, Housing Unit Count, and Land Area
Mean Household Income	Highly correlated with Median Household Income, which is a better indicator
Weighted Average Population Density of CBG in Urban Area	Highly correlated with Urban Area Population, Housing Unit Count, CBSA Employment, CBSA Population, CBSA Workers, Weighted Average Residential Density, and Weighted Average Employment Density
Number of workers earning $1250/month or less [both home and work location]	Highly correlated with Urban Area Population, Housing Unit Count, and Urban Area Employment
Auto Oriented Network Density	Highly correlated with Total Road Network Density
Annual Passenger Miles [all modes]	Highly correlated with Annual Unlinked Passenger Trips
Average number of bus stations with all day auto parking	Highly correlated with Average number of rail stations with all day auto parking
Average Number of stops with auto lot [both rail and bus]	Correlated with Total Road Network Density
Annual Unlinked Passenger Trips [Ferry and Demand Response]	Lack of Data
Average Adult Base Fare [Ferry and Demand Response]	Lack of Data
Transfer Surcharge Score [All Modes]	Lack of Data
Zone Surcharge Score [All Modes]	Lack of Data

Variable Removed	Reason for Removal
Peak-Hour Surcharge Score [All Modes]	Lack of Data
Cash Surcharge Score [All Modes]	Lack of Data
Multi-agency Regional Pass Score [All Modes]	Lack of Data
Stored Value Media Score [All Modes]	Lack of Data
Regional Smart Card Score [All Modes]	Lack of Data
Smart Card Use Score [All Modes]	Lack of Data
Guaranteed Ride Home Plan Score [All Modes]	Lack of Data
Average Number of ADA Accessible Stations per Total Stations [Rail, Ferry, DR]	Lack of Data
Average Number of Attended stations per Total Stations [Rail, Ferry, DR]	Lack of Data
Average Number of Stations with Security Cameras per Total Stations [Rail, Ferry, DR]	Lack of Data
Average Number of stations with emergency call buttons per total stations [Rail, Ferry, DR]	Lack of Data
Average Number of Stations with Concessions per Total Stations [Rail, Ferry, DR]	Lack of Data
Average Number of stations with all-day auto parking per total stations [Rail, Ferry, DR]	Lack of Data
Average Number of stations with lockers or secure bike facility per Total Station [Rail, Ferry, DR]	Lack of Data
Average Number of stations with bike racks per total stations [Rail, Ferry, DR]	Lack of Data
Average Number of stations with Bikeshare station per total stations [Rail, Ferry, DR]	Lack of Data
Average Number of Stops with Shelters per total Stops [Rail, Ferry, DR]	Lack of Data
Average Number of Stops with Benches per total stops [Rail, Ferry, DR]	Lack of Data
Average Number of Stops with Lighting per total stops [Rail, Ferry, DR]	Lack of Data
Average Number of stops with auto lot per total stops [Rail, Ferry, DR]	Lack of Data
Average Number of All-day Parking Spaces per Station [Rail, Ferry, DR]	Lack of Data
Average Number of Part-day Parking Spaces per Station [Rail, Ferry, DR]	Lack of Data
Average Number of Bicycle Rack Spaces per Station [Rail, Ferry, DR]	Lack of Data
Average Number of bicycle spaces in secure facilities - locker/facility per Station [Rail, Ferry, DR]	Lack of Data

Appendix D

FACTOR ANALYSIS

Table 6: Factor Scores for Each City

Name	Factor A	Factor B	Factor C	Factor D	Factor E	Factor F	Factor G	Factor H	Factor I	Factor J	Factor K
Akron, OH	-1.70	0.77	-0.85	1.62	-1.55	-0.27	0.00	-2.24	1.21	-1.24	-0.78
Albany--Schenectady, NY	-1.87	-0.50	1.09	1.58	-1.85	0.59	0.00	3.26	-0.53	-0.58	-1.03
Albuquerque, NM	-1.63	1.28	-0.99	-2.85	0.22	-0.18	0.00	0.00	0.00	0.00	0.00
Allentown, PA--NJ	-1.62	0.54	0.41	2.17	-1.51	-0.50	0.00	-2.24	-0.53	-1.29	2.61
Atlanta, GA	5.38	-1.69	-1.27	0.00	-1.52	1.23	-3.20	-2.24	-0.53	-1.14	-1.35
Augusta-Richmond County, GA--SC	-2.42	1.62	-2.32	-0.04	-2.26	-0.50	0.00	0.00	0.00	0.00	0.00
Austin, TX	0.22	-3.03	-1.14	-0.41	-0.55	0.83	3.97	5.34	0.00	0.00	0.00
Bakersfield, CA	-2.37	0.68	-1.14	-1.61	1.90	-3.01	0.00	-2.24	-0.53	0.00	-0.24
Baltimore, MD	0.93	-1.35	2.11	0.10	0.94	0.06	2.32	3.12	-0.53	-0.85	-3.01
Baton Rouge, LA	-1.80	0.26	-1.64	0.39	-1.83	-0.10	0.00	0.00	0.00	0.00	0.00
Birmingham, AL	-1.45	1.38	-1.68	0.01	-2.41	0.46	0.00	-2.24	-0.53	-1.41	-3.01
Boston, MA--NH--RI	3.72	-2.08	4.26	6.26	0.06	1.50	5.57	5.34	19.41	-0.64	2.61
Bridgeport--Stamford, CT--NY	-0.67	-3.21	0.72	1.51	-1.69	1.51	0.00	-2.24	-0.53	-1.12	1.64
Buffalo, NY	-1.38	1.86	1.63	1.94	-0.61	0.07	-3.20	-2.24	-0.53	-1.25	-0.11
Cape Coral, FL	-2.38	3.17	-2.53	-2.71	-1.61	-1.14	0.00	0.00	0.00	0.00	0.00
Charleston--North Charleston, SC	-2.21	-0.85	-1.63	0.71	-2.01	0.72	0.00	1.04	-0.53	-1.25	0.71
Charlotte, NC--SC	0.12	-2.09	-2.07	0.81	-1.85	1.78	-3.20	-2.24	0.63	-0.94	1.68
Chattanooga, TN--GA	-2.15	1.37	-1.97	-2.22	-3.09	-0.44	-3.20	-2.24	-0.53	1.80	1.98
Chicago, IL--IN	9.02	-1.50	3.85	-0.64	3.66	0.29	-3.01	5.34	-0.53	-0.69	-2.85
Cleveland, OH	0.34	1.49	0.45	0.64	-0.05	0.08	-3.20	-2.24	-0.53	-0.90	-1.50
Colorado Springs, CO	-2.20	-0.79	-1.70	-1.05	-0.27	0.77	0.00	-2.24	-0.53	0.27	-0.43
Columbia, SC	-1.86	-0.35	-1.82	-0.25	-2.12	0.62	0.00	0.00	0.00	0.00	0.00
Columbus, OH	0.14	-1.20	-0.56	0.94	0.85	1.01	0.00	-2.24	-0.53	-0.53	0.70

