2017

STRIDE | Southeastern Transportation Research, Innovation, Development and Education Center

Final Report

Using Crowdsourcing to Prioritize Bicycle Route Network Improvements

(Project # 2013-083)

Authors: Jeffrey J. LaMondia, Ph.D., Auburn; and Kari Watkins, Ph.D., Georgia Institute of Technology

May 2017

U.S. DOT DISCLAIMER AND ACKNOWLEDGMENT OF SPONSORSHIP

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

As the emphasis placed on cycling as a means of transportation is increasing in the United States, so is the need for adequate facilities that provide cyclists with a comfortable and connected facility. In order for these facilities to be built and encourage community residents to cycle, the city planners and engineers need to understand what type of facilities are appropriate and where they should be placed. This report uses data collected using the Strava, CycleDixie and CycleAtlanta crowdsourced cycling smartphone applications to determine factors that influence route choice. Specifically, these factors are studied through a) modeling cycling facility prioritization preferences, b) modeling cycling route segment and path choices, and c) developing route suitability score and preference models. This comprehensive research uniquely includes work from both suburban areas, represented by Auburn, AL and urban cores, represented by Atlanta, GA. From the analyses it was found that demographics, roadway characteristics and surrounding land-use had a significant impact on whether a particular street segment would be used.

CHAPTER 1: BACKGROUND

PROBLEM STATEMENT

Cities across the United States are becoming more interested in developing cycling infrastructure to foster sustainable livability, reduce traffic congestion, and improve the environment. It has been recognized that cycling can benefit communities by decreasing the amount of congestion on the roadways which not only decreases the air pollution in those communities but also cuts down on the gas consumption as well. While cutting down the vehicle emissions being issued into the air in communities, cycling also has a beneficial effect on the obesity rates in those areas by getting residents outside and exercising. It has been found that homes near bike trails have slightly higher home prices than those that don't have good access to cycling trails and facilities (Shinkle 2008). Recognizing the benefits of cycling on communities, the amount of federal funding and number of cycling projects has significantly increased over the past 20 year. In 1992 the number of cycling facility projects numbered only 50, with a funding of about $22.9 million. This has drastically increased to 2,485 projects totaling $820.5 million in federal funding for the year 2014 (FHWA 2015).

However, in order to promote the use of these facilities, it is critical to understand why cyclists choose to use specific routes. As such, route choice models based on finding suitable alternatives have become important measures. Building upon past research focused on modeling the choice of routes between the selected route and choice of alternatives, the main objective of this research is to model whether individual links within the road network will likely be used as part of commute cycling travel as well as identify the relative importance of the link characteristics on this this decision. Additionally, this work incorporates measures of land use access (e.g. for shopping, office, educational, etc.) to describe how connected (and relevant) each roadway link is

to the city. In this research it is hypothesized that the more connected a link is to the roadway network, the higher the likelihood that the link will be chosen as part of the cyclists' route.

Along with having links that are well connected to the roadway network, the links need to be designed in a way to encourage the use of cyclists and that those cyclists feel safe and comfortable on that link. An issue that often gets overlooked is which user group of the system the facilities should be designed for in order to encourage use of the facility. Some researchers suggest designing for all users, which allows them to not outright say which group should be the target design group (Bhat and Stinson 2005;Mekuria et al. 2012). The Vermont Pedestrian and Bicycle Facility Design Manual advises planners to design a facility for a "Design Cyclist", but also goes on to state that, "As a goal, a particular bicycle facility design should be chosen to encourage use by the lowest caliber bicyclist expected to frequently use the facility." (Vermont 2002). The only other definitive answer that was found was from the Federal Highway Administration, and states that "…DOT encourages transportation agencies to go beyond the minimum requirements, and proactively provide convenient, safe, and context-sensitive facilities that foster increased use by bicyclists and pedestrians of all ages and abilities, and utilize universal design characteristics when appropriate." (FHWA 2014)

While Vermont and FHWA chose to focus their design groups on the experience level of the cyclists, AASHTO chose to mention that design should be based on a number of purposes. In the *Guide for the Development of Bicycle Facilities*, it stated "… roads and pathways should be designed to facilitate various bicycle trip purposes." (AASHTO 2012) While this statement doesn't seem to suggest a group to design for, if the road or pathway is designed for various purposes, it will cover multiple groups of users as different groups will use a facility for varying purposes.

To model the route selection, an ordinal logistic regression model was used. The likelihood that a link was selected was based on roadway characteristics, connectivity to various access groups, and connectivity to various socio-demographic groups. The roadway characteristic variables were based on data obtained from the City of Auburn GIS databases. The access groups and socio-demographic groups were created using data from the U.S. census, utilizing the 2000 and 2010 census and American Community Survey, and the road network of the City of Auburn. The model also looked into the types of facilities present, and whether parallel facilities were present that could provide a better route alternative. Bicycle Level of Service was also considered in the analysis of the cyclists' route choice, with the links being rated an A-F.

RESEARCH OBJECTIVES

This project completes the following objectives:

- Collect crowdsourced cycling data through smartphone applications
- Compare and analyze measures describing the quality of bicycle facilities, in terms of level of service and level of stress
- Model the factors affecting bicycle route choice in urban and suburban areas
- Develop route suitability score and route preference models

SCOPE OF STUDY

Past research has studied whether individuals will commute via cycling (and the reasons for doing so) as well as individuals' preferences for different facility types (e.g. pathways, bike lanes, sharrows, etc.). However, less work has considered route choices as part of a larger network, and even less has completed choice models of commute cycling routes. This chapter summarizes

past work on cyclist classifications, design groups, data sources, influencing factors and choice models to inform the model developed in this report.

Roadway Factors Related to Cyclist Commute Routes

The majority of the factors considered in past route choice research attempt to describe the characteristics of the potential routes that cyclists choose among. The characteristics most often studied include travel time, continuity of bike facilities, number of traffic signals, and gradients (Bhat et al. 2005; Hood et al. 2011; Menghini et al. 2009; Fricker and Kang 2013; Aultman-Hall et al. 1997). From the previous research conducted, it was found that the continuity of the bike facilities had a positive impact on the likelihood of a route being selected, resulting in that route being used more often by cyclists. Due to the emphasis placed on continuity, the number of traffic signals had a negative impact as they caused the cyclists to have to stop before proceeding through an intersection. The travel time and roadway grade were also found by past researchers to cause the likelihood of a specific route being used to decrease due to the effort needed to traverse steep grades and the value placed on time. The perceived safety of the route, along with the adjacent land use was also studied in some of the past literature (Gliebe et al. 2009; Beheshtitabar et al. 2014). The route length (along with its relationship to the shortest path distance), travel times, and the steepness of the gradients along the route were found to have the greatest impact on route choices. (Bhat et al. 2005; Hood et al. 2011; Gliebe et al 2009; Krenn et al. 2014).

The majority of work aggregates or summarizes these roadway characteristics over the entire route, rather than consider variations across each link individually. This is most likely due to limited cyclist data records, where it is necessary to consider each route individually. When more detailed and widespread regional cycling trip data is available, such as this work, researchers are able to study whether each roadway link is important to the cycling network. For example,

Bhat et al. (2005) modeled link characteristics including roadway classification, presence of parallel parking, and pavement type and condition. Pavement type, whether the roadway was paved or unpaved, along with pavement condition were highlighted as important to cyclists, due to a bicycle not having the suspension capabilities of a car. Therefore, the cyclist will feel every bump and pothole in the road, and will favor roads that are smoother over roadways that are not paved or have not received adequate maintenance. Parallel parking was found to have a deterrent effect as the possibility of a cyclist being hit by an opening car door is increased as the number of cars parked along a stretch of roadway increases (Bhat et al. 2005). In another study, tied into roadway classification, the number of trucks and buses utilizing the roadway was found to have a negative impact on the number of cyclists willing to use a particular link as their perceived safety and quality of ride was diminished, suggesting that cyclists avoid busier roads in favor of roads with less vehicular traffic (Segadilha and Sanches 2014). A few researchers went on to look into cyclists' characteristics, built environment, and socio-demographics as well as the roadway characteristics (Bhat et al. 2009; Ma and Dill 2013; Urban et al. 2014). These researchers found that cyclists preferred routes that had continuous facilities, low amounts of on-street parking, lower speed limits, bike facilities present, and less cross-streets. The results also showed that travel time was important with shorter travel times preferred, especially in the 18 to 34 year old groups.

The Bicycle Level of Service (BLOS) was also used by a few researchers in order to see how suitable roadways were for cyclists (LaMondia and Moore 2015; Zolnik and Cromley 2007; Robinson et al. 2014). The BLOS "quantifies the perceived safety and comfort level of bicyclists on a shared roadway with respect to motor vehicle traffic" (Robinson et al. 2014). While the BLOS gives a rank from A through F of a roadway, that ranking can be used to determine which routes

are most likely to be used due to its perceived safety and the level of comfort that cyclists are likely to experience on that roadway.

Connectivity is another facet of cycling that needs to be considered in route choice. Past studies have looked into network connectivity by looking at how well the street network is connected, or the amount of street links connected to a node. The studies that looked into network connectivity did so based on Intersection Density, Link-Node Ratio, and the Road Type/Classification (Dill 2004; Hou et al. 2010). Intersection density is defined as the number of intersections per unit of area, with the higher the value the better as it assumes that the more intersections there are the more connected the road network is in that particular area. As it names suggests, the Link-Node Ratio measures connectivity based on the number of links, or roadway segments, in an analysis area to the number of nodes, or intersections, in that defined area. A higher number suggests better connectivity as there are more routes to choose from in the area due to the higher number of links to choose among.

The road functional classification also has a significant impact on the connectivity of a roadway and its appropriateness for cycling facilities. The highest classification is Arterial, which includes interstates and freeways. These roads have high mobility but to obtain this high degree of mobility these roads have low land access. The next classification of roads, collectors, relies on a balance of mobility and land access. The collectors link arterials to the final classification group of local roads. Local roads make up the majority of the roads in a community and provide the highest land access but also have the lowest mobility as they are generally designed to have lower speeds and are often found in neighborhood settings. While the street network being well connected is important, in order to give the cyclists multiple route options, it is also important that

the network be well connected to different types of areas that the cyclists may want to travel to, like shopping or office spaces for example.

Personal Factors Related to Cyclist Commute Routes

In addition to roadway characteristics, we can consider how individuals perceive these different components. A recurring technique for this is to break cyclists up into different categories based on how experienced the cyclist is and how comfortable that cyclist is with being in close proximity to vehicular traffic. Often times, researchers will also group cyclists based on their comfort level when traveling within traffic. A common typology of cyclists used in past research was the grouping of cyclists into the following categories: Strong and Fearless, Enthused and Confident, Interested but Concerned, and then finally No Way No How (Geller 2009; Dill and McNeil 2013; LaMondia and Moore 2015). While this method of grouping cyclists together provided some initial information, based off of the group name, of how the cyclists felt about participating in cycling, it does not necessarily group cyclists together based on how they use the road network.

Another common categorical system found allowed researchers to classify cyclists based on how they used bicycle facilities, grouping them into dedicated cyclists, path-using cyclists, fair-weather utilitarians, and leisure cyclists (Damant-Sirois et al. 2013). While these classifications focus on the way cyclists use the network and the perceived comfort level, Mekuria et al. uses the four category system to classify streets based on the amount of stress, traffic wise, each road presents. These traffic stress levels, when mapped, correspond to the common four cyclists groups in the above paragraph, with No Way No How corresponding to Level of Stress 1, Interested but Concerned corresponding to Level of Stress 2, Enthused and Confident to Level of Stress 3, and finally Strong and Fearless to Level of Stress 4 (Mekuria et al. 2012).

While the above classification schemes were developed by researchers in an attempt to better group similar cyclists together, Federal Highway Administration (FHWA) also published its own scheme, with it being simple to understand. The scheme developed by FHWA has three groups of cyclists, A: Advanced Cyclists, B: Basic Cyclists, and C: Children. While this classification is easy to understand, deciding whether a cyclist is an advanced cyclists or basic cyclists leaves room for subjectivity, and can make it difficult to form groups of similar riders. The American Association of State Highway and Transportation Officials (AASHTO), in their *Guide for the Development of Bicycle Facilities*, briefly mentioned that cyclists can often fall into two groups, Experienced and Confident or Casual and Less Confident. Not only does this classification scheme group cyclists into a group based on their experience, it also takes into account the cyclists' confidence level with cycling with traffic and other obstacles (AASHTO 2012).

Finally, Bhat et al. developed a three group system in their paper researching the preferences of bicycle commuters. Their classification took into account whether the cyclist was an experienced or inexperienced commuter and whether or not an individual was interested in commuting by bicycle (Bhat and Stinson 2005). This allowed the researchers to not only group the experienced individuals together, but also get a sense of how inexperienced users who are interested in commuting perceive the road network and what factors are keeping those that aren't interested in commuting from becoming interested in commuting by bicycle.

To further classify cyclists using road and bicycle facilities, researchers also gather socio-demographic information, including age, sex, education, access to motor vehicles, and health condition (Ma and Dill 2013; Urban et al. 2014; Poulas et al. 2015). The adjacent land use was also studied to see the effect that various land uses had on the frequency and type of trips being made. It was found that those living closer to a bicycle trail are more likely to cycle for recreation,

whereas those living closer to multiple trails increase their likelihood of commuting by bike (Urban et al. 2014). It was also found that high land-use mixing had a favorable impact on the likelihood of a route being used. On the negative side, it was found that areas with large amounts of high traffic areas, such as those areas containing restaurants and shopping, had a negative impact on the likelihood of a route being chosen, with cyclists avoiding those areas, most likely due to the increased presence of vehicles (Krenn et al. 2014).

Collecting Complete Regional Cycling Path Data

Until recently, the most common method of obtaining data on how cyclists were using cycling facilities was through the use of stated and revealed preference surveys (Hood et al. 2011). These surveys were conducted by phone, both land line and mobile, and through questionnaire surveys (Ma and Dill 2013; Yang and Mesbah 2013). This surveying method relies on not only people who have access to phones but who are also willing to complete the surveys and questionnaires. Another issue involved with this surveying method is the reliability of the information being reported, due to the respondent having to remember he routes that they chose and the characteristics of those routes, which can be tough depending on how far back the respondent is being asked to remember.

Alternatively, two methods for data collection have emerged as technology becomes more widespread and accessible. The first method is the use of web-based surveys. In many of these surveys, a list of individuals are emailed with a link to the survey, allowing for a large number of individuals to be contacted in the hopes of obtaining a larger sample size (Bhat et al. 2009;Poulos et al. 2015). These web-based surveys were interested in gaining an individual's preferences for a particular route, or interested in determining factors influencing bicycle usage (Sousa et al. 2014; Segadila and Sanches 2014; Krenn et al. 2014; Wang et al. 2014). While this surveying type is

effective for when a large number of individuals needs to be contacted, it relied on the response from those that had internet access and the time to complete the survey, often relying on individuals to remember the routes that were taken and other specific information pertaining to the route.

As the availability of smartphones and GPS has grown, many researchers have found the benefit of using GPS data to collect information on where individuals are choosing to cycle (Hood et al. 2011; Gliebe et al. 2009; Menghini et al. 2009; Seghadilla et al. 2014; Qing Shen et al. 2014). By using GPS, researchers can get coordinate data and map it in Geographic Information System (GIS) programs, such as ArcGIS provided by the company ESRI. The data collected can also be used to see what kind of facilities are being used and to see if cyclists are going out of their way to avoid certain areas or roads that are busy and have a high traffic volume. While GPS can give information about where the cyclists are choosing to travel, additional surveys are needed to obtain information about the cyclists and information about the roadway.

Crowdsourcing Data

Crowdsourcing has been alternately defined as: the outsourcing of a job (typically performed by a designated agent) to a large undefined group in the form of an open call (Howe 2006); a process that "enlists a crowd of humans to help solve a problem defined by the system owners" (Doan et al. 2011); or "a sourcing model in which organizations use predominantly advanced Internet technologies to harness the efforts of a virtual crowd to perform specific organizational tasks" (Saxton et al. 2013). Common across these alternate definitions is the notion that crowdsourcing invites all interested people to form an open forum of ideas that can eventually lead to a solution of the assigned problem. As Howe (2006) states, crowdsourcing utilizes the "latent potential of crowd" to achieve a solution to a problem that the crowd can relate to.