Name	Factor A	Factor B	Factor C	Factor D	Factor E	Factor F	Factor G	Factor H	Factor I	Factor J	Factor K
Concord, CA	-0.85	-4.70	-0.89	-0.87	0.61	2.76	0.00	-2.24	0.00	0.00	0.00
Dallas--Ft Worth--Arlington, TX	5.42	-1.82	-0.79	-1.56	1.92	0.25	3.97	4.43	-0.08	0.13	0.34
Dayton, OH	-1.78	1.89	-0.98	1.47	-0.75	-0.16	-1.65	0.00	0.00	0.00	0.00
Denver--Aurora, CO	0.90	-2.52	0.43	-1.92	2.45	1.08	0.00	0.00	0.00	0.00	0.00
Des Moines, IA	-2.15	-1.66	-1.22	3.40	-0.72	1.24	0.00	-2.24	0.00	0.00	0.00
Detroit, MI	2.40	1.61	-0.08	-1.41	1.88	0.05	3.55	3.58	-0.53	-1.21	1.31
El Paso, TX--NM	-1.72	2.67	-0.81	-3.43	0.14	-2.27	0.00	-2.24	-0.53	1.53	1.66
Fresno, CA	-1.85	2.56	-0.11	0.69	2.52	-2.48	0.00	0.00	0.00	0.00	0.00
Fresno, CA	0.02	-0.21	-0.28	0.48	-1.35	0.47	0.00	1.04	-0.53	-1.32	-0.82
Grand Rapids, MI	-1.80	0.01	-0.60	1.75	-0.36	0.36	0.00	-2.24	0.00	0.00	0.00
Greenville, SC	-2.14	1.89	-1.95	2.40	-2.83	-0.35	0.00	0.00	0.00	0.00	0.00
Harrisburg, PA	-2.12	-0.07	0.14	3.10	-1.50	0.35	0.00	-2.24	-0.53	-1.26	-0.24
Hartford, CT	-1.02	-0.40	0.76	0.46	-1.82	-0.07	0.00	-2.17	-0.53	-1.37	1.65
Houston, TX	6.46	-2.04	-0.54	-0.54	2.08	-1.02	-3.20	-2.24	-0.53	-0.49	-0.23
Indianapolis, IN	0.33	0.00	-1.59	0.80	-0.42	0.36	0.00	-2.24	-0.53	-1.11	-3.01
Jacksonville, FL	-0.52	0.85	-1.65	0.56	-1.44	-0.26	4.02	0.00	-0.53	7.16	-3.01
Kansas City, MO--KS	-0.08	-0.23	-1.35	0.47	-0.09	1.37	0.00	-2.24	-0.53	-1.15	0.34
Knoxville, TN	-1.79	2.40	-2.26	-2.02	-2.58	0.17	0.00	0.00	0.00	0.00	0.00
Lancaster, PA	-2.32	0.39	0.09	5.05	-2.73	-0.36	0.00	0.00	0.00	0.00	0.00
Las Vegas--Henderson, NV	-0.19	0.26	-0.04	-4.20	2.93	-1.55	0.00	0.00	0.00	0.00	0.00
Little Rock, AR	-2.14	1.40	-1.75	-0.26	-1.67	0.28	-3.20	-2.24	-0.53	-1.24	2.61
Los Angeles--Long Beach--Anaheim, CA	13.48	-1.07	2.70	0.88	10.80	-1.06	1.99	-1.81	-0.49	0.06	-2.56
Louisville/Jefferson County, KY--IN	-1.07	1.31	-0.87	-0.62	-1.26	-0.23	0.00	0.00	-0.53	-1.17	2.61
Madison, WI	-2.21	3.52	0.53	-1.01	-0.61	-3.83	0.00	1.04	-0.53	-0.81	0.70
McAllen, TX	-2.06	-1.83	-0.99	-2.77	-1.42	2.80	0.00	0.00	0.00	0.00	0.00
Memphis, TN--MS--AR	-0.67	1.46	-1.08	-0.50	-0.56	-0.51	3.92	0.00	0.00	0.00	0.00
Miami, FL	3.02	0.89	2.00	-1.90	2.79	-0.53	-3.20	-2.20	-0.53	-0.28	-1.12
Milwaukee, WI	-0.46	0.48	1.57	-1.64	0.47	0.34	0.00	5.34	0.00	0.00	0.00
Minneapolis--St. Paul, MN--WI	1.62	-2.29	1.62	-0.08	0.80	2.22	-3.20	1.38	5.21	8.93	2.48
Mission Viejo--Lake Forest--San Clemente, CA	0.35	-4.74	-1.07	-1.97	3.12	2.22	0.00	0.00	0.00	0.00	0.00
Nashville-Davidson, TN	-1.27	-0.53	-1.72	-1.68	-1.78	0.59	-3.20	-2.24	-0.53	-0.58	0.71
New Haven, CT	-1.80	-0.76	1.03	1.19	-1.81	0.65	-3.20	-2.24	0.00	0.00	0.00
New Orleans, LA	-1.75	1.43	1.14	-1.83	1.83	-0.43	-3.20	-2.24	-0.53	-1.29	-3.01

Name	Factor A	Factor B	Factor C	Factor D	Factor E	Factor F	Factor G	Factor H	Factor I	Factor J	Factor K
New York--Newark, NY--NJ--CT	25.58	-1.32	16.27	-0.79	5.99	-0.02	-3.19	-1.95	-0.53	0.16	-0.53
Oklahoma City, OK	-0.99	1.04	-1.54	0.04	-0.45	-0.34	0.00	-2.24	-0.53	3.53	2.61
Omaha, NE--IA	-1.44	-1.50	-0.78	-0.82	0.63	0.65	0.00	0.00	0.00	0.00	0.00
Orlando, FL	-0.16	-0.40	-1.36	0.75	-0.52	0.08	0.00	1.04	-0.53	2.35	0.37
Palm Bay--Melbourne, FL	-2.66	2.35	-2.77	-1.38	-1.41	-0.09	0.00	0.00	-0.53	-0.24	-3.01
Philadelphia, PA--NJ--DE--MD	6.27	-0.27	3.41	0.95	0.12	-0.32	2.01	1.76	-0.53	-1.08	-1.97
Phoenix--Mesa, AZ	3.09	0.08	-0.71	-2.89	2.10	0.35	0.00	-2.24	-0.53	0.48	0.36
Pittsburgh, PA	0.38	1.06	1.11	2.76	-1.55	0.84	0.18	-0.81	1.37	-0.84	-0.08
Portland, OR--WA	-0.01	-0.48	1.05	0.98	3.06	0.75	5.57	-1.09	-0.15	0.12	-0.41
Poughkeepsie--Newburgh, NY--NJ	-2.07	-1.02	-0.28	0.35	-3.46	0.10	0.00	0.00	0.00	0.00	0.00
Providence, RI--MA	-1.07	1.28	1.30	4.79	-0.22	-1.50	0.00	-2.24	-0.53	-1.24	0.71
Raleigh, NC	-0.74	-3.18	-2.12	1.08	-2.30	2.30	0.00	0.00	0.00	0.00	0.00
Reno, NV--CA	-2.37	0.24	-0.43	-3.54	-0.02	-0.55	0.00	3.83	-0.53	0.74	1.34
Richmond, VA	-1.44	-0.60	-0.88	-0.04	-1.79	-0.16	0.00	0.00	-0.53	-0.62	-3.01
Riverside--San Bernardino, CA	0.56	-0.85	-1.35	-1.57	2.25	-2.56	0.00	2.40	-0.36	1.42	-0.50
Rochester, NY	-1.69	0.66	0.74	0.31	-1.86	0.60	0.00	-2.24	0.00	0.00	0.00
Sacramento, CA	-0.33	0.15	-0.34	-0.90	2.07	-0.03	0.00	0.00	0.00	0.00	0.00
Salt Lake City--West Valley City--Ogden--Layton--Provo--Orem, UT	-0.10	-7.90	-0.77	0.49	1.45	-12.31	3.97	5.34	-0.53	-0.65	-1.20
San Antonio, TX	0.53	-2.50	-1.35	-1.86	0.53	0.27	3.97	5.34	0.51	1.04	2.04
San Diego, CA	1.60	1.05	0.72	-1.50	3.13	-1.23	2.39	4.87	-0.53	0.51	1.58
San Francisco--Oakland, CA	3.00	-1.51	5.46	-0.67	7.28	-0.08	2.07	-2.18	-0.53	-1.19	0.69
San Jose, CA	0.41	-2.97	2.07	-0.98	5.82	1.74	4.02	3.83	-0.53	1.66	-1.11
Sarasota--Bradenton, FL	-2.22	12.37	0.85	-0.10	-1.00	-1.58	0.00	-2.24	-0.53	0.03	1.89
Scranton, PA	-2.66	2.07	-0.60	2.70	-1.79	0.21	0.00	-2.24	-0.53	-1.22	-3.01
Seattle, WA	1.58	2.66	2.31	-1.47	1.93	-0.05	0.02	0.09	0.16	4.75	-0.46
Spokane, WA	-2.46	-1.62	-0.64	-0.95	0.10	2.22	0.00	3.58	-0.53	-1.21	0.96
Springfield, MA--CT	-1.46	2.48	-0.38	5.61	-2.07	-0.06	0.00	-2.24	0.00	0.00	0.00
St. Louis, MO--IL	0.52	1.77	0.64	0.97	0.33	-0.41	-3.20	-2.24	-0.53	-1.09	0.71
Syracuse, NY	-2.29	1.54	1.19	-0.13	-1.80	0.35	0.00	5.34	-0.53	-1.10	2.07
Tampa--St. Petersburg, FL	0.80	1.68	-0.69	-1.02	0.22	0.06	-3.20	0.83	-0.53	0.06	0.49
Toledo, OH--MI	-2.10	2.14	-0.61	0.92	-0.53	-0.31	0.00	3.83	-0.53	-1.29	0.71