According to Saxton et al., crowdsourcing systems are characterized by three main features – the process of outsourcing the problem, the crowd, and a web-based platform for collaboration (Saxton et al. 2013). Outsourcing a problem generally implies getting a task done by outside sources even when it could have been performed by people within a system; in crowdsourcing, outsourcing is done in cases where either the in-house expertise has failed to produce a solution, or is an expensive means to produce a solution, or there is no in-house expertise available to use for solving the issue. Crowdsourcing systems also rely largely on an anonymous unidentified group of people ("the crowd") to come together willingly instead of the business sub-contract model of outsourcing where the task is performed by a previously identified and designated group of people or a company (Saxton 2013).

An important subset of the general crowdsourcing idea is the concept of citizen science, in which amateurs contribute to research projects in conjunction with the professional scientists. Goodchild used the term "citizen science" in describing crowdsourced geo-mapping, referring to the fact that information generated through crowdsourcing, although not of the level of a professional, helps in expanding the reach of science (Goodchild 2008). The nature of participation of the people in citizen science projects takes different forms depending on the type of the project and can range from data collection to data analysis, from instrument building to taking part in scientific expeditions. Recent citizen science projects tend to focus on utilizing the ever-increasing reach and availability of electronic gadgets, particularly mobile phones and sensors, for data collection and monitoring purposes. In their experiments, Kuznetsov and Paulos (2010) and Kuznetsov et al.(2011) provided citizen scientists with sensors to monitor air and environmental quality, while the CycleTrack project in San Francisco used GPS enabled mobile devices to record cyclist trip data (Hood et al. 2011). Citizen science projects are gaining

popularity as an alternative to cost intensive data collection efforts, particularly in cases where the information needed is global in character, and are thus being increasingly used for planning and monitoring purposes.

Despite the advantages, crowdsourcing can only be successful if a platform exists that can provide open access to incorporate, modify, and synthesize data. There are four different versions of this shared platform – the wiki system, open source software, geocrowd mapping, and mash-ups using crowdsourcing data (Kitchin and Dodge 2011). Wiki systems are mainly centered on authoring information; open source software provides a platform to share and co-develop program source code; geocrowd mapping entails collecting, cleaning, and uploading GPS data; and mash-ups are combinations of some or all of these. While maintaining coordination among people coming from different backgrounds and motivations is a significant challenge, this voluntary coming together of a mass of people for a purpose is particularly useful in tackling problems that are large scale, e.g., mapping of a country.

Steinfield et al. (2013) categorized public participation as either general purpose or domain specific systems. General purpose systems do not require any special expertise from the contributors and are not targeted to any user group in particular, while domain specific systems are designed for a special purpose user group. For example, most crowdsourced service quality feedback does not require any special expertise on the part of the participants and are hence, general purpose systems. Conversely, developing or beta-testing open source software through crowdsourcing requires expertise in particular programming languages and platforms and are hence, domain specific systems.

Crowdsourcing systems are further classified based on whether the system is local or global in scope and whether the system is time bound or not (Erickson 2010). For crowdsourcing systems

where the participants are at the same place at the same time, the system is termed as audience-centric (e.g., clickers used in class discussions). For systems where participants can be at different places but the crowdsourced event is time bound i.e., it has a start and end time between which the collaboration has to happen, such systems are termed as event-centric. An example of event-centric crowdsourcing is organized online brainstorming sessions triggered by an event and spanning over a limited period of time. Systems where collaboration can happen between people from different places and over an indefinite period of time are termed global crowdsourcing systems (e.g., Wikipedia). Finally, systems where people are at the same place but the crowdsourcing is an ongoing process are termed as geo-centric crowdsourcing – an example is bicycle route choice data collection for a city.

The characteristic of crowdsourcing that makes it suitable and useful for transportation planning is that it voluntarily brings together a large group of people on the same platform to address common issues that affect them. The use of crowdsourcing works successfully for local purposes through localized knowledge and acquired experiences (Brabham 2009) because people in a region tend to identify themselves with the region where they live, work, and socialize, and are generally more interested in the systems that affect them (Erickson 2010).

A survey of existing transportation systems which use crowdsourcing reveals that the predominant purposes of using crowdsourcing in these projects are either data or feedback collection from the users. For example, one popular use of crowdsourcing is in collecting route choice data from bicyclists using the GPS functionality of the user's cell phone – such data are not readily available through the standard data collection procedures and designing a separate survey for a small population of users is often not cost effective for regional planning agencies. Crowdsourcing in this case helps the geographically dispersed and diverse population of cyclists

to work together on a common interest without financially burdening the planning agencies. Similarly, crowdsourcing can also help in collecting feedback from a socio-demographically diverse range of users of any transit system that can be immensely useful for improving transit service quality and standards.

Transportation related crowdsourcing systems designed to date can be implicit or explicit standalone systems as defined by Doan et al. (2011) and discussed in the previous section. They may also be either geocentric systems where only local users are engaged or global systems where any person can contribute to the system. Extending the categorization of public participation as defined by Steinfield et al. (2013), transportation crowdsourcing systems may be further classified as either general purpose or domain specific systems. General purpose crowdsourcing systems do not require any special expertise from the contributors and are not targeted to any user group in particular, while domain specific systems are designed for a special purpose user group.

Modeling Where Cyclists Travel

To build a model to determine the most attractive route for cyclists, a few common methods were found in the past literature. The first method chosen by researchers was the Binary Logit Model (Bhat et al. 2005; Ma and Dill 2013; Urban et al. 2014). In two of the papers found using this method, the Binary Logit model was first used as a predictor of whether a cyclist would bike within a defined period, and then another model, such as a regression, was then employed to determine the frequency, based off a set of influences (Ma and Dill 2013; Urban et al. 2014). Bhat et al. (2005) used the binary logit model to estimate the impact of the studied variables on an individual's selection of a route.

Another common method found in the previous literature was the Multinomial Model (Hood et al. 2011; Bhat et al. 2009; Gliebe et al. 2009; Menghini et al. 2009; Akar and Clifton

2009; Ben-Akiva and Bierlaire 1999). These models were designed to determine the attractiveness of a route compared with a set of alternative routes not selected. Since the set of alternative routes can overlap on segments of the alternatives, the researchers had to overcome the correlation of the error terms by incorporating a similarity measure into the used utility functions. The most common similarity measure used was based off of the Path-Size measure presented by Ben-Akiva and Bierlaire (1999) (Hood et al. 2011; Gliebe et al. 2009; Menghini et al. 2009). The multinomial models were also used to determine the factors that influenced a person to cycle, as well as the selection of the route (Akar and Clifton 2009).

One of the key steps in the use of Multinomial models is the generation of choice sets to model the different route options available to the user. To generate the choice set of alternative routes, a few common methods were seen in the literature. The first method, discussed by Hood et al (2011), was the "doubly stochastic" method. In this method, both the link attributes and cost function coefficients were randomized for each search of the shortest path. To get accurate cost function parameters, the researchers developed the distributions that the coefficients were pulled from base on the road network. This methodology provided routes that were similar to those that were chosen, but bias and error can easily be introduced if the proper calibrations of the coefficient distributions are not performed.

While the above methods produced shortest paths for the inclusion in a choice set, these paths may not necessarily be completely unique. To overcome this limitation, the path-size factor was used to capture the similarity between the alternative shortest paths generated (Hood et al. 2011; Gliebe et al. 2009; Menghini et al. 2009).

Menghini et al. (2009) chose to use a broad search technique in their research to find the suitable alternative routes for the use in the choice set that they generated. To search for these

routes, they employed the use of the Multi-Agent Transport Simulation Toolkit (MATSim). The search was conducted using a certain detour threshold and a cost attribute of link length. To ensure that unique routes were found, overlap was controlled by the link elimination procedure in which up to three links were removed from a previously found shortest path. This correction factor slightly adjusts the utility placed on each of the shortest paths, which allows the researchers to avoid the use of more complex modeling techniques.

While the above models looked at modeling the route choice of cyclists by studying the route as a whole, some research has been done in modeling the route of a cyclist on the individual link, or segment, level. These link level models considered the route chosen by drivers as a sequential choice of links from the origin to the destination (Fosgerau et al. 2013, 2009). To determine the probability of choosing the next link of the route, the link level methods use the same modeling techniques as those that model whole routes, but do a sequential method, which allows for smaller set generations of alternatives, or in this case the next link. While these models were geared toward the study of driver behavior, these models are helpful to study for cyclists' route choice since the data provided for this paper was in the form of route segments and not full routes.

Summary vis-à-vis This Project

This research extends beyond the past work in a number of ways: First, it considers unique crowdsourced datasets to answer route choice questions. Second, it considers the questions of route choice at both the suburban and urban levels, which are recognized in past literature as being significantly different. Third, it models route choice based on a verity of variables, including accessibility and mobility along routes. Fourth, it incorporates ideas such as level of service and level of stress to help differentiate locations for facility improvement.

Report Organization

The report chapters highlight each of the study topics, and each chapter then describes the analyses and results from the suburban and urban contexts. Specifically, the following chapters discuss: a) colleting route choice cyclist records and supporting data, b) models of cycling facility prioritization preferences, c) models of cycling route segments and path choices, and d) models of route suitability scores and cyclist self-identification. The report concludes with thoughts on how the results from the two areas are related.

CHAPTER 2: COLLECTING ROUTE CHOICE CYCLIST RECORDS AND SUPPORTING DATA

This chapter discusses both what data was collected in each region to describe cyclist behavior and its surrounding land uses/roadways as well as the methods used to collect this data.

SUBURBAN APPLICATION ANALYSIS

Cycling Application Travel Data

The routing data used in the suburban component of the project was obtained from two different sources. First, data was collected from Strava, a technology company that developed a smartphone application that allows cyclists to record, via the GPS located in the phone, the routes that they cycle (Strava 2015). A screenshot of the application interface can be seen in Figure 1, which also shows some of the information that the app displays to the user after a route has been recorded. The application is available for use by any person who has a GPS device and access to the internet, with the majority of users comprised of cyclists and runners. As the cyclist uses the app, information such as duration, speed, elevation change, and distance are collected, along with the GPS route information. This allows the user to be able to look and see not only where they went but they can also analyze how well they performed and compare with other users. Second, data was collected using a smartphone app developed by the Auburn University research group, called CycleDixie (also seen in Figure 1). This app too worked for those with a smartphone that accessed the internet. Cyclists could swipe to start recording their trip, with their location recorded via GPS taken every 5 seconds, and then sent to the research team when the trip was completed.

The accuracy of the GPS data from both apps depends on the connection to the GPS satellites, with more satellites available the better the accuracy. Having an unobstructed signal to the satellites is also important to having high quality accuracy, with dense tree foliage and tall buildings obscuring and scattering the GPS signal. Fortunately, the accuracy in the suburban areas was of a higher level than the urban areas due to the lack of any urban canyon effect.

The research team worked with the local cycling community to collect the cycling travel data in 4 main ways. First, cyclists recorded trips using the CycleDixie app. Second, cyclists recorded new Strava trips using their existing app, and emailed to direct the team to download the trip from the Strava database. Third, cyclists provided access to their Strava accounts so the team could download existing trips directly. Fourth, the team purchased commute data directly from Strava for the Auburn, AL, region. The team worked closely with the local cycling community to evaluate the quality of the Strava data. While this data represents travel recorded from serious cyclists, it only included commute data with no recreational trips, and the cycling community felt it was representative of their travel patterns as well as those trips of cyclists who do not use the app. Overall, the research conducted in this report is one of the first to utilize the route data collected by Strava. The data from all 4 sources was formatted to roadway segments, resulting in a dataset that provided counts on roadway segments throughout the day.

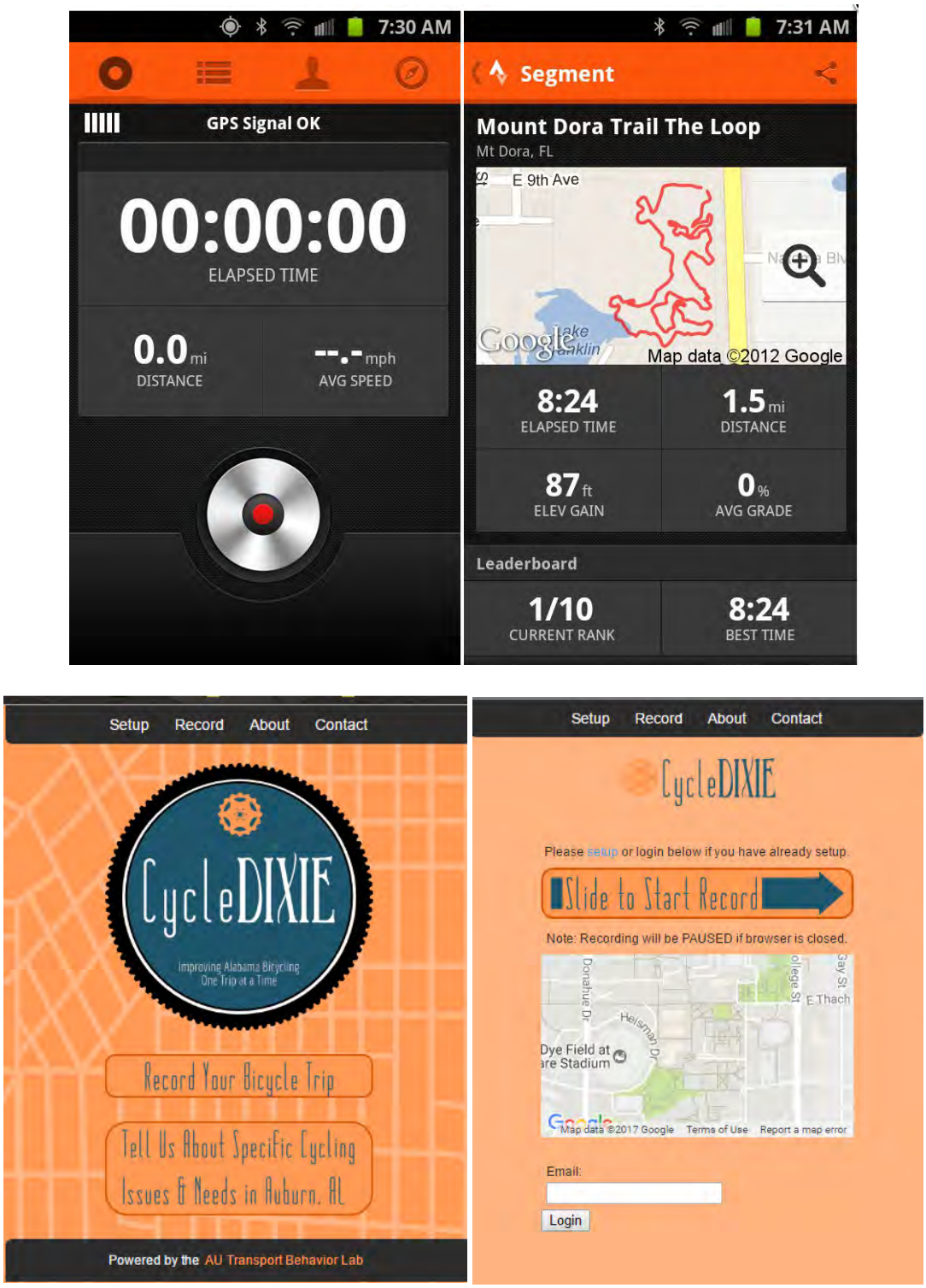


Figure 1: Strava & CycleDixie App Screenshots

Roadway Characteristics

To model the likelihood of a link being chosen as part of a cyclists' route, a number of roadway characteristics were considered. The variables included: speed limit, traffic volume (vehicles per hour, vph), pavement condition, presence of bike facility, width of outside lane, width of paved shoulder, number of driveways present, and whether medians were present. These variables were obtained from the City of Auburn, AL GIS database, and were attached to a particular link by assigning each link a unique identifying number. Additional information, including number of driveways, identified using Google Maps, roadway speed limits, and bike facility presence, determined from the City of Auburn Master Plan, were also collected.

The above variables were contained in multiple GIS layers, with many of the variables being their own separate layer. Using the unique identifier for each link, the road characteristic information for each of the above variables could be merged together creating a single GIS layer. Street links having majority of their associate information missing were removed from the dataset, as they provided no usable information. A total of 856 records were contained in this file, with one row of characteristics per street link.

Land Use Characteristics and Accessibility Measures

Along with the roadway characteristics that were considered for incorporation into the route choice model, land-use accessibility was also taken into account. The land-use variables that were considered were as follows: Shopping, Community, Educational, Governmental, Health Care, Mixed Development, Office Spaces, Parking, Residential, Restaurants. The information on where these particular land-uses are present in the City of Auburn, AL was also found using the City of Auburn's GIS database, utilizing the existing parcel ownership records layer.

To determine how well connected each roadway link in the city was to each of these land-uses, an accessibility measure was calculated. The form of the accessibility used can be seen below where A_i is the accessibility of link i to a particular land-use, x_z is the amount of land available for a particular a land-use in zone z, and d_{iz} is the average distance from link i to census zone z following the road network.

$$A_i = \sum_{n=1}^{z} x_z d_{iz}^{-1.5}$$

To calculate the distance between a roadway link and a census zone in Auburn, AL, the network analyst in ArcGIS was utilized. By setting the origins as the centroid of the road link, and the destination as the centroid of the census zone, an average distance, following the road network of Auburn, could be calculated for each origin/destination pair. The Auburn road network layer contained a total of 5,238 links, and the census layer contained 2,354 zones. The final dataset for this set of land-use information contained one row per street link with the corresponding calculated accessibility measures matched to each link by the link's unique identifier.

Regional Demographics

Similarly to the land-use variables, the accessibility to different socio-demographic groups was important to the model as well. Using U.S. census data, information concerning age, and household size was obtained from the 2010 census. Since the census information utilized was obtained from the American Community Survey (ACS), it is important to note that the ACS uses the definition of a household as: includes all people who occupy a housing unit as their usual place of residence (US Census 2015). This is important to note since the City of Auburn has a relatively high population of students, leading to some students being categorized as a household since a group of students may reside in the same residential unit. Utilizing the information from the 2000 census, commute time, income, and number of vehicles owned could be found for each census

zone in the City of Auburn, matching the census zone to the corresponding census block group to attach the census information collected the GIS layer containing the census zones. It is also important to note that since census data was used to gain demographic data, this data is not necessarily representative of Strava users, and that those using Strava may not be in the representative demographic groups for the City of Auburn.

The accessibility for each link to these socio-demographic groups was found using the same procedure as above, but using the demographic variables instead of the land-use variables for x_z. To use the information, care was taken to make sure that the census zone information matched the same zones used for the land-use calculations. The dataset for these set of variables also included one row per link with the associated accessibility measure for the socio-demographic groups, matched together using the links unique identifier.

Final Dataset and Geographical Distribution

The data obtained from Strava included an ID for each roadway segment, along with the number of cyclists, Strava users, which had traversed that roadway segment during the study period. Along with the number of cyclists who used the road segment, the number of activities, or number of one-way trips, for each roadway segment was also found in the dataset. The number of activities and cyclists per roadway segment were also listed for the peak morning and evening rush hours, as well as each direction of travel for the given road segment. For the scope of this research, the total number of cyclists per roadway segment over the 3 month period was used for the modeling process.

Since the Strava data was already processed by the Strava researchers, little cleaning was needed to be able to use the data. Screening was performed to verify that there were no abnormalities in the data provided, for example checking the roadway segments to make sure that

adjacent roadway segments had similar numbers of users and that there were no drastic differences in number of users between connecting segments, such as one segment having 3 users and the next having 30 users without there being a trip generator adjacent to those segments. The Excel file that contained all the weekday trips was saved as an SPSS file for the analysis to be performed quicker. The roadway segments were then given a usage rank based on the number of people using each roadway segment. Table 1 below shows the usage groups that were considered in the model developed later, with the groupings found using the natural breaks in the data. Along with Table 1 showing the Strava usage groupings, Figure 2 shows on which segments these groups chose to travel.

Table 1: Strava Usage Groups

Usage Group	Number of Cyclists
Low	0-13
Low-Average	14-34
Average	35-58
High-Average	59-93
High	94-157

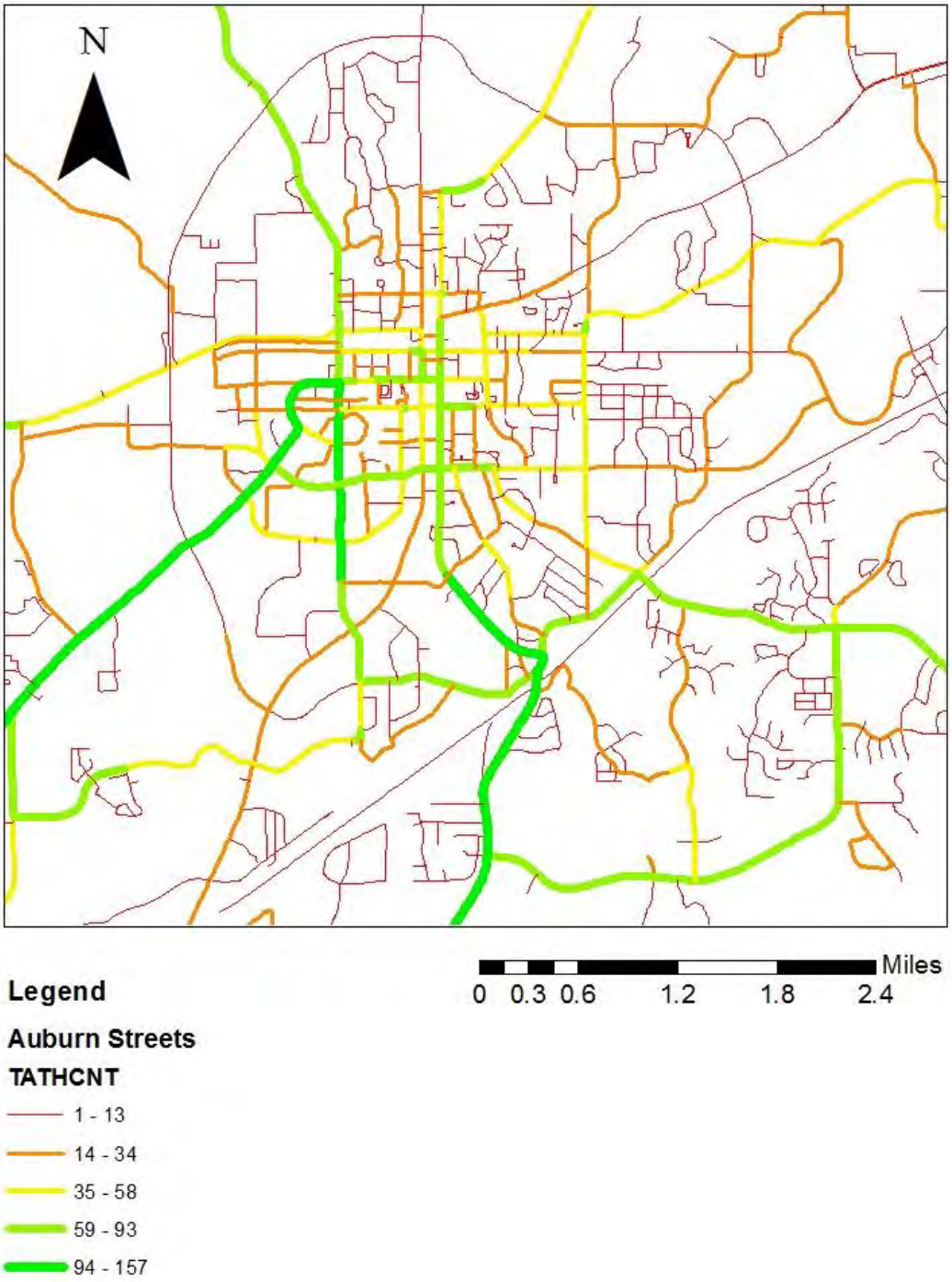


Figure 2: Strava User Counts per Roadway Segment

From the datasets provided by Strava, the number of bicycle trips taken in the Auburn area from January 2013 to December 2013 was a total of 5,201 trips recorded by Strava users. These trips were taken by 458 different cyclists. Looking at the number of trips per cyclists and taking an average, the average number of trips per cyclists was found to be about 11.4 trip/cyclists for the year 2013. The number of trips per cyclists per year seems low, but that is likely due to the majority of users recording only 1 to 5 trips during the year. The highest number of trips taken by a cyclist in this time period was found to be 377 trips. Figure 2 below shows the number of trips and the frequency of cyclists who cycled that many trips.

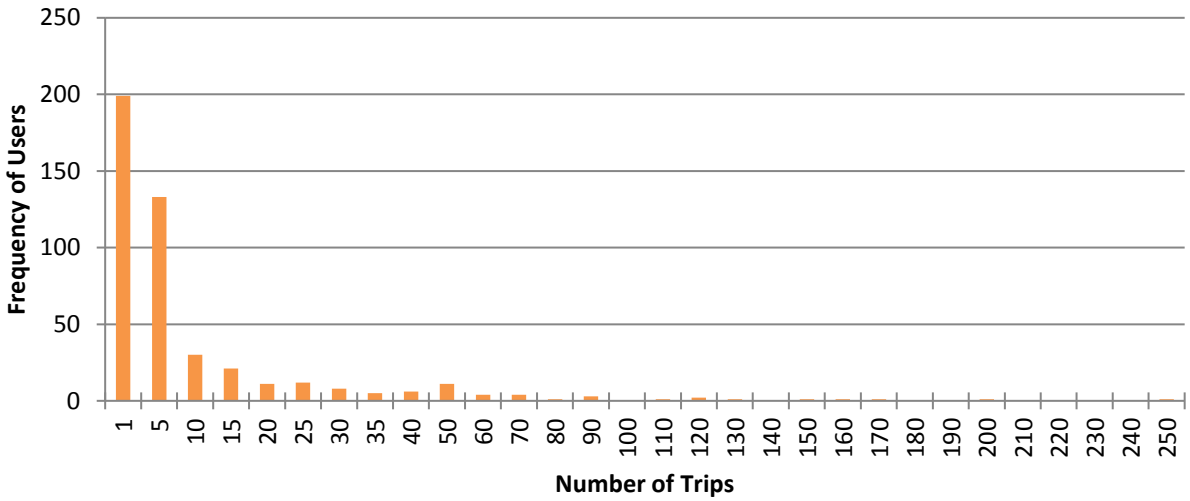


Figure 3: Trip Frequencies

The number of commute trips and non-commute trips could also be determined from the data provided. The number of commute trips was found to be low with only 887 trips of the total 5,201 trips taken being classified as a commute trip. This percentage breakdown can be seen in Figure 4. This percentage breakdown suggests that cyclists are more concerned about tracking their recreational trips and not their commute trips.

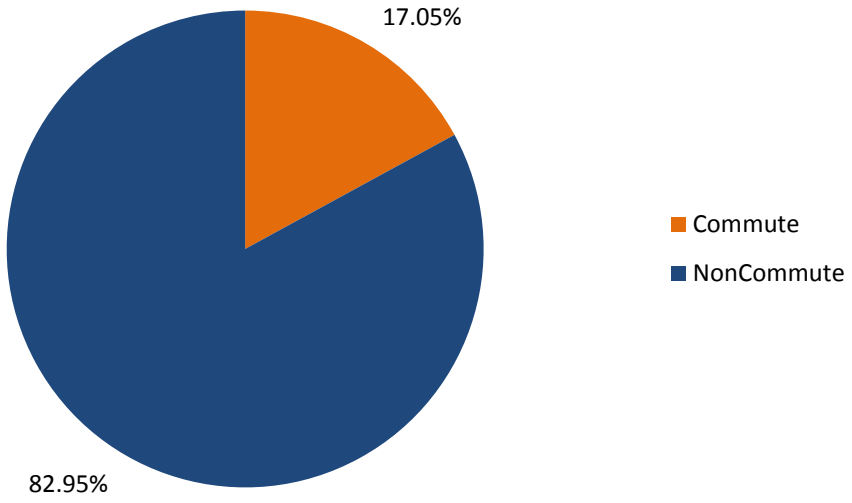


Figure 4: Percentage of Commute vs Non-Commute Trips

To obtain a final dataset for use in the modeling process, the four individual files were merged together, using each road links unique ID, to create the final data file. The final data set contained a record for each street link with the associated cyclist usage rank, roadway characteristics, land-use accessibility, and socio-demographic accessibility variables. A total of 856 links were in the final dataset which was used for the route choice modeling process.

URBAN APPLICATION ANALYSIS

Cycling Application Travel Data

The GPS data used for this urban component of the study were collected via the smartphone application Cycle Atlanta. Launch of the app in October 2012 was announced by the Mayor of the City of Atlanta and the app was widely publicized through various cycling advocacy groups and social media. Participation in using the app is voluntary and no reward was offered to record trips. The app is designed for both Android and iPhone GPS-enabled smartphones and is freely available for download from the app stores. The user has to turn on the app at the start of the trip and geolocation of the user is recorded from that point until the user indicates a trip end. The trip is not saved unless the person wants to save the trip which s/he can indicate via the 'save' button. At that point, the trip is uploaded to the secured database maintained by Georgia Tech. For each trip, the app records latitude, longitude, altitude, speed, time, and horizontal and vertical accuracy at an interval of 1 second. Figure 5 shows an example of the original uncleaned data from the Cycle Atlanta app.

(a) (b)

Figure 5: Original Uncleaned Data: (a) Raw GPS Points (b) Trip Lines Constructed from GPS points

Data Cleaning and Noise Filtering

The data issues found were classified as (1) systemic, (2) operational and (3) random. Systemic errors include issues that occur because of the use of GPS capability and are general in nature across all studies using GPS data. For example, cold and warm start problems, signal loss issues and urban canyon effects will be classified as systemic errors within our classification system. Operational errors are often errors introduced in the system by the users. These issues include forgetting to turn off the app after trip completion, using the app for non- cycling trips, using shortcuts and bypasses that are not part of the street network, etc. These errors will depend on the purpose of data collection and consequently on the participants. Random errors are most often related to systemic errors brought into the data due to use of GPS, but the nature of the errors are specific to each instant of recording and, hence, no standardized method can be applied to remove such errors.

The data cleaning was done following established practices from the literature. However, knowing the difficulty of map matching with noisy data, a lot of effort was put into early cleaning stages before running the snapping algorithms and therefore, the standard practices were modified and customized to suit our needs. Some additional criteria were also implemented keeping in mind the specific nature of the dataset. Efforts were made to attain a balance between retaining as much necessary information as possible in contrast to retaining data that is erroneous and can increase the computational burden for a later stage of analysis. It should also be noted that the app did not report the number of satellites, so that information could not be used for data processing in our case.

Operational Error Handling

As the study focused on bicyclists in Atlanta, at first, the data were checked for geographical limits – since the app is freely available to anyone owning a smartphone, it was suspected that the data might have trips that are not Atlanta based. Therefore, any point with latitude and longitude beyond the latitude and longitudinal boundaries of Atlanta [NW: 33.886823, -84.551068; SE: 33.647808, -84.28956] was removed from the dataset. Some trips were recorded over multiple days which can happen if the user forgets to turn off the app at the end of a trip and the app continues to record trips as continuation of the first trip until it is turned off. In such cases, the day with maximum number of recorded points was retained and data from other days were discarded.

Random Error Handling

Duplicate removal and basic data filtering: Two types of duplicates were identified: (1) points within same trip having same timestamp but different latitude and longitude and (2) identical latitude, longitude, timestamp and user id but different trip id. So, while in the first case, all points except the first point are removed, in the second case, the trip with the lower trip id is retained and the duplicates are removed. Some points were recorded with invalid timestamp (0000-00-00, 00:00:00) – these points were also removed during this step.