Name	Factor A	Factor B	Factor C	Factor D	Factor E	Factor F	Factor G	Factor H	Factor I	Factor J	Factor K
Tucson, AZ	-1.59	2.81	-0.73	-2.67	-0.45	-0.20	0.00	0.00	0.00	0.00	0.00
Tulsa, OK	-1.49	0.37	-1.39	1.12	-1.11	-0.16	0.00	0.00	0.00	0.00	0.00
Urban Honolulu, HI	-1.86	-1.50	4.79	-3.98	2.15	0.55	0.00	3.26	-0.53	4.55	-1.01
Virginia Beach, VA	-0.89	-0.95	-1.11	-1.71	0.08	-0.13	-3.20	-2.24	-0.53	-0.93	-0.97
Washington, DC--VA--MD	4.86	-4.95	3.36	-0.94	1.35	3.06	5.57	4.96	0.24	-0.36	0.09
Wichita, KS	-1.86	-0.15	-1.57	0.00	-0.33	-0.10	0.00	0.00	0.00	0.00	0.00
Winston-Salem, NC	-2.02	1.44	-2.13	-0.59	-3.18	0.06	-3.20	-2.24	-0.53	-1.15	2.61
Worcester, MA--CT	-1.71	-0.51	-0.01	5.04	-2.85	0.40	0.00	-2.24	0.00	0.00	0.00
Youngstown, OH--PA	-2.52	4.03	-1.26	2.21	-1.98	-1.92	0.00	0.00	0.00	0.00	0.00

Appendix E

CLUSTER ANALYSIS

Table 7: Cities by Cluster and Factor Scores

Name	Cluster	Factor A	Factor B	Factor C	Factor D	Factor E	Factor F	Factor G	Factor H	Factor I	Factor J	Factor K	Choice Riders	Land Use Mix (Com, Res, Ind)	Average Base Adult Bus Fare	Average Base Adult Rail Fare
Akron, OH	Cluster 11	-1.70	0.77	-0.85	1.62	-1.55	-0.27	0.00	-2.24	1.21	-1.24	-0.78	-0.88	0.00	3.11	0.00
Albany--Schenectady, NY	Cluster 2	-1.87	-0.50	1.09	1.58	-1.85	0.59	0.00	3.26	-0.53	-0.58	-1.03	-0.83	0.49	-0.26	0.00
Albuquerque, NM	Cluster 11	-1.63	1.28	-0.99	-2.85	0.22	-0.18	0.00	0.00	0.00	0.00	0.00	0.90	0.49	-1.31	-0.06
Allentown, PA--NJ	Cluster 11	-1.62	0.54	0.41	2.17	-1.51	-0.50	0.00	-2.24	-0.53	-1.29	2.61	-0.52	0.00	0.24	0.00
Atlanta, GA	Cluster 1	5.38	-1.69	-1.27	0.00	-1.52	1.23	-3.20	-2.24	-0.53	-1.14	-1.35	0.19	-1.17	-0.01	-0.07
Augusta-Richmond County, GA--SC	Cluster 11	-2.42	1.62	-2.32	-0.04	-2.26	-0.50	0.00	0.00	0.00	0.00	0.00	-0.40	0.00	-0.57	0.00
Austin, TX	Cluster 7	0.22	-3.03	-1.14	-0.41	-0.55	0.83	3.97	5.34	0.00	0.00	0.00	0.78	1.27	0.09	0.60
Bakersfield, CA	Cluster 9	-2.37	0.68	-1.14	-1.61	1.90	-3.01	0.00	-2.24	-0.53	0.00	-0.24	0.38	1.04	-0.30	0.00
Baltimore, MD	Cluster 6	0.93	-1.35	2.11	0.10	0.94	0.06	2.32	3.12	-0.53	-0.85	-3.01	-0.34	1.42	0.23	0.73
Baton Rouge, LA	Cluster 11	-1.80	0.26	-1.64	0.39	-1.83	-0.10	0.00	0.00	0.00	0.00	0.00	-0.08	0.00	0.00	-0.17
Birmingham, AL	Cluster 11	-1.45	1.38	-1.68	0.01	-2.41	0.46	0.00	-2.24	-0.53	-1.41	-3.01	-0.05	0.00	-0.57	0.00
Boston, MA--NH--RI	Individual Cluster	3.72	-2.08	4.26	6.26	0.06	1.50	5.57	5.34	19.41	-0.64	2.61	0.78	0.00	-0.54	-0.01
Bridgeport--Stamford, CT--NY	Cluster 4	-0.67	-3.21	0.72	1.51	-1.69	1.51	0.00	-2.24	-0.53	-1.12	1.64	1.82	0.00	-0.29	0.00
Buffalo, NY	Cluster 11	-1.38	1.86	1.63	1.94	-0.61	0.07	-3.20	-2.24	-0.53	-1.25	-0.11	-1.42	0.00	0.24	-0.06
Cape Coral, FL	Cluster 11	-2.38	3.17	-2.53	-2.71	-1.61	-1.14	0.00	0.00	0.00	0.00	0.00	0.14	0.34	-0.30	0.00
Charleston--North Charleston, SC	Cluster 11	-2.21	-0.85	-1.63	0.71	-2.01	0.72	0.00	1.04	-0.53	-1.25	0.71	-1.04	-0.55	0.16	0.00
Charlotte, NC--SC	Cluster 11	0.12	-2.09	-2.07	0.81	-1.85	1.78	-3.20	-2.24	0.63	-0.94	1.68	-0.36	-1.24	0.52	0.03
Chattanooga, TN--GA	Individual Cluster	-2.15	1.37	-1.97	-2.22	-3.09	-0.44	-3.20	-2.24	-0.53	1.80	1.98	-1.49	0.00	-0.30	5.67
Chicago, IL--IN	Individual Cluster	9.02	-1.50	3.85	-0.64	3.66	0.29	-3.01	5.34	-0.53	-0.69	-2.85	0.69	0.18	1.01	0.69