Horizontal Accuracy: As mentioned in NCHRP report and used in other research, the horizontal accuracy (haccuracy) threshold could be between 5 and 20 for a point to be a valid point. For this research, haccuracy limit was set to 30 – any point with horizontal accuracy more than 30 was removed from the database. The higher-than-standard limit was set after experimenting with haccuracy values of 10, 20 and 30. Since the data are from cyclists who tend to use bypasses, cut throughs and underpasses which do not always have a good signal, setting a higher accuracy

expectation resulted in removing too many points and created connectivity issues as well as sparse data problem for shorter trips.

Systemic Error Handling

Speed, Distance and Heading: The app recorded instantaneous speed at each point as well as latitude and longitude. Since the app is designed for cyclists, points with instantaneous speed more than 12 mph were discarded. Points with zero speed were further checked for distance and bearing from a point preceding 10 points upstream and the point succeeding 10 points downstream. If either distance or bearing change remained zero, the point was removed from the database.

Sparse Data: Some trips were found to have too few points for proper identification. The threshold ratio of distance to number of points was set such that speed between two consecutive points should not exceed 100 feet per second. If more than 50% of a trip consisted of points that did not match this criterion, the trip was discarded.

Noise Filtering: To filter points that are mainly signal jumps, a criterion similar to sparse data was used. If the distance from the point 10 steps before and/or 10 steps ahead of the point being checked is such that it cannot be traversed in the time between the timestamps at a speed of 70 feet per second, then that point is removed from the dataset. An additional check, if a large group of 10 or more points are major deviations, was used to remove any GPS point that was over 5,280 feet from the point that is 10 positions prior to it.

Data Reduction: The Cycle Atlanta dataset consists of about 15,000 trips, with each trip on average recording more than 1000 GPS points. One of the concerns was using such a large amount of data for map matching and our initial experiments of map matching in ArcGIS and R proved to be significantly slow and often problematic. Therefore, we decided to apply the Douglas-Peukar

algorithm to remove points for a trip that aren't necessary to identify its true shape and distance. The algorithm first identifies the starting and ending point. Then it finds the point in the line that is furthest perpendicularly from that line. If that distance from the point to the line is greater than the tolerance, then that point is kept and it remaps the "line" from the starting point to that furthest point. That new line then finds the point that is furthest from itself and does the same check. If it is within the tolerance, then that point is dropped and the algorithm checks for the next furthest point. It iterates over the whole line until all points have been checked. The tolerance used for our purposes was 5 feet, with the projection of the NAD83, UTM18 (North American Datum 1983, Universal Transverse Mercator, Zone 18). This means that any point that varied by more than 5 feet from the line between the points before and after, was kept, and any point that was under 5 feet was removed. This struck a good balance between ensuring that much of the route shape was kept while limiting the number of points needed. In addition, it ensured that for snapping purposes, no streets were skipped that were clearly traveled on. For a street to be snapped, there had to be a point near it. Therefore, reducing the number of points with too large of a tolerance would have resulted in long straight segments of a path with no points kept. The 5 foot tolerance allows for enough precision while clearly reducing the number of points required.

With this simplified line, we can then interpolate the points in it for snapping purposes and determine whether there are any path duplicates. The function ST_DumpPoints in PostGIS takes the simplified line and returns the points of that line, thus reducing the number of points to snap from roughly 15 million to 2 million.

Final Dataset and Geographical Distribution

Users who indicated that they lived outside of the Atlanta metropolitan area were purged from the database. This was done by sorting the table of users by the home zip

code they reported and deleting the records that contained zip codes outside of the Atlanta area.

For geographic analysis using ArcGIS, a shapefile of Atlanta zip codes was obtained from the Atlanta Regional Commission (ARC). However, ARC's zip code shapefile did not contain all of the zip codes reported by Cycle Atlanta users. For example, the zip code 30332, which contains part of Georgia Tech's campus, was not part of the ARC zip code shapefile. To rectify this, missing zip codes were drawn into the shapefile using Google Maps and a shapefile of city streets for guidance. The chosen study area comprised of zip codes located either completely or partially within Atlanta city limits and/or the Perimeter (I-285), as shown by the red shading in Figure 6.

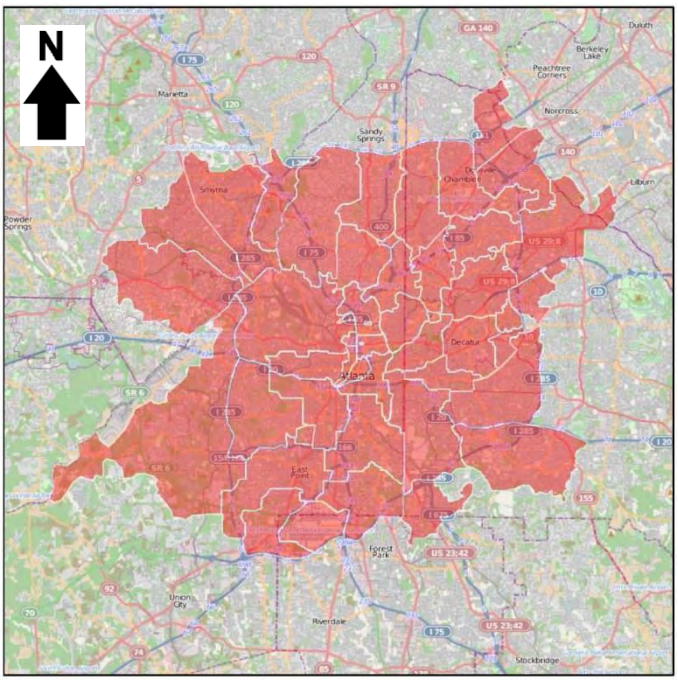


Figure 6: Zip Codes completely or partially within Perimeter (I-285) or City of Atlanta limits

The Cycle Atlanta datasets were queried to return a list of each discrete home zip code in that dataset as well as how many people in the dataset reported that zip code as their home

zip code. The datasets were then joined to the study area zip code shapefile using home zip code as the common field. This resulted in a data table containing fields for home zip code and number of Cycle Atlanta users living in the zip code.

A map was created in ArcGIS to show the percent of cyclists within the study area who reported each zip code as the one they resided in. The map was shaded in such a way that darker zip codes had a greater percentage of the dataset's cyclists (within the study area). For context, an OpenStreetMap basemap was added to each map.

To analyze the relationships between Cycle Atlanta user home zip codes and demographic traits associated with those zip codes, four maps were generated using census data. For each map, the shade of the zip code polygon represents the demographic variable (zip code median age, median annual income, percent of non-white residents, and population density). The size of the black dot over a zip code represents the percent of Cycle Atlanta users residing there. The median age and percent non-white data were obtained from American Community Survey table DP05, "Demographic and Housing, 2007-2011 5-Year Estimates". The median income data were obtained from American Community Survey table S1903, "Median Income in the Past Twelve Months (In 2011 Inflation-Adjusted Dollars), 2007-2011 5-Year Estimates". The population density data were obtained from American Community Survey table B01003, "Total population, 2007-2011 American Community Survey 5-Year Estimates".

Figure 7 shows that cyclists are concentrated in the "intown" part of Atlanta, near the center of the Perimeter. Specifically, zip codes east of the Downtown Connector (the north-south running Interstate near the center of the study area) have the highest percentages of cyclists living within them.

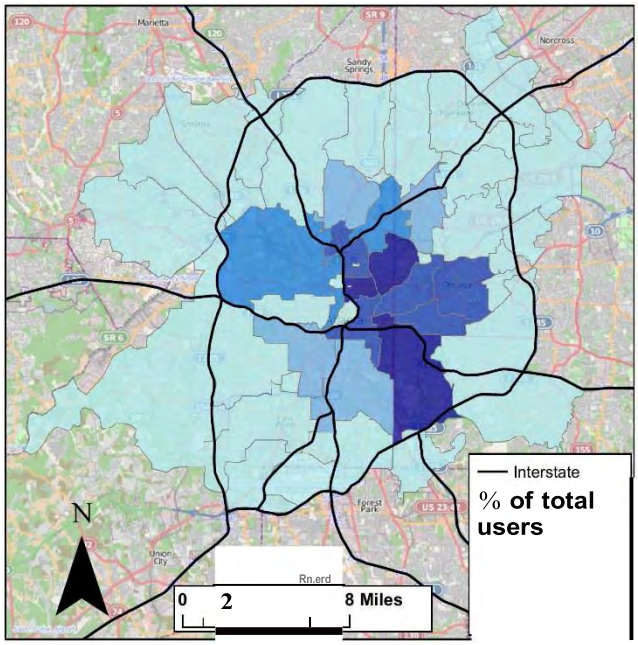


Figure 7: Distribution of Cycle Atlanta Users by Home Zip Code

Spatial Correlation

This section examines the correlation between the percent of Cycle Atlanta users residing in a zip code and several demographic characteristics of the overall population in that zip code – median income, median age, non-white population, and population density.

Figure 8 shows a comparison between the percent of Cycle Atlanta users living in a zip code and the percent of non-white, non-Hispanic residents living in that zip code. The darker the zip code, the greater the percentage of non-white residents; the bigger the black dot over a zip code, the higher the percentage of Cycle Atlanta users living there. It is difficult to see a clear relationship between the two variables. Some zip codes have a low percentage of non-white residents and a high percent of Cycle Atlanta users living there such as 30306 and 30307 (located between E4 and E5). However, some zip codes have a high percentage of non-white residents and a high percentage of Cycle Atlanta users, such

as 30316 (located between E5 and E6).

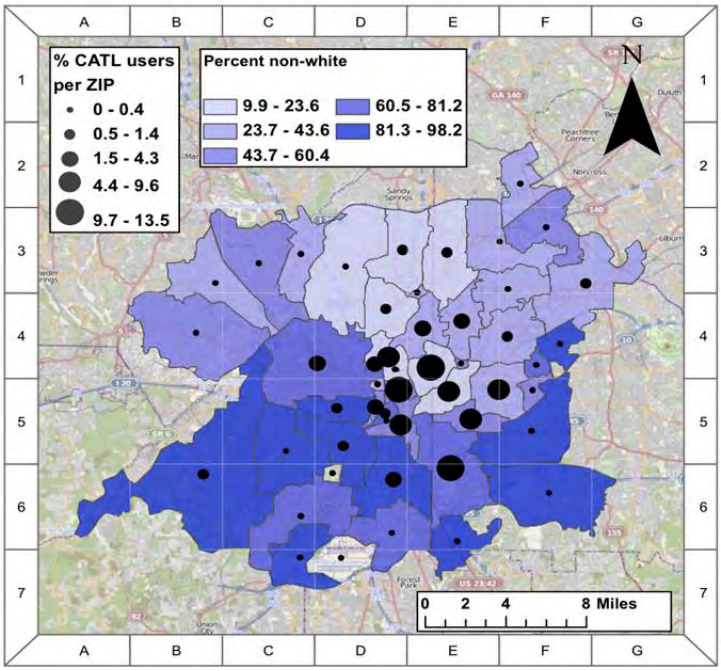


Figure 8: Cycle Atlanta Users Home Zip Code Distribution across Ethnicity Distribution in Atlanta

Figure 9 shows a comparison between the percent of Cycle Atlanta users living in a zip code and the median income of households in the zip code. The darker the zip code, the greater the median incomes of households there; the bigger the black dot over a zip code, the higher the percentage of Cycle Atlanta users living there. Although a high percentage of Cycle Atlanta users are from the high income group (greater than $100,000), that is not reflected in the geographical representation. Zip code 30327, for example, has the highest median income of any zip code (between $100,000 and $130,270, the income group that had the greatest number of Cycle Atlanta users in it). However, 30327 also has one of the lowest percentages of Cycle Atlanta users residing in it, at less than 0.353.

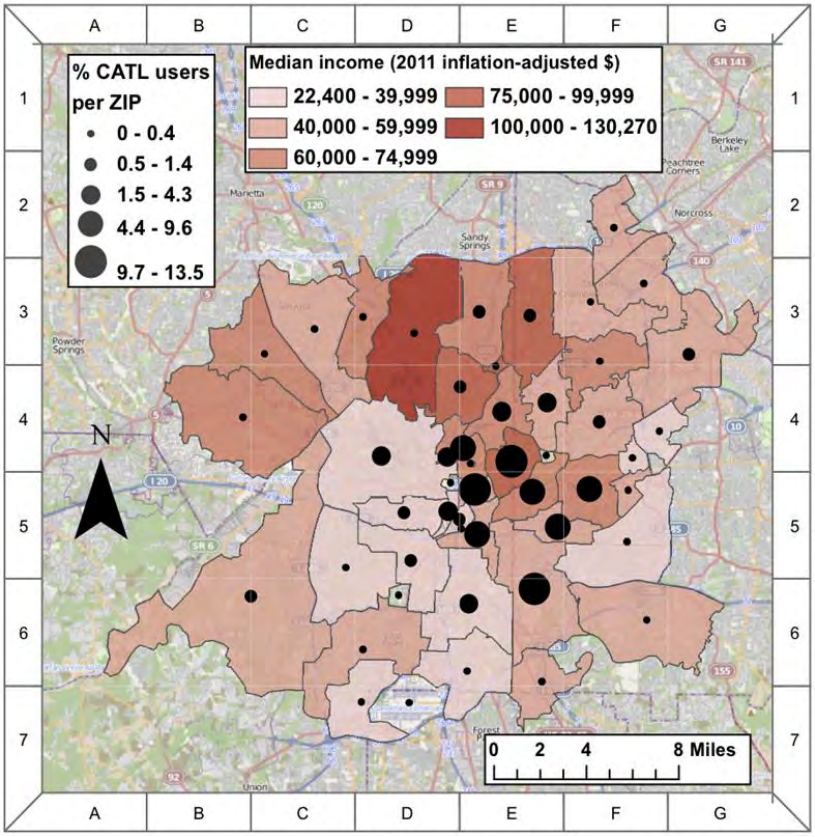


Figure 9: Cycle Atlanta Home Zip Code Distribution across Median Household Income Distribution in Atlanta

Figure 10 shows a comparison between the percent of Cycle Atlanta users living in a zip code and the median age of people living in that zip code. The darker the zip code, the greater the median age of people living there; the bigger the black dot over a zip code, the higher the percentage of Cycle Atlanta users living there. The researchers could expect zip codes with median ages between 25 and 34 to have the greatest percentage of Cycle Atlanta users residing in them, since this was the age category with the greatest percent of Cycle Atlanta users. While this is somewhat true, it appears that zip codes with median ages between 35 and 44 have greater percentages of Cycle Atlanta users living in them than zip codes with median ages between 25 and 34 (which is the age category that has the second highest percentage of Cycle Atlanta users in it).