Name	Cluster	Factor A	Factor B	Factor C	Factor D	Factor E	Factor F	Factor G	Factor H	Factor I	Factor J	Factor K	Choice Riders	Land Use Mix (Com, Res, Ind)	Average Base Adult Bus Fare	Average Base Adult Rail Fare
Cincinnati, OH--KY--IN	Cluster 11	0.02	-0.21	-0.28	0.48	-1.35	0.47	0.00	1.04	-0.53	-1.32	-0.82	-0.46	0.04	-0.07	0.00
Cleveland, OH	Cluster 11	0.34	1.49	0.45	0.64	-0.05	0.08	-3.20	-2.24	-0.53	-0.90	-1.50	-0.55	0.00	0.49	0.05
Colorado Springs, CO	Cluster 11	-2.20	-0.79	-1.70	-1.05	-0.27	0.77	0.00	-2.24	-0.53	0.27	-0.43	-0.53	0.00	-0.03	0.00
Columbia, SC	Cluster 11	-1.86	-0.35	-1.82	-0.25	-2.12	0.62	0.00	0.00	0.00	0.00	0.00	-0.96	0.00	0.00	0.00
Columbus, OH	Cluster 11	0.14	-1.20	-0.56	0.94	0.85	1.01	0.00	-2.24	-0.53	-0.53	0.70	-0.33	0.15	-0.84	0.00
Concord, CA	Cluster 4	-0.85	-4.70	-0.89	-0.87	0.61	2.76	0.00	-2.24	0.00	0.00	0.00	2.87	0.00	0.24	0.00
Dallas--Fort Worth--Arlington, TX	Cluster 7	5.42	-1.82	-0.79	-1.56	1.92	0.25	3.97	4.43	-0.08	0.13	0.34	0.50	-1.41	1.37	-0.84
Dayton, OH	Cluster 11	-1.78	1.89	-0.98	1.47	-0.75	-0.16	-1.65	0.00	0.00	0.00	0.00	-0.89	0.00	0.00	-0.94
Denver--Aurora, CO	Cluster 4	0.90	-2.52	0.43	-1.92	2.45	1.08	0.00	0.00	0.00	0.00	0.00	1.33	-0.34	0.00	0.00
Des Moines, IA	Cluster 11	-2.15	-1.66	-1.22	3.40	-0.72	1.24	0.00	-2.24	0.00	0.00	0.00	0.48	-0.73	-0.03	0.00
Detroit, MI	Cluster 7	2.40	1.61	-0.08	-1.41	1.88	0.05	3.55	3.58	-0.53	-1.21	1.31	-0.59	-0.53	0.07	-0.61
El Paso, TX--NM	Cluster 8	-1.72	2.67	-0.81	-3.43	0.14	-2.27	0.00	-2.24	-0.53	1.53	1.66	-0.09	-0.21	-0.30	0.00
Fresno, CA	Cluster 9	-1.85	2.56	-0.11	0.69	2.52	-2.48	0.00	0.00	0.00	0.00	0.00	0.00	1.06	0.00	0.00
Grand Rapids, MI	Cluster 11	-1.80	0.01	-0.60	1.75	-0.36	0.36	0.00	-2.24	0.00	0.00	0.00	0.33	0.00	-0.30	0.00
Greenville, SC	Cluster 11	-2.14	1.89	-1.95	2.40	-2.83	-0.35	0.00	0.00	0.00	0.00	0.00	0.02	0.00	-0.30	0.00
Harrisburg, PA	Cluster 11	-2.12	-0.07	0.14	3.10	-1.50	0.35	0.00	-2.24	-0.53	-1.26	-0.24	-0.10	0.00	-0.03	0.00
Hartford, CT	Cluster 11	-1.02	-0.40	0.76	0.46	-1.82	-0.07	0.00	-2.17	-0.53	-1.37	1.65	-0.61	1.23	0.35	0.49
Houston, TX	Cluster 1	6.46	-2.04	-0.54	-0.54	2.08	-1.02	-3.20	-2.24	-0.53	-0.49	-0.23	0.43	0.84	-0.17	-0.39
Indianapolis, IN	Cluster 11	0.33	0.00	-1.59	0.80	-0.42	0.36	0.00	-2.24	-0.53	-1.11	-3.01	-0.54	0.00	-0.03	0.00
Jacksonville, FL	Individual Cluster	-0.52	0.85	-1.65	0.56	-1.44	-0.26	4.02	0.00	-0.53	7.16	-3.01	-0.98	0.00	-1.94	-0.94
Kansas City, MO--KS	Cluster 11	-0.08	-0.23	-1.35	0.47	-0.09	1.37	0.00	-2.24	-0.53	-1.15	0.34	0.03	-1.46	0.90	0.00
Knoxville, TN	Cluster 11	-1.79	2.40	-2.26	-2.02	-2.58	0.17	0.00	0.00	0.00	0.00	0.00	-1.15	0.00	0.00	0.00
Lancaster, PA	Cluster 10	-2.32	0.39	0.09	5.05	-2.73	-0.36	0.00	0.00	0.00	0.00	0.00	-0.90	0.00	0.00	0.00