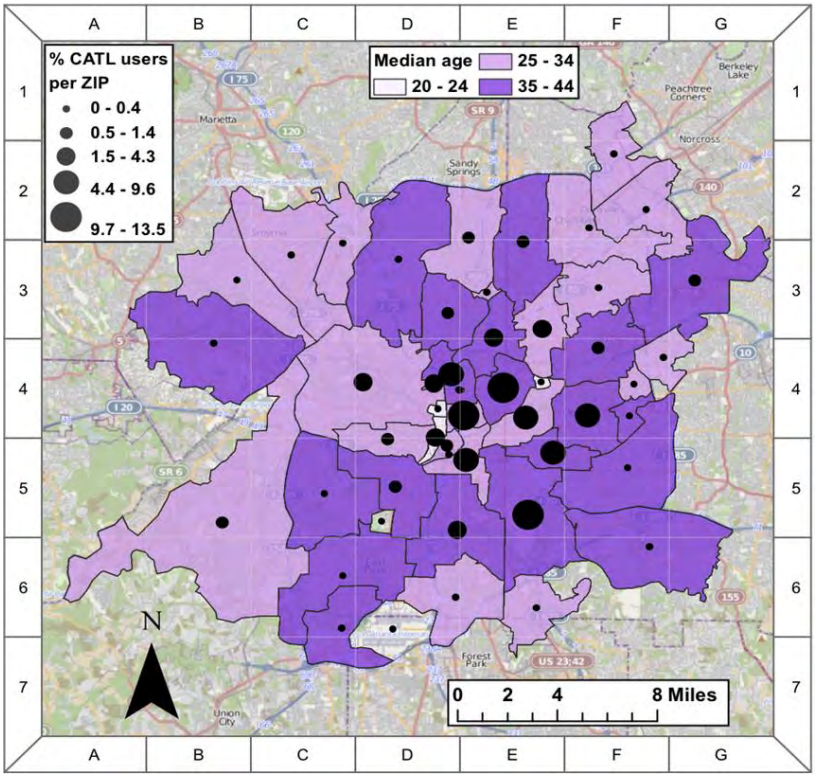


Figure 10: Cycle Atlanta Users Home Zip Code Distribution across Median Age Distribution in Atlanta

Figure 11 shows a comparison between the percent of Cycle Atlanta users living in a zip code and the population density of that zip code. The darker the zip code, the greater the population density; the bigger the black dot over a zip code, the higher the percentage of Cycle Atlanta users living there. Aside from a few outliers, this map suggests that as the population density of a zip code increases, so does the percentage of Cycle Atlanta users living in that zip code. This makes sense, since high-density urban areas are often the most bikeable. One note-worthy outlier is the zip code 30316, located between E5 and E6. This zip code contains dense areas such as East Atlanta Village and Reynoldstown in the northern part, but also less dense areas such as Gresham Park in the southern part. It is likely that if this zip code were separated into a north part and a south part, the north part would show high density as well as a high percentage of Cycle Atlanta users residing in it, and the south part would show low density and a low percentage of Cycle Atlanta users.

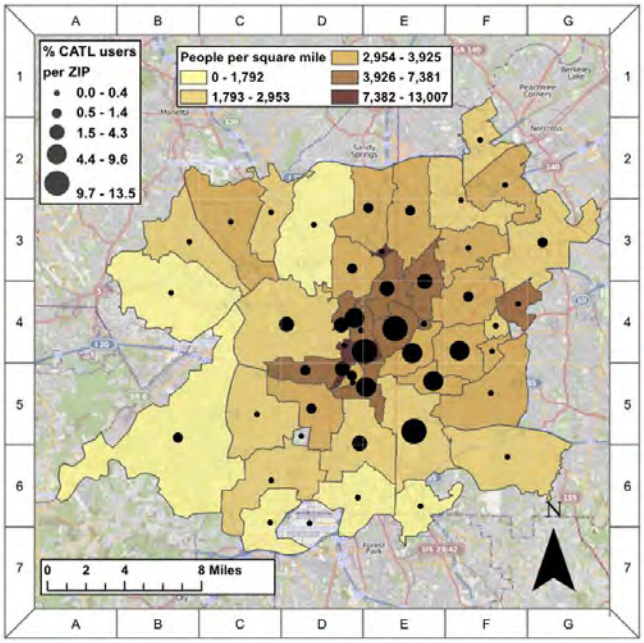


Figure 11: Cycle Atlanta Users Home Zip Code Distribution across Population Density Distribution in Atlanta

CHAPTER 3: MODELING CYCLING FACILITY PRIORITIZATION PREFERENCES

This chapter studies how roadways environments are evaluated and quantified in terms of support for cyclists. These analyses are critical because these measures are used frequently to evaluate how well a roadway supports cycling as well as whether improvements are needed.

SUBURBAN APPLICATION ANALYSIS

Common Types of Bicycle Level of Service Measures

Despite the fact that many bicycle LOS measures exist, most rank roadway segments in terms of bicycle safety and/or comfort based on the same core set of roadway characteristics. These core roadway characteristics include: traffic volumes, roadway widths, number of lanes, vehicle speeds, and presence of bicycle facilities. As one would expect, increased volumes, smaller lanes, higher speeds and lack of facilities all lead to lower LOS scores (Yamaka and Namerikawa, 2007, Lowry et al. 2003, Jones and Carlson, 2003, Landis et al, 1997, Kang and Lee, 2012, Li et al., 2012). Some less prominent, but still important, factors include on-street parking, access/egress points, and pavement quality (Landis et al, 1997, Kang and Lee, 2012). These vary in the impact on bicycle level of service: for example, some measures score the presence of on-street parking as a positive for the cycling environment (by helping delineate spaces) whereas others score it as a negative (due to the chance of getting "doored") (Lowry et al. 2003, Jensen, 2007).

Data collection used to generate the ranking scales for these different methods may be conducted in a number of ways as well. This may influence a) what factors are considered and b)

how individuals respond to the factors. Stated preference surveys for roadways done through paper, video, web-based, observation, and simulations (Jones and Carlson, 2003, 17, Landis et al, 1997, Li et al., 2012, Parkin et al., 2007). Each reaches a different group of respondents (e.g. cyclists under observation versus non-cyclists via paper surveys) and gauges reactions differently (e.g. real time responses in simulations versus memory in paper surveys).

Still, the most important difference between LOS measures occurs in *how* they weight these factors in the scoring method. This has much to do with how each researcher chooses to interpret cyclists' perception of comfort and safety, regardless of the data collection technique. In fact, Lowry *et al.* recognized that there are different ways to interpret cycling comfort and safety (Lowry et al. 2003). They identified a number of definitions, including "bicycle suitability", "bikeability", and "bicycle friendliness", in addition to "bicycle level of service". Additionally, as transportation systems shifts from dense urban areas to suburban areas, the impacts that certain factors have on perceived cycling environment may also change (Jones and Carlson, 2003, Parkin et al., 2007). Four common types of bicycle level of service measures have been identified from the literature: index measures, applied measures, point measures, and scaled measures. A summary of each type, highlighted examples, and a representative method (that will be calculated for roadway segments in this study) are discussed below as a means to describe these different approaches. Representative methods were selected to be as similar to each other in terms of input factors as possible.

Index Measures: Bicycle Compatibility Index (Harkey et al, 1998)

Index LOS measures weight roadway characteristics in a simple regression equation and then groups are scores are assigned LOS values A through F (See: Jones and Carlson, 2003, Jensen, 2007, Transportation Research Board, 2010). These measures do not add complexity beyond

simple weighting of bicycle facility factors. The Bicycle Compatibility Index was selected to represent the index measure category. It calculates a comfort level score for cyclists (A-F) based on presence of a bike facility, lane widths, traffic volume, speed and other factors.

Interaction Measures: Bicycle Applied Model (Landis et al, 1997***)***

Interaction LOS measures are similar to indices but use more advanced equations, such as ordered probit or non-linear regression models to allow for complex interactions between factors (See: Yamaka and Namerikawa, 2007, Kang and Lee, 2012, Parkin et al., 2007). Groups of equation values are still assigned LOS values A through F. The Bicycle Applied Model was selected to represent the interaction measure category. It is based on user responses to roadway and traffic scenarios (scored 0 to 5.5, proportioned out to A through F), and considers thirty-four different factors, including volumes, lane widths, speeds, land uses and others. This method weighted heavy vehicles significantly in the calculation, but this data was assumed to not be a significant factor for the region. As such, a modified version of this measure with that factor removed was also considered in the analysis.

Point Measures: Bicycle Scoring System (Dixon, 1996***)***

Point LOS measures award points to a segment based on how different factors are represented. Then points are added up to a final score, which is tied to assigned LOS values A through F (See: Ophartd, 2005, Charlotte Department of Transportation, 2014). In several methods, some factors earn more points than others to weight that factor as more important, but most use a similar point scale for each factor. The Bicycle Scoring System was selected to represent the point measure category. It measures quality of the bicycle accommodations on a point

scale of 0 to 21, which correspond to grades of A through F. This measure also considers the presence of bicycle facilities, conflict points, speed differentials and perceptions of maintenance/roadway quality.

Scaled Measures: Multimodal Measure (Mozer, 2014***)***

Scaled LOS measures are similar to the point measures, where each factor is awarded points based on how well it performs. However, rather than sum the points this method averages the points, making each factor weighted identically (See: Kentucky Transportation Cabinet, 2014). While the point measure allows one high-performing factor to "make-up" for another deficient factor, the scaled measure takes an average across the different factors (allowing one factor to skew the final LOS value). The Multimodal Measure was selected to represent the scaled measure category. This method assigns an A through F score for common factors separately like lane widths, volumes, speeds, on-street parking and others. These scores are then averaged to calculate a final LOS for the segment.

Calculating Bicycle Level of Service Measures

Cycling Infrastructure Data

The four common Bicycle LOS measures require a wide range of data to be collected, which includes roadway volumes, lane widths, speeds, bike lane widths, land uses, maintenance issues and pavement conditions. The city of Auburn was selected as a case study for comparison due to its size (over 58,000 residents within 39.6 square-miles), its interest in cycling (earning a bronze designation from the league of American cyclists) and its younger, active community (81% of the population is 45 years or younger) (US Census Bureau, 2003).

To complete the comparisons, the majority of the roadway characteristics and traffic information were collected from the City of Auburn, AL, and formatted within an ArcGIS spatial database. Supplemental geometric data, such as the access and egress points along segments, on street parking, lane widths and shoulder widths, was calculated from Google Earth aerial photographs. Environmental factors, such as inadequate sight distances, bikeway barriers, pavement conditions and maintenance issues were gathered by cyclist volunteers within the Auburn Bicycle Committee.

Some assumptions were made when the segment-level detailed data was not available. For example, traffic counts were given in a point format, so volumes were averaged over entire roadways and assumed to be the same for all segments of that roadway. Also, in order to address the issue of two-way travel, LOS for each segment was calculated for the side of the road that would cause that segment to be rated lower. Pavement conditions were assumed to be a level 4 out of 5 on the ranking scale, unless otherwise noted by a volunteer. On-street parking time limits were fixed, as per city guidelines, at 2 hours. Also, while heavy vehicles are present in many of these measure calculations, it was assumed that, due to the minimum presence of these vehicles that the percentage was near zero.

Working with the City of Auburn, the research team selected 565 segments representing 66 roads (of over 95 miles) that covered the core cycling routes of Auburn. These road segments did not include local or minor residential roadways nor did they include segments on the outskirts of the city. Of the 565 segments, the mean length was about 770 ft. and the maximum and minimum lengths were 5539 ft. and 26 ft., respectively. A bicycle facility was available on 24.5% of the segments studied. The mean vehicular traffic volume on the roads in the study was about

4900 ADT with a maximum and minimum count of over 18,000 ADT and 117 ADT, respectively. Most arterials and collectors within the city have similar geometric design and volumes.

Finally, the equations and procedures for each of the LOS measures were applied to each of the 565 segments. As a result, each segment received five LOS measure scores from "A" through "E/F". One can see the spatial distribution of these scores for the roadway segments in Figure 12.

Perceived Bike Route Suitability Data

To supplement the common bicycle LOS measures, additional surveys of perceived bike route suitability were collected. This survey attempted to collect information on what the LOS measures were trying to calculate: where cyclists of different types think are suitable places to bicycle. If answered honestly, these are the places that cyclists currently think are the routes that should be used to travel throughout the city via bicycle.

The survey was disseminated to participants of the Auburn Bicycle Committee and the East Alabama Cycling Club as well as administered in two locations in downtown Auburn with the busiest foot-traffic from 9am to 3pm on multiple weekdays, across multiple weeks (the main downtown intersection in Auburn and the main concourse on the campus of Auburn University). Respondents included students, faculty, staff and other city residents and the research team worked diligently to explain to participants that they did not need to a cyclist to answer. Some surveys were returned incomplete or with obvious indications that they were haphazardly answered, which were discarded. A total of 565 complete surveys were collected and included in the analysis.

To improve response quality, the survey focused on two questions and a map of the city. Participants were first asked how often they rode a bicycle ("everyday", "a few times a week, "a few times a month", and "not often/never") and how they would classify themselves ("Strong and

Fearless", "Enthused and Confident", "Cautious but Comfortable", "Interested but Concerned", and "No Way No How"). These categorical questions were designed to mimic the cyclist types found in the literature, and an additional category was added during pilot testing when participants requested a "middle category". The sample included 10.3% who cycle every day, 17.5% who cycle a few times a week, 23.8% who cycle a few times a month and 48.4% who cycle rarely. The respondents also classified themselves as 8.8% who are "Strong and Fearless", 25.8% who are "Enthused and Confident", 24.6% who are "Cautious but Comfortable", 25.3% who are "Interested but Concerned", and 15.5% who are "no Way No How".

Respondents then highlighted or circled those cycling routes within the city that they felt were suitable for cycling, regardless if they used them or not. In pilot testing, the term "suitable for cycling" garnered the same response from both cyclists and non-cyclists, which was that it was appropriate and acceptable to bike on. In most cases, it was clear individuals were selecting long routes going between places that they had either biked, had seen others bike, or just thought was reasonable to bike. Results from the surveys were summed in ArcGIS by segment and by cyclist types. Segments were then scored in a similar fashion to the LOS measures, and for each category of cyclists, natural breaks were used to separate the segments into five naturally-forming categories, labeled "A" through "E/F". The total (i.e. all cyclist types) spatial distribution of these scores can be seen in Figure 12.

While the sample is well distributed across the various cyclist types, it is important that the sample is biased towards the campus community due to the collection technique. However, more than 80% of students live off-campus and more than 55% of all city residents are associated with the university, so it is a reasonable assumption that respondents are familiar with the roadway environment both on- and off-campus.

Comparing Common Bicycle Level of Service Measures

Once each of the representative common bicycle level of service measures was calculated and scored for each segment within Auburn, AL, they were plotted, as seen in Figure 12. In these diagrams, the major streets the city of Auburn consider part of the cycling network are colored to show the score it received: the darkest color represents LOS A and the lightest color represents LOS E/F. The urban core of the city is located in the center of each map, where the central north/south and east/west streets intersect. Most of the population lives within the ring road around the city, and the university is notated with a cross-hatch shading in the southwest quadrant of the map. Visually comparing the roadways LOS measures, one can see that the roadways on the outskirts of town with less traffic volume are highlighted with the best LOS values for all four methods. Across the four common LOS measures, the scores emphasize disjointed pieces of the network that are/ are not adequate rather than entire routes, which makes it difficult to identify areas that need improvements. Further, the areas in which cyclists are using most often, as seen in the bottom heatmap of recorded cyclist trips (*30*), do not follow a strong relationship with any of the segment LOS values. In fact, many of the most traveled segments have low LOS scores in all four maps.

Figure 12: Variations in Cycling Level of Service across Measures

Table 2 defines the correlation between the LOS measures and roadway characteristics. The results show a somewhat strong correlation between most pairs of LOS measures, indicating that they are roughly ranking segments similarly. The strongest relationship is between the index, point, and scaled methods. These do not add in additional interactions between factors, but do include different factors in their calculations. The interaction measures have the weakest correlation with the other three techniques (but are correlated with each other despite the modification of the heavy vehicle parameter).

Table 2: Correlations with Common Bicycle Level of Service Measures

		Common Bicycle LOS Measures				
		Index	Interaction	Interaction Modified	Point	Scaled
Common Bicycle LOS Measures	Index	1.000	-	-	-	-
	Interaction	0.510	1.000	-	-	-
	Interaction Modified	0.491	0.901	1.000	-	-
	Point	0.646	0.406	0.383	1.000	-
	Scaled	0.772	0.525	0.562	0.712	1.000
Roadway Characteristics	Total Volume	0.655	0.656	0.654	0.352	0.579
	Total # Of Lanes	0.397	0.215	0.244	0.290	0.378
	Speed Limit	0.436	0.354	0.444	0.307	0.387
	Bike Facility Identifier	-0.408	0.289	0.304	-0.338	-0.318
	Width Of On-Street Parking	0.141	-0.122	-0.041	0.044	0.028
	Curb Lane Width	-0.383	-0.582	-0.641	-0.441	-0.613
	Access/Egress	-0.018	0.001	-0.008	0.180	0.067

The table also compares each measure with different critical roadway characteristics. The most significant correlations are with traffic volumes (except for the point method). Interestingly, the interaction measure is highly correlated with the curb lane width, despite all the measures using this factor. Perhaps this is closely tied with other factors in this model. Otherwise, there is not a

strong correlation with any specific characteristic, suggesting that the combinations of these factors are more significant than any one.