Name	Cluster	Factor A	Factor B	Factor C	Factor D	Factor E	Factor F	Factor G	Factor H	Factor I	Factor J	Factor K	Choice Riders	Land Use Mix (Com, Res, Ind)	Average Base Adult Bus Fare	Average Base Adult Rail Fare
Las Vegas--Henderson, NV	Cluster 3	-0.19	0.26	-0.04	-4.20	2.93	-1.55	0.00	0.00	0.00	0.00	0.00	-0.70	0.51	0.00	1.26
Little Rock, AR	Cluster 11	-2.14	1.40	-1.75	-0.26	-1.67	0.28	-3.20	-2.24	-0.53	-1.24	2.61	-0.96	0.00	-0.47	-0.50
Los Angeles--Long Beach--Anaheim, CA	Individual Cluster	13.48	-1.07	2.70	0.88	10.80	-1.06	1.99	-1.81	-0.49	0.06	-2.56	0.47	0.12	-1.16	-0.10
Louisville/Jefferson County, KY--IN	Cluster 11	-1.07	1.31	-0.87	-0.62	-1.26	-0.23	0.00	0.00	-0.53	-1.17	2.61	-1.13	0.00	-0.03	0.00
Madison, WI	Cluster 9	-2.21	3.52	0.53	-1.01	-0.61	-3.83	0.00	1.04	-0.53	-0.81	0.70	0.73	0.14	0.24	0.00
McAllen, TX	Cluster 4	-2.06	-1.83	-0.99	-2.77	-1.42	2.80	0.00	0.00	0.00	0.00	0.00	1.77	0.00	-0.85	0.00
Memphis, TN--MS--AR	Cluster 11	-0.67	1.46	-1.08	-0.50	-0.56	-0.51	3.92	0.00	0.00	0.00	0.00	-1.48	0.00	0.00	-0.50
Miami, FL	Cluster 3	3.02	0.89	2.00	-1.90	2.79	-0.53	-3.20	-2.20	-0.53	-0.28	-1.12	0.24	-0.10	0.33	0.15
Milwaukee, WI	Cluster 2	-0.46	0.48	1.57	-1.64	0.47	0.34	0.00	5.34	0.00	0.00	0.00	-0.68	0.00	1.26	0.00
Minneapolis--St. Paul, MN--WI	Individual Cluster	1.62	-2.29	1.62	-0.08	0.80	2.22	-3.20	1.38	5.21	8.93	2.48	1.25	-1.50	1.16	-0.17
Mission Viejo--Lake Forest--San Clemente, CA	Cluster 4	0.35	-4.74	-1.07	-1.97	3.12	2.22	0.00	0.00	0.00	0.00	0.00	2.30	0.00	-1.12	0.00
Nashville-Davidson, TN	Cluster 11	-1.27	-0.53	-1.72	-1.68	-1.78	0.59	-3.20	-2.24	-0.53	-0.58	0.71	0.45	0.00	0.11	-0.06
New Haven, CT	Cluster 11	-1.80	-0.76	1.03	1.19	-1.81	0.65	-3.20	-2.24	0.00	0.00	0.00	0.13	0.00	0.57	-0.94
New Orleans, LA	Cluster 11	-1.75	1.43	1.14	-1.83	1.83	-0.43	-3.20	-2.24	-0.53	-1.29	-3.01	-1.14	-0.25	-0.38	-0.39
New York--Newark, NY--NJ--CT	Individual Cluster	25.58	-1.32	16.27	-0.79	5.99	-0.02	-3.19	-1.95	-0.53	0.16	-0.53	-1.46	0.75	0.11	-0.54
Oklahoma City, OK	Cluster 8	-0.99	1.04	-1.54	0.04	-0.45	-0.34	0.00	-2.24	-0.53	3.53	2.61	-0.20	0.00	-0.03	0.00
Omaha, NE--IA	Cluster 11	-1.44	-1.50	-0.78	-0.82	0.63	0.65	0.00	0.00	0.00	0.00	0.00	-0.16	-0.33	-0.57	0.00
Orlando, FL	Cluster 11	-0.16	-0.40	-1.36	0.75	-0.52	0.08	0.00	1.04	-0.53	2.35	0.37	-0.27	0.04	-0.87	0.00
Palm Bay--Melbourne, FL	Cluster 11	-2.66	2.35	-2.77	-1.38	-1.41	-0.09	0.00	0.00	-0.53	-0.24	-3.01	0.58	0.00	-0.30	0.00
Philadelphia, PA--NJ--DE--MD	Cluster 6	6.27	-0.27	3.41	0.95	0.12	-0.32	2.01	1.76	-0.53	-1.08	-1.97	0.18	1.68	-0.62	1.10
Phoenix--Mesa, AZ	Cluster 3	3.09	0.08	-0.71	-2.89	2.10	0.35	0.00	-2.24	-0.53	0.48	0.36	0.03	-1.43	-0.57	-0.50
Pittsburgh, PA	Cluster 10	0.38	1.06	1.11	2.76	-1.55	0.84	0.18	-0.81	1.37	-0.84	-0.08	0.76	0.00	-0.37	0.16
Portland, OR--WA	Cluster 5	-0.01	-0.48	1.05	0.98	3.06	0.75	5.57	-1.09	-0.15	0.12	-0.41	1.15	0.21	1.89	-0.19

Name	Cluster	Factor A	Factor B	Factor C	Factor D	Factor E	Factor F	Factor G	Factor H	Factor I	Factor J	Factor K	Choice Riders	Land Use Mix (Com, Res, Ind)	Average Base Adult Bus Fare	Average Base Adult Rail Fare
Poughkeepsie--Newburgh, NY--NJ	Cluster 11	-2.07	-1.02	-0.28	0.35	-3.46	0.10	0.00	0.00	0.00	0.00	0.00	1.83	0.00	-0.85	0.00
Providence, RI--MA	Cluster 10	-1.07	1.28	1.30	4.79	-0.22	-1.50	0.00	-2.24	-0.53	-1.24	0.71	1.10	1.63	0.64	0.00
Raleigh, NC	Cluster 4	-0.74	-3.18	-2.12	1.08	-2.30	2.30	0.00	0.00	0.00	0.00	0.00	-0.06	-0.49	-0.50	0.00
Reno, NV--CA	Cluster 2	-2.37	0.24	-0.43	-3.54	-0.02	-0.55	0.00	3.83	-0.53	0.74	1.34	-1.05	0.00	4.58	0.00
Richmond, VA	Cluster 11	-1.44	-0.60	-0.88	-0.04	-1.79	-0.16	0.00	0.00	-0.53	-0.62	-3.01	-0.60	1.35	-0.03	0.00
Riverside--San Bernardino, CA	Cluster 3	0.56	-0.85	-1.35	-1.57	2.25	-2.56	0.00	2.40	-0.36	1.42	-0.50	1.81	0.00	0.32	-0.94
Rochester, NY	Cluster 11	-1.69	0.66	0.74	0.31	-1.86	0.60	0.00	-2.24	0.00	0.00	0.00	-1.83	0.00	-0.85	0.00
Sacramento, CA	Cluster 3	-0.33	0.15	-0.34	-0.90	2.07	-0.03	0.00	0.00	0.00	0.00	0.00	1.65	-0.17	1.50	-0.94
Salt Lake City--West Valley City--Ogden--Layton--Provo--Orem, UT	Individual Cluster	-0.10	-7.90	-0.77	0.49	1.45	-12.31	3.97	5.34	-0.53	-0.65	-1.20	1.64	2.25	2.28	0.16
San Antonio, TX	Cluster 7	0.53	-2.50	-1.35	-1.86	0.53	0.27	3.97	5.34	0.51	1.04	2.04	-0.64	-0.44	-0.52	-0.94
San Diego, CA	Cluster 7	1.60	1.05	0.72	-1.50	3.13	-1.23	2.39	4.87	-0.53	0.51	1.58	0.99	0.58	-1.19	0.21
San Francisco--Oakland, CA	Individual Cluster	3.00	-1.51	5.46	-0.67	7.28	-0.08	2.07	-2.18	-0.53	-1.19	0.69	0.98	0.92	0.87	0.22
San Jose, CA	Cluster 5	0.41	-2.97	2.07	-0.98	5.82	1.74	4.02	3.83	-0.53	1.66	-1.11	1.90	-0.85	0.24	-0.06
Sarasota--Bradenton, FL	Individual Cluster	-2.22	12.37	0.85	-0.10	-1.00	-1.58	0.00	-2.24	-0.53	0.03	1.89	-0.31	0.00	-0.54	0.00
Scranton, PA	Cluster 11	-2.66	2.07	-0.60	2.70	-1.79	0.21	0.00	-2.24	-0.53	-1.22	-3.01	-0.54	0.00	-0.03	0.00
Seattle, WA	Cluster 3	1.58	2.66	2.31	-1.47	1.93	-0.05	0.02	0.09	0.16	4.75	-0.46	1.65	-1.10	-1.43	0.17
Spokane, WA	Cluster 4	-2.46	-1.62	-0.64	-0.95	0.10	2.22	0.00	3.58	-0.53	-1.21	0.96	1.06	-2.09	-0.30	0.00
Springfield, MA--CT	Cluster 10	-1.46	2.48	-0.38	5.61	-2.07	-0.06	0.00	-2.24	0.00	0.00	0.00	-0.30	0.00	-0.57	0.00
St. Louis, MO--IL	Cluster 11	0.52	1.77	0.64	0.97	0.33	-0.41	-3.20	-2.24	-0.53	-1.09	0.71	-0.79	0.00	0.19	0.16
Syracuse, NY	Cluster 2	-2.29	1.54	1.19	-0.13	-1.80	0.35	0.00	5.34	-0.53	-1.10	2.07	-1.86	0.00	-0.20	0.00
Tampa--St. Petersburg, FL	Cluster 11	0.80	1.68	-0.69	-1.02	0.22	0.06	-3.20	0.83	-0.53	0.06	0.49	-0.31	-0.75	0.39	0.16
Toledo, OH--MI	Cluster 2	-2.10	2.14	-0.61	0.92	-0.53	-0.31	0.00	3.83	-0.53	-1.29	0.71	-0.90	0.00	-0.57	0.00
Tucson, AZ	Cluster 11	-1.59	2.81	-0.73	-2.67	-0.45	-0.20	0.00	0.00	0.00	0.00	0.00	-0.57	0.00	-0.30	-0.28