Beyond correlations and spatial relationships, Figure 13 presents the distribution of LOS scores over segments for each method. The charts show that the distributions are rather different: the index and point methods feature a normal distribution, but lack many segments with LOS A scores. The interaction measure over-identified segments with LOS A while the modified version addressed this issue, but both feature a majority of segments with LOS D. The scaled measure was the most pessimistic approach, ranking the most segments at the E/F LOS.

There are two major conclusions from these comparisons. First, index and point bicycle LOS measures may be interchangeable, but the others are not. Each over-represents certain LOS values and interprets the importance of factors differently. The interaction measures are focused on complex sets of factors influencing cycling and the scaled measures flip to allow a single factor to dominate the LOS score. Second, traffic volumes are critical to LOS, but no one other factor dominates the bicycle LOS scores. Therefore, it is worthwhile to have equations and methods that incorporate these sets of variables.

Comparing Roadway Characteristics and Route Suitability

While comparing common LOS measures is important, it is also useful to compare these measures with the cyclists' reality. In this study, this perspective was defined by asking different types of cyclists (defined by frequency of cycling and confidence level) to identify roadway segments suitable for cycling. As noted before, a comparable A through F score was created for each roadway segment based on natural breaks in the total number of cyclists (of each type) that identified a segment was suitable for cycling.

Interestingly, when the segment scores for graphed for each type of cyclist (9 in total) were plotted, they were nearly identical to the overall suitability score map seen in Figure 12. This means that different groups of cyclists, from those that were avid cyclists to novices interpreted the cycling network the same way. Comparing with the four common LOS measures, the suitability rankings are very different. Here, the roadway segments connecting the university, downtown areas, and other major activity centers within the city were identified as the most suitable for cycling. These same segments received rather low LOS scores from the common LOS measures, mainly due to the traffic volumes on them. However, comparing these core segments with the heatmap, these *are* the roadways where cyclists are recording their movements. Also noteworthy is the fact that rankings are consistent for routes rather than the piecemeal approach that the common LOS measures generated.

The distribution of suitable segments is also quite different from the common LOS measure scores, as seen in Figure 13. The overall distribution of scores is skewed left, with a strong emphasis on a set of very suitable routes and an increasing number less suitable ones. This trend is rather consistent across all types of cyclists, as seen in Figure 14 as well. This suggests that while LOS measures based on roadway characteristics provide insight into some factors, cyclists of all types are adjusting their opinions based on what options are currently available and then identifying suitable routes that get them where they need to go regardless. Interestingly, those individuals that cycle every day and those that never cycle have the strongest opinions about E/F segments, with more of them than any other cyclist group. Those cyclists that are not interested in cycling have the least number of A-scored segments, suggesting the strongest bias of the groups. Weekly riders provided the most C- and D-scored segments, which conveys the fact that they are beginning to explore and define their comfort-levels more than less frequent cyclists.

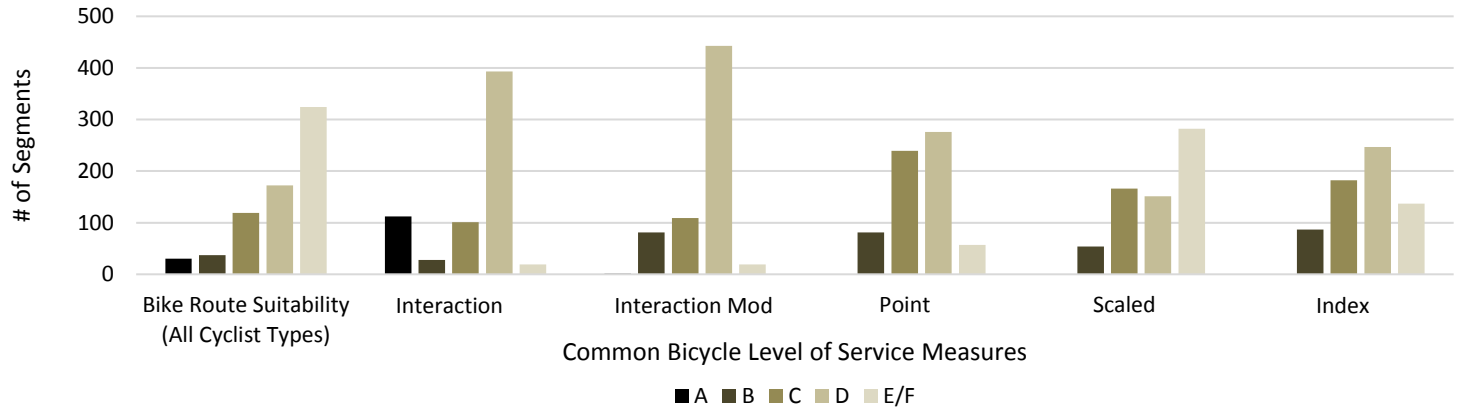


Figure 13: Cycling Level of Service Scores for Road Segment across Common Measures

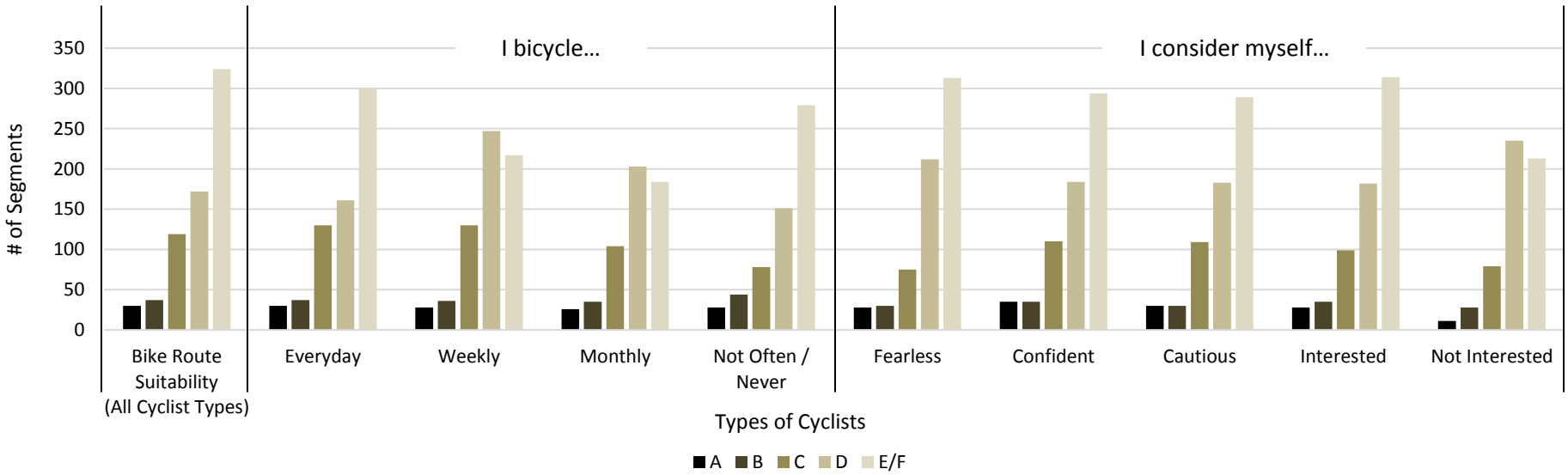


Figure 14: Perceived Bike Route Suitability Scores for Road Segments by Cyclist Type

Finally, Table 3 extends the correlation analysis to include perceived suitability rankings. As to be expected from the mapping, there is a strong correlation among the perceived bicycle route suitability across all cyclist types and little correlation with the other common LOS measures. This difference is most likely due to the fact that the full routes were promoted in the core of city as suitable were not ranked as well by the other measures. There is also little correlation with roadway characteristics. Perhaps this definition of bicycle route suitability is more about connections than traditionally defined level of service; future work into land uses and suitable routes will be explored.

Again, there are two main conclusions to be drawn regarding LOS and bike route suitability. First, bike route suitability is defined by more than just roadway characteristics, although those characteristics are still influential. The fact that LOS measures and suitability do not match indicates that land uses and activity access are also important. It would seem that cyclists seeking routes will accept a certain lower level of service on suitable routes if it is the best route available. It then follows that LOS and suitability are defining two separate cyclist needs. Second, the majority of cyclist types interpreted the bicycle transport system the same way, with similar suitable locations and distributions of preferences. However, this implies that non-cyclists and avid cyclists have different thresholds as to where it is acceptable for them to ride. Even if a novice agrees a roadway is suitable for cycling, there are more factors affecting their willingness to bicycle than roadway characteristics or suitability. This is evidence that infrastructure improvements are not the only means to generate more cyclists.

Table 3: Correlations with Perceived Bike Route Suitability Ranking

		Perceived Bike Route Suitability Ranking by Cyclist Type									
			I cycle…				I consider myself…				
		All Bicyclists	…Everyday	… a Few Times a Week	…a Few Times a Month	…Not Often / Never	…Fearless	…Confident	…Cautious	…Interested	…Not Interested
Perceived Bike Route Suitability Ranking by Cyclist Type	*I cycle…*										
	…Everyday	0.914	1.000	-	-	-	-	-	-	-	-
	…a Few Times a Week	0.888	0.841	1.000	-	-	-	-	-	-	-
	…a Few Times a Month	0.918	0.887	0.870	1.000	-	-	-	-	-	-
	…Not Often / Never	0.944	0.891	0.852	0.895	1.000	-	-	-	-	-
	I consider myself…										
	…Fearless	0.950	0.901	0.857	0.880	0.931	1.000	-	-	-	-
	…Confident	0.945	0.887	0.902	0.908	0.924	0.907	1.000	-	-	-
	…Cautious	0.931	0.884	0.859	0.910	0.928	0.896	0.893	1.000	-	-
	Interested	0.893	0.834	0.859	0.887	0.879	0.861	0.884	0.856	1.000	-
	…Not Interested	0.903	0.890	0.831	0.877	0.899	0.883	0.891	0.862	0.852	1.000
Common Bicycle LOS Measures	Index	0.199	0.215	0.235	0.221	0.235	0.266	0.157	0.216	0.237	0.250
	Interaction	0.415	0.428	0.486	0.454	0.403	0.399	0.405	0.423	0.416	0.508
	Interaction Modified	0.364	0.377	0.442	0.407	0.349	0.363	0.355	0.367	0.382	0.497
	Point	0.071	0.070	0.134	0.107	0.106	0.115	0.066	0.073	0.107	0.078
	Scaled	0.195	0.157	0.244	0.176	0.222	0.263	0.149	0.176	0.194	0.233
Roadway Characteristics	Total Volume	0.372	0.341	0.341	0.358	0.383	0.417	0.309	0.368	0.365	0.481
	Total # of Lanes	0.105	0.113	0.022	0.052	0.146	0.189	0.030	0.094	0.073	0.196
	Speed Limit	-0.194	-0.121	-0.132	-0.226	-0.155	-0.115	-0.228	-0.205	-0.257	-0.059
	Bike Facility Identifier	0.136	0.206	0.179	0.211	0.114	0.085	0.155	0.148	0.114	0.239
	Width Of On-Street Parking	0.108	0.090	0.144	0.124	0.117	0.116	0.118	0.107	0.116	0.094
	Curb Lane Width	-0.159	-0.228	-0.248	-0.261	-0.177	-0.175	-0.147	-0.196	-0.203	-0.271
	Access/Egress	-0.013	-0.012	-0.003	0.016	-0.014	-0.012	0.005	0.025	0.024	0.000

URBAN APPLICATION ANALYSIS

Perceived Level of Traffic Stress

The quality-of-service measure that is most relevant to this research is the MTI Level of Traffic Stress (LTS). The Mineta Transportation Institute study classified roadways and bikeways into four levels of traffic stress according to a modified version of Geller's four types of bicyclists. LTS 1 included facilities suitable for children; LTS 2 facilities characteristics were based on the Dutch CROW (Center for Research and Contract Standardization in Civil and Traffic Engineering (Netherlands)) Design Guide and were intended to be comfortable for most adults; and LTS 3 and LTS 4 present tolerance for characteristics of higher stress (Mekuria et al. 2012). LTS criteria were developed for the following facility types: physically separated bikeways, bike lanes, and shared travel lanes. LTS criteria were developed for right-turn only motor vehicle lanes and unsignalized intersections also. High stress roadways at unsignalized intersections and limited access roadways were identified as the main barriers to low stress bicycling.

The MTI LTS takes into consideration the following variables; number of through lanes, bicycle facilities, posted speed, width of bike lane, width of parking lane, bike lane blockage, right turn lane geometric information, on street parking (alongside bicycle facilities), signalized intersections, and median (Mekuria 2012). The two main strengths of the MTI LTS are being more intuitively understandable to the public and decision makers and considering both current and potential bicyclists. MTI LTS has already been deployed in numerous bicycle and pedestrian plans. The MTI LTS requires the most readily available data out of the quality-of-service models discussed here. Requiring easily accessible data makes the analysis of roadways and bikeways much easier for jurisdictions. Unlike other quality-of-service tools, the MTI LTS categorizes

facilities based on the preferences of the entire adult population who currently bike and who would consider biking.

The MTI LTS has two primary weaknesses: data that requires manual collection and lack of research used to validate traffic and roadway characteristics that affect perceived stress. Another weakness is the approximation of bike lane blockage by assuming that bike lane blockage is frequent in commercial areas and rare in all other areas when it is unknown how effective this method is for approximating bicycle lane blockage by motor vehicles (Mekuria et al. 2012). Manual data collection is required to measure bicycle lane and parking lane width since most jurisdictions do not collect these data. Manual data collection can be very time consuming and may not be feasible. The majority of criteria used to classify roadways and bikeways by LTS level were based on Dutch bicycle design criteria and not through research measuring the perceived stress or comfort of roadway, bikeway, and traffic characteristics for current and potential bicyclists.

As the literature review illustrates, existing bicycle quality-of-service measures often require data that are labor intensive and costly to obtain, lack transparency and are difficult for the public and decision makers to read, and are unable to analyze innovative bicycle facilities such as protected cycle tracks. To help agencies and decision makers have access to a quality-of-service tool that is easily understood and not data intensive, yet effective, this study proposes a modified quality-of-service measure which can be easily implemented throughout the United States.

The modified LTS is built based on the concept that facilities may be associated with different levels of perceived safety and the perception depends on the type of bicyclist and his/her tolerance level for traffic stress. There have been several studies that have classified bicyclists into different categories based on their skill level (Dill and McNeil 2013, AASHTO 2012) and

bicycling frequency (Winters et al. 2011, Dill and Voros 2007, Sanders 2013, Ahmed et al. 2013). However, this study uses the bicyclist classification introduced by Roger Geller (Geller 2006) and later modified by Misra et al. (2014).

Geller (2006) categorized current and potential bicyclists of Portland by their level of comfort riding on different types of roadway and bikeway facilities. The four bicyclist types suggested by Geller are (i) Strong and Fearless (less than one percent of bicyclists), (ii) Enthused and Confident (seven percent), (iii) Interested but Concerned (60 percent), and (iv) No Way No How (33 percent).The Cycle Atlanta typology is a modified version of the Geller typology, in which the No Way No How type was dropped, because the typology includes only descriptions of those who are currently bicycling or who are interested in bicycling. In addition, the Interested but Concerned type used in the Geller typology was split into two types with Comfortable but Cautious category intended to include bicyclists such as females and/or older travelers who are bicycle enthusiasts, but may be more risk adverse (Misra et al. 2014). See Table 4 for descriptions of all four Cycle Atlanta types. People who identify as LTS 2 Comfortable but Cautious are estimated to be the largest type present in the population and will not bike on shared roadways with high motor vehicle speeds and traffic volume, will only bike on roadways with low speeds and low traffic volumes like local or neighborhood roads, and prefer to bike on bicycle or shared-use paths. The Cycle Atlanta typology is used in this research as the basis for the modified LTS roadway and bikeway criteria which are discussed in more detail later.