Name	Cluster	Factor A	Factor B	Factor C	Factor D	Factor E	Factor F	Factor G	Factor H	Factor I	Factor J	Factor K	Choice Riders	Land Use Mix (Com, Res, Ind)	Average Base Adult Bus Fare	Average Base Adult Rail Fare
Tulsa, OK	Cluster 11	-1.49	0.37	-1.39	1.12	-1.11	-0.16	0.00	0.00	0.00	0.00	0.00	-0.40	0.00	-0.30	0.00
Urban Honolulu, HI	Individual Cluster	-1.86	-1.50	4.79	-3.98	2.15	0.55	0.00	3.26	-0.53	4.55	-1.01	1.18	-0.67	0.79	0.00
Virginia Beach, VA	Cluster 11	-0.89	-0.95	-1.11	-1.71	0.08	-0.13	-3.20	-2.24	-0.53	-0.93	-0.97	-0.63	-0.02	-1.68	-0.28
Washington, DC--VA--MD	Individual Cluster	4.86	-4.95	3.36	-0.94	1.35	3.06	5.57	4.96	0.24	-0.36	0.09	0.94	-1.79	2.86	-0.17
Wichita, KS	Cluster 11	-1.86	-0.15	-1.57	0.00	-0.33	-0.10	0.00	0.00	0.00	0.00	0.00	-1.34	0.00	-0.03	0.00
Winston-Salem, NC	Cluster 11	-2.02	1.44	-2.13	-0.59	-3.18	0.06	-3.20	-2.24	-0.53	-1.15	2.61	-2.20	-0.90	-0.85	-0.50
Worcester, MA--CT	Cluster 10	-1.71	-0.51	-0.01	5.04	-2.85	0.40	0.00	-2.24	0.00	0.00	0.00	0.59	0.00	-0.30	0.00
Youngstown, OH--PA	Cluster 9	-2.52	4.03	-1.26	2.21	-1.98	-1.92	0.00	0.00	0.00	0.00	0.00	0.12	1.80	-0.57	0.00

Appendix F

REGRESSION ANALYSIS

Table 8: Regression Analysis of All Modes

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	0.00	0.04	0.07	0.95	-0.08	0.08	-0.08	0.08
Factor A	**0.17**	0.02	8.66	0.00	0.13	0.21	0.13	0.21
Factor B	-0.02	0.02	-0.73	0.47	-0.06	0.03	-0.06	0.03
Factor C	**0.23**	0.03	7.67	0.00	0.17	0.29	0.17	0.29
Factor D	**-0.07**	0.03	-2.50	0.01	-0.12	-0.01	-0.12	-0.01
Factor E	**-0.13**	0.03	-4.29	0.00	-0.18	-0.07	-0.18	-0.07
Factor F	-0.03	0.03	-1.02	0.31	-0.10	0.03	-0.10	0.03
Factor G	0.02	0.03	0.69	0.49	-0.03	0.07	-0.03	0.07
Factor H	**-0.05**	0.02	-2.13	0.04	-0.09	0.00	-0.09	0.00
Factor I	-0.02	0.02	-0.77	0.44	-0.07	0.03	-0.07	0.03
Factor J	0.01	0.03	0.23	0.82	-0.05	0.06	-0.05	0.06
Factor K	0.03	0.03	0.78	0.44	-0.04	0.09	-0.04	0.09
Choice Riders	-0.08	0.05	-1.53	0.13	-0.18	0.02	-0.18	0.02
Land Use Mix (Com, Res, Ind)	0.05	0.08	0.58	0.56	-0.11	0.20	-0.11	0.20
Average Base Adult Bus Fare	-0.03	0.05	-0.66	0.51	-0.13	0.07	-0.13	0.07
Average Base Adult Rail Fare	-0.09	0.06	-1.41	0.16	-0.22	0.04	-0.22	0.04

Table 9: Regression Analysis of Bus-Related Factors

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	0.00	0.03	0.07	0.94	-0.06	0.07	-0.06	0.07
Factor A	0.19	0.02	12.20	0.00	0.16	0.22	0.16	0.22
Factor B	0.01	0.02	0.39	0.70	-0.03	0.04	-0.03	0.04
Factor C	0.16	0.02	6.75	0.00	0.11	0.20	0.11	0.20
Factor D	-0.02	0.02	-0.95	0.34	-0.06	0.02	-0.06	0.02
Factor E	-0.02	0.02	-0.86	0.39	-0.06	0.02	-0.06	0.02
Factor F	-0.01	0.02	-0.54	0.59	-0.06	0.04	-0.06	0.04
Factor H	-0.02	0.01	-1.04	0.30	-0.05	0.01	-0.05	0.01
Factor I	-0.02	0.02	-1.33	0.19	-0.06	0.01	-0.06	0.01
Factor J	0.02	0.02	0.80	0.42	-0.03	0.06	-0.03	0.06
Factor K	-0.01	0.03	-0.24	0.81	-0.06	0.04	-0.06	0.04
Choice Riders	-0.07	0.04	-1.80	0.07	-0.15	0.01	-0.15	0.01
Land Use Mix (Com, Res, Ind)	0.02	0.06	0.40	0.69	-0.09	0.14	-0.09	0.14
Average Base Adult Bus Fare	-0.03	0.04	-0.90	0.37	-0.11	0.04	-0.11	0.04

Table 10: Regression Analysis of Rail-Related Factors

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	0.00	0.05	0.04	0.97	-0.09	0.10	-0.09	0.10
Factor A	0.16	0.02	7.05	0.00	0.11	0.20	0.11	0.20
Factor B	-0.01	0.02	-0.35	0.73	-0.06	0.04	-0.06	0.04
Factor C	0.24	0.03	7.14	0.00	0.17	0.30	0.17	0.30
Factor D	-0.07	0.03	-2.52	0.01	-0.13	-0.01	-0.13	-0.01
Factor E	-0.16	0.03	-4.69	0.00	-0.23	-0.09	-0.23	-0.09
Factor F	-0.02	0.04	-0.67	0.50	-0.10	0.05	-0.10	0.05
Factor G	-0.02	0.03	-0.61	0.54	-0.07	0.03	-0.07	0.03
Choice Riders	-0.08	0.06	-1.31	0.19	-0.19	0.04	-0.19	0.04
Land Use Mix (Com, Res, Ind)	0.08	0.09	0.90	0.37	-0.09	0.25	-0.09	0.25
Average Base Adult Rail Fare	-0.12	0.07	-1.67	0.10	-0.27	0.02	-0.27	0.02