Table 4: Cycle Atlanta LTS Typology

LTS Type		Description
LTS 1	Interested, but concerned	I have heard a lot about cycling and I am curious to try it, but I require facilities geared to cyclists before I would do so
LTS 2	Comfortable but cautious	I am comfortable on most roads, but strongly prefer facilities geared to cyclists and will chose another mode depending on facilities
LTS 3	Enthused and confident	I am confident sharing the road with vehicles but prefer facilities geared to cyclists
LTS 4	Strong and fearless	I am willing to bike in any situation and being a cyclist is part of my identity

MODIFIED LTS MEASURE

The modified LTS quality-of-service measure builds upon the MTI LTS and classifies roadways and bikeways by one of four levels of traffic stress based on traffic and geometric characteristics such as traffic volume, posted speed limit, number through lanes per direction, presence of on street parking, and bicycle facility type. Roadways and bikeways categorized at LTS 1 are the least stressful and have low traffic volumes and low speed limits, while roadways and bikeways categorized as LTS 4 are the most stressful and have the highest traffic volumes and speed limits. It is estimated that the majority of current and potential bicyclists find LTS 1 and LTS 2 facilities comfortable. Table 5 provides a description of the characteristics of roadways and bikeways for each LTS. This table is a modified version of a similar table used by the Mineta Transportation Institute to describe the roadway and traffic characteristics of its LTS measure. MTI LTS classifies protected shared paths, cycle tracks, and side paths as LTS 1, however, the modified LTS re-classified protected cycle tracks and side paths as LTS 2 due to the increased presence of conflict zones such as driveways and intersections for these facilities as opposed to the presence of few conflict zones for most shared paths. MTI LTS considered LTS 1 facilities suitable for children; however, the modified LTS does not make assessments for children since there is very limited research on perceived stress for children. The modified LTS also introduced buffered bicycle lanes as a facility type since this facility type was not considered by the MTI LTS.

Table 5: LTS Roadway and Bikeway Characteristics

LTS Level	Modified LTS Roadway and Bikeway Descriptions
LTS 1	Considered comfortable and low stress by almost all cyclists. Includes shared paths which separate cyclists from motor vehicle traffic and present few conflict zones such as intersections and driveways. Shared travel lanes are only tolerable if traffic volume is so low that cyclists only occasionally interact with motor vehicles and there is little difference in travel speed between cyclists and motor vehicles due to a posted speed limit of 25 mph or below. Intersections are low stress to approach and cross.
LTS 2	Considered low stress by all cyclists except for people who identify as LTS 1. Includes side paths and protected cycle tracks which are low stress, but present some conflict zones at driveways and intersections. Shared travel lanes can only have one lane per direction, a speed limit of 30 mph or below, and must be classified as local. Conventional bike lanes and buffered bike lanes allow for slightly higher traffic volume, speed, and classification as local or collector.
LTS 3	Conventional bike lanes or buffered bike lanes are located on roadways with moderate traffic volume and speed and can be classified as minor arterial or lower. Shared travel lanes must be classified as collector or lower and 35 mph or lower. Roadways of LTS 3 can have 2 lanes or less per direction.
LTS 4	Any level of stress beyond LTS 3 excluding limited access roadways. Includes all roadways with a posted speed limit above 40 mph and/or 3 or more lanes per direction with or without bicycle lanes.

Calculating Level of Traffic Stress

The details of traffic stress classification for separated bicycle facilities are presented below. The criteria tables for shared travel lanes and on-road bicycle facilities are also given. Note that criteria tables follow the rule that the aspect of a link with the highest LTS determines the LTS of that segment. For example, a conventional bicycle lane with no adjacent motor vehicle parking with one through lane per direction (LTS 1), a posted speed of 35 mph (LTS 3), a functional class of collector (LTS 2), and a traffic volume of 10,000 vehicles per day (LTS 2) would be classified as LTS 3 for the link as a whole. The notation "(no effect)" means that the factor does not cause an increase to that LTS.

Criteria for Separated Bicycle Facilities

Research has shown that people prefer separated bicycle infrastructure (Winters et al. 2010, Kremm et al. 2014, Broach et al. 2012, Misra et al. 2014, Dill and Voros 2007, Monsesre et al. 2014). MTI LTS classified all separated bicycle facilities (shared-use paths, side paths, and protected cycle tracks) as LTS 1. However, this method does not consider the potential stress of bicycle and motor vehicle interaction at driveways, intersections, and loading areas. Therefore, in this study, separated bicycle facilities or shared-use paths, which are the most separated from motor vehicle traffic, are classified as LTS 1. Protected bicycle facilities such as side paths, one and two way cycle tracks, and raised cycle tracks are classified as LTS 2 due to the potential interaction of motor vehicles and bicycles at midblock driveways, intersections, and loading bays.

Traffic, Roadway, and On-Road Bikeway Characteristics

The roadway and traffic characteristics which are considered include: number of through lanes per direction, traffic volume or annual average daily traffic (AADT), functional class, and posted speed limit. The focus on traffic volume and speed is supported by Winters' survey of current and potential bicyclists in Metro Vancouver. This study found that high traffic volume and traffic speed were major deterrents from riding (Winters et al. 2011). Thus, for conventional bicycle lanes, buffered bicycle lanes, and shared travel lanes, the level of traffic stress for a link increases as those variables increase. The perceived stress caused by the presence of or lack of on street motor vehicle parking was also considered.

Traffic Volume or Annual Average Daily Traffic (AADT) and Functional Class

MTI LTS does not include traffic volume or functional class when classifying facilities. However, research has shown that the majority of people who want to bicycle more list "too much traffic" as the top environmental barrier (Dill and Voros 2007). Therefore, traffic volume and functional class were included in this study. Number of travel lanes and functional class have a strong relationship, as the USDOT FHWA Highway Functional Classification Concepts, Criteria, and Procedures states, "roadways are designed and constructed according to their expected function" (USDOT 2013). For example, an arterial is designed to be a high capacity roadway and would likely have more travel lanes, while a collector would likely have less travel lanes than an arterial and a local road even less travel lanes than a collector. Research by Winters et al. (2010) also found that when comparing shortest route to actual route, bicyclists traveled significantly less along arterial roads than predicted by the shortest route model and significantly more along local roads.

Number of Through Lanes per Direction

Multilane streets, as opposed to those with one lane in each direction, promote higher motor vehicle traffic speed and decreases the visibility of bicyclists for left-turning and cross motor vehicle traffic at intersections and driveways (Mekuria et al. 2012). The MTI LTS based its LTS criteria for number of lanes on the Dutch CROW Design Manual and modified the Dutch standards by allowing more lanes per direction if the roadway had a median. This study did not consider medians due to the lack of data on the location of medians in the case study area. However, roadways were categorized using the basic number of through lanes criteria used by MTI.

Posted Traffic Speed

High motor vehicle travel speeds have been rated by current and potential bicyclists as a deterrent to bicycling (Winters et al. 2011). Measures of observed speed when available are the best data to use especially when observed traffic speed and the posted speed limit differ. However, observed traffic speed is typically not available. Data on posted speed limit are readily available and for this reason, were used in the study. The posted speed limit criteria used in this study follow the methodology used by MTI for conventional bicycle lanes. This study modified the conventional bicycle lane criteria table to create a buffered bicycle lane table since MTI did not include criteria for buffered bicycle lanes in its analysis. The criteria table for buffered bicycle lanes allow for a slighter higher posted speed limit and functional classification; however, the AADT and number of through lanes per direction remain the same.

On Street Parking

Winters' survey of Metro Vancouver residents found that respondents preferred streets without on street parking to those with on street parking (Winters 2011). It would be preferable to consider if the width of the bicycle lane and parking lane were adequate to reduce perceived stress due to the potential of "dooring". However, parking and bicycle lane width data are typically not readily available. Data collection for on street parking was limited to conventional bike lanes and buffered bike lanes due to the potential that these facilities would position riders in the "dooring" zone.

Beltline Case Study

The modified LTS measure was used to classify roadway and bikeway facilities within a six-mile buffer of the Atlanta BeltLine Eastside Trail. The Eastside Trail is a small part of a much

larger transportation and economic development project which will provide parks, shared use paths and transit along a 22-mile historic railroad corridor in Atlanta, Georgia (Atlanta Beltline 2015). The completed Atlanta BeltLine will connect 45 neighborhoods. Four sections of the trail are currently completed, and the Eastside Trail, which is the focus of this case study, was the first segment to be completed (Atlanta Beltline 2015). The case study area was limited to six-miles around the Eastside Trail as research has shown that routes over six miles are perceived as a strong deterrent in the choice to bicycle for many people (Winters et al. 2011).

Data

Three primary data sources were used in this analysis. The NAVTEQ Streets 2014 shapefile was obtained by Atlanta Regional Commission (ARC) from the company HERE. It includes a comprehensive inventory of roadways, especially local roadways that are omitted from other data sources. The other roadway database used in the research, RC_ROUTES_ARC, is a modified version of the roadway database maintained by the Georgia Department of Transportation (GDOT) and focuses on state managed roadways rather than locally managed roadways and bikeways. The third data source was the Metro Atlanta Bicycle Facility Inventory, which was compiled from information provided by local governments in the region and verified with Google Earth and Bing Imagery. The location of on street parking on roadways with conventional bicycle lanes and buffered bicycle lanes was manually coded in ArcGIS using Google Earth imagery.

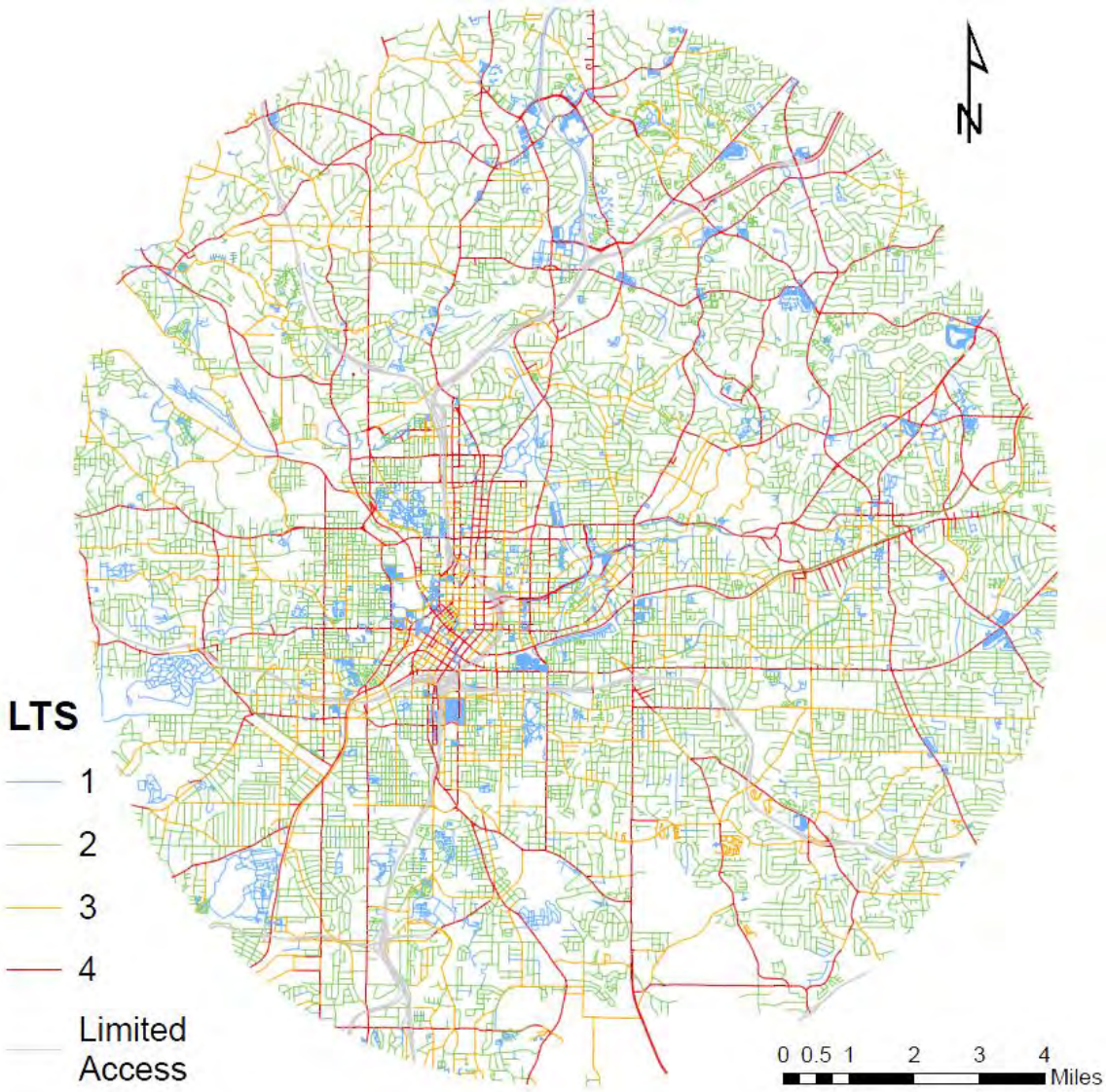


Figure 15: LTS Measure Applied in Case Study Area

An overview of the case study area with the modified LTS measure applied can be seen in Figure 15. LTS is coded by color with blue = LTS 1, green = LTS 2, orange = LTS 3, red = LTS 4, and grey indicating limited access roadways. While only 15% of facilities are LTS 1, a little over half, 54%, of the facilities are LTS 2. The robust presence of LTS 2 facilities in general was also noted in the MTI study (Mekuria et al. 2012) and indicates the prevalence of local or neighborhood streets in the case study area. Approximately 69% of the roadways and bikeways in

the case study area are classified at LTS 1 or LTS 2 which are considered low stress for the majority of current and potential bicyclists. As can be seen in Table 6, 344 miles of facilities are LTS 1 and 1,223 miles are LTS 2 in the case study area for a total of 1,567 miles of low stress facilities. 1,524 of those miles are low stress local roadways. The shared travel roadway criteria can be seen in Table 6. For a shared travel roadway to be classified as low stress, it must be a local street with a maximum speed limit of 30 mph, have a traffic volume of 6,000 vehicles per day or less, and have a maximum of one through lane per direction of travel.

Table 6: Distribution of Centerline Miles by Level of Traffic Stress and Facility Type

	LTS 1	LTS 2	LTS 3	LTS 4	N/A	Total Miles	Total %
Conventional Bicycle Lanes	0%	16%	**59%**	25%	-	35	100%
Buffered Bicycle Lanes	0%	6%	12%	**82%**	-	2	100%
Shared Travel Roadways	16%	**59%**	12%	13%	-	2028	100%
Side Paths	**100%**	-	-	-	-	10	100%
Protected Cycle Tracks	-	**100%**	-	-	-	1	100%
Shared-Use Paths	-	**100%**	-	-	-	27	100%
Limited Access Roadways	-	-	-	-	**100%**	164	100%
Total Miles	344	**1223**	270	266	164	2267	

Figure 16 presents a zoomed-in version of Figure 15 to provide a more detailed image of the LTS classification of roadways and bikeways around the Atlanta BeltLine Eastside Trail, the focus of this case study.

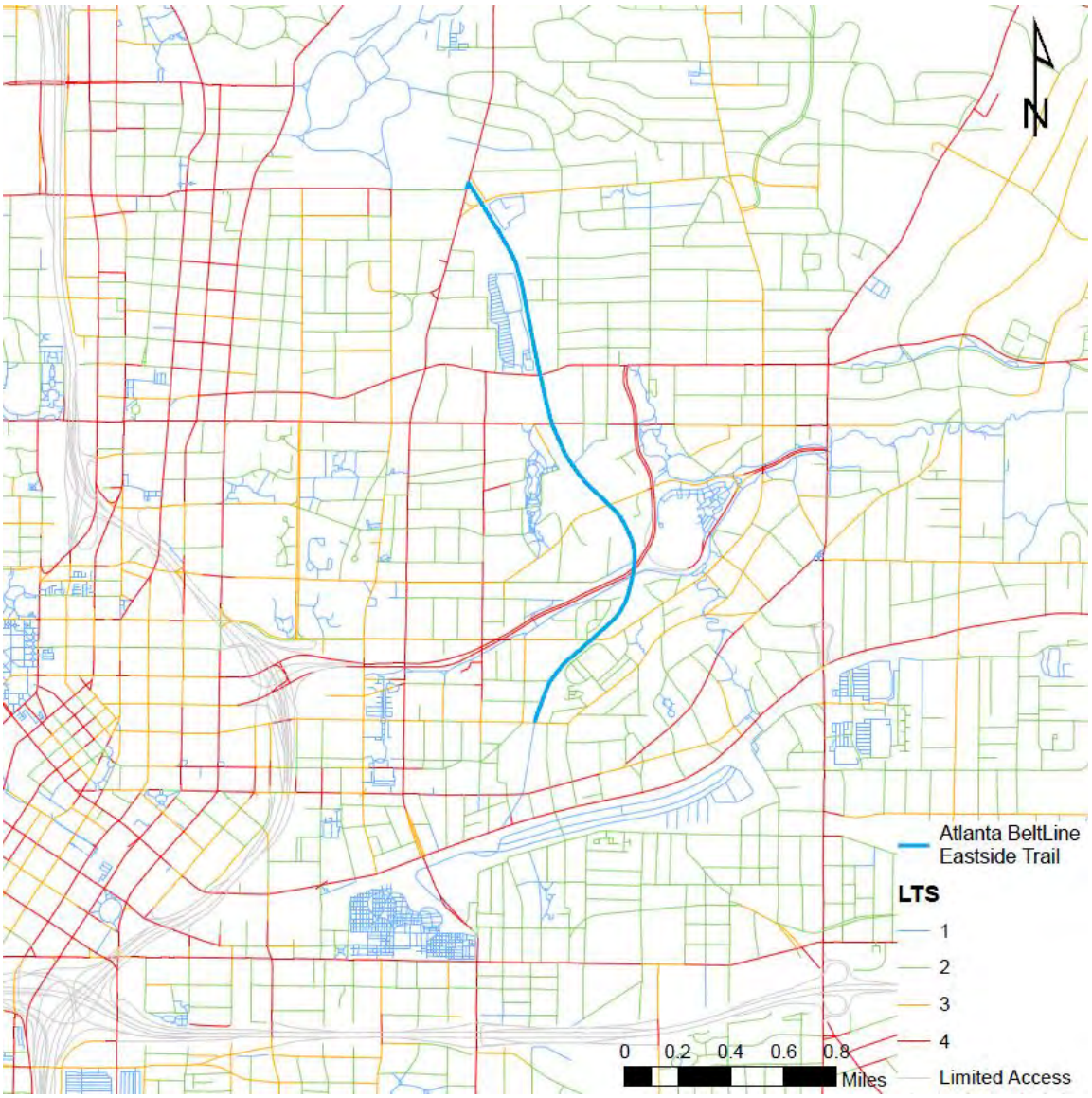


Figure 16: Closer View of LTS in Case Study Area (Atlanta Beltline)

While the majority of shared travel roadways are categorized as low stress facilities (LTS 1 or 2), any collector or arterial functional class roadway without a bicycle facility is categorized as high stress (LTS 3 or 4). The majority of the conventional bicycle lanes were categorized as high stress due to high traffic volume and high number of through lanes per direction. The buffered bicycle lane criterion allows for a higher threshold for traffic volume and number of through lanes

per direction; however, the two miles of facilities which have currently been built in the case study area are on roadways with traffic volumes that exceed the LTS 1 and 2 threshold.

Even though a majority of the facilities are classified as low stress by the modified LTS measure, it does not mean that the facilities are well connected, as shown in Figure 16. Connectivity in the study area is reduced as a result of two factors; limited access roadways which do not allow bicycle traffic and collector and arterial functional class roadways which trigger the high stress classification. A total of 164 miles of limited access roadways exist throughout the case study area. While there are only 419 miles of collectors and arterials in the study area, they present barriers to a connected bicycle network. Investment in strategic bicycle facilities may be needed to create connected low stress facilities across interstates and other limited access roadways.

A map of roadways and bikeways classified as LTS 1 or LTS 2 is shown in Figure 17. This map reveals that while a majority of the roadways and bikeways in the study area are classified as LTS 1 and LTS 2, these facilities appear to not be well connected. This concept is explored further in the map in Figure 7-4 where the Atlanta BeltLine Eastside Trail's bikeshed is considered for LTS 1 and LTS 2 facilities. The overview map, Figure 18, shows that the bikeshed does not spread very far outward and includes gaps within the bikeshed.

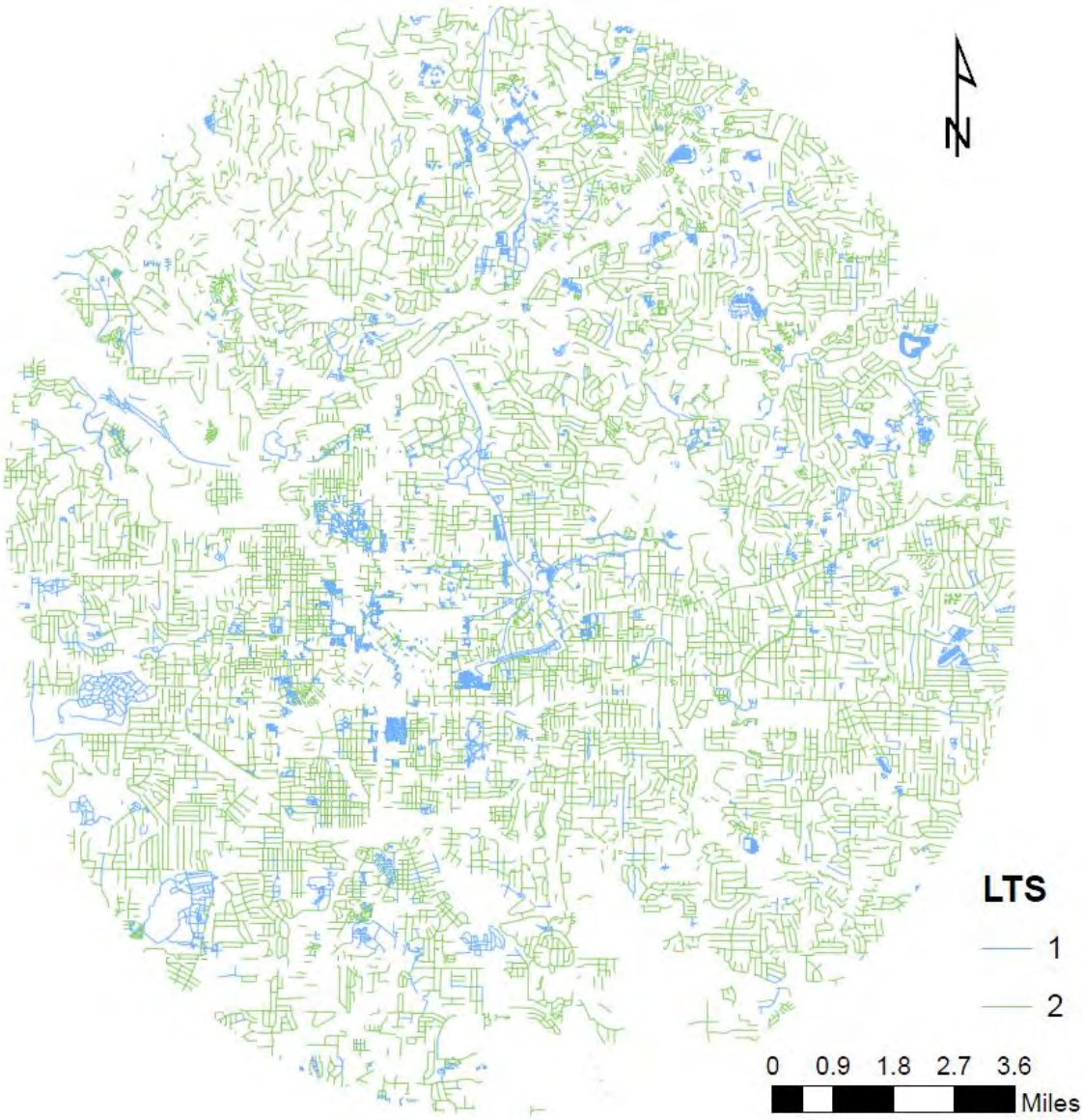


Figure 17: Case Study Area LTS 1 and LTS 2 Facilities Only

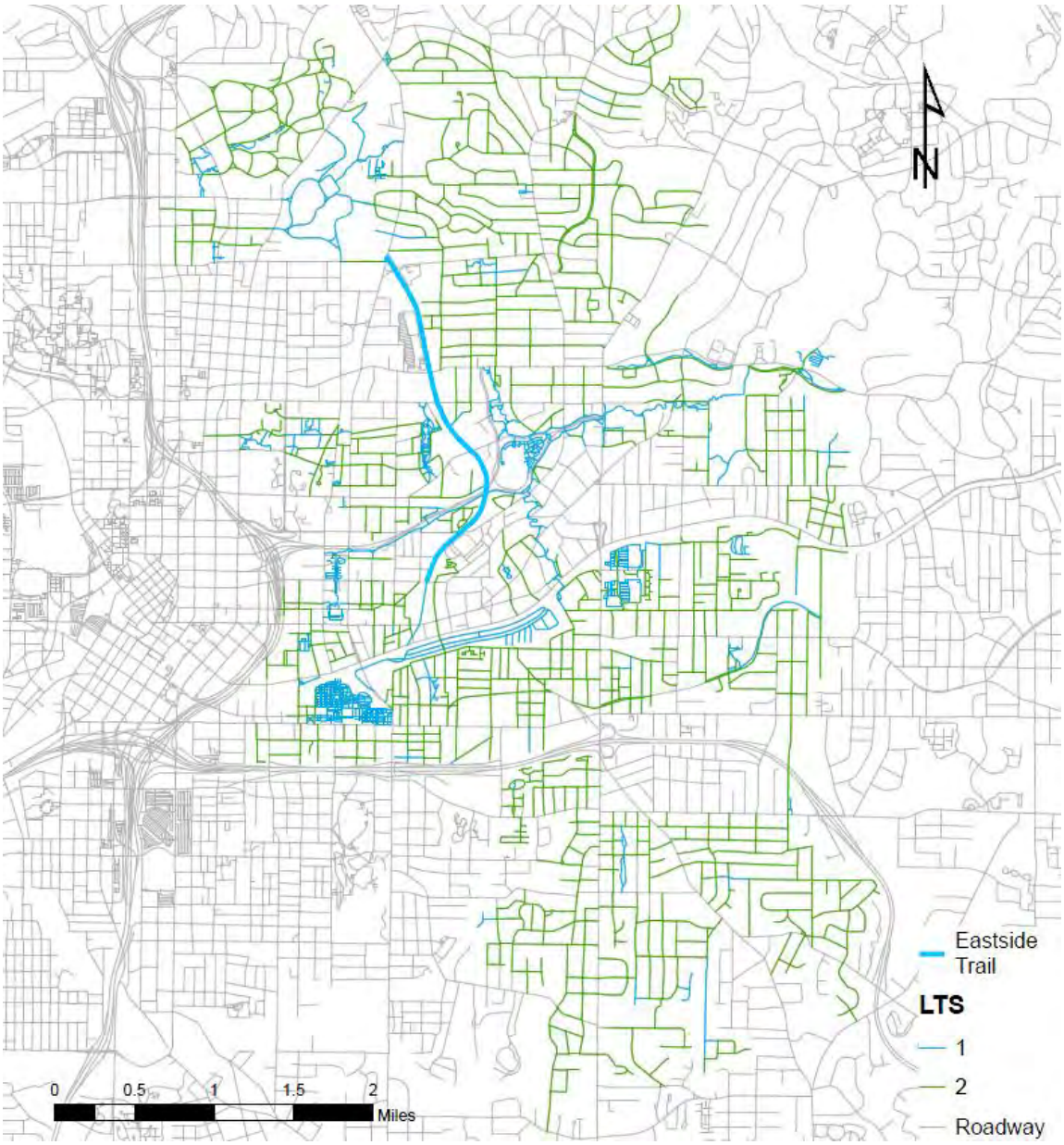


Figure 18: Closer View of Eastside Trail Bikeshed with LTS 1 and LTS 2 Facilities Only

Although a cursory overview of the study area reveals a large amount of LTS 1 and 2 roadways and bikeways, further analysis reveals that the facilities are not well connected. This means that for people who are comfortable using facilities at LTS 1 and LTS 2, estimated to be a majority of current and potential bicyclists, the bike network is disconnected. The case study also shows that while local roadways are an important part of a low stress bicycle network, a well-connected bikeway network cannot be achieved with local streets alone. Collector and arterial roadways provide the connectivity of a roadway network, yet they are too stressful for the majority of current and potential bicyclists without bicycle facilities that provide separation from motor vehicle traffic. Conventional bicycle lanes and buffered bicycle lanes are appropriate to install on collector and arterial roadways when the traffic volume is lower. However, collector and arterial roadways with high traffic volume require greater separation through the use of protected cycle tracks or side paths.

MARTA Stations Case Study

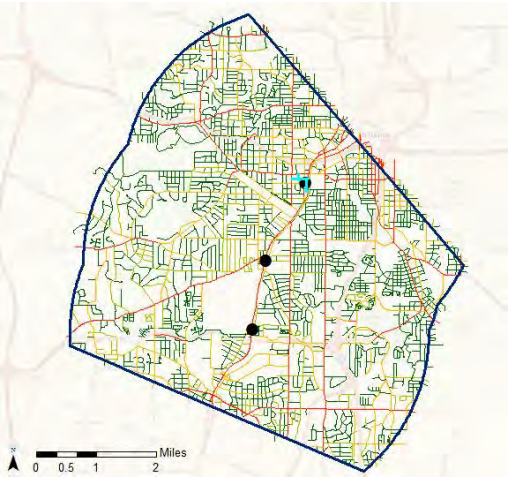
A further investigation was conducted to demonstrate the utility of using LTS methodology in evaluating the impact of bicycle infrastructure investments. In this case, a similar analysis was undertaken for a 3 miles buffer around the MARTA West End, Oakland City, and Lakewood/Ft. McPherson stations. Improving the bicycle network around MARTA stations can directly increase the bike catchment area for that station and, as a result, could substantially change the commute environment around that station. These stations were chosen specifically for the current development strategies based on market strength and social equity.

To evaluate the low stress bike networks accessing the West End, Oakland City, and Lakewood/Ft. McPherson MARTA stations, three low stress (LTS 1-2) networks as well as the entire (LTS 1-4) bike network were compared based on total network length, accessible area, and

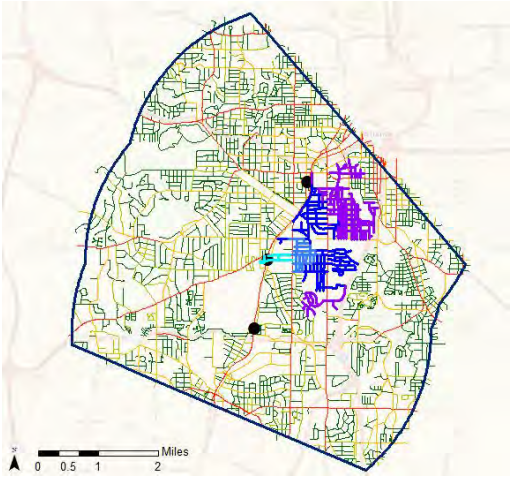
accessible population. The accessible area and population were determined based on the 2010 census blocks that intersected each network. The 2010 census was used instead of the 2009-2013 5-year American Community Survey (ACS) estimates because the 5-year estimates are only available at the block group level. The study area population was only 1.7% larger based on the 2013 5-year ACS census block group estimates compared to the 2010 census blocks, and so the 2010 census blocks were chosen for the analysis to allow higher precision. The block group was not granular enough to provide a precise enough definition of the study area.

The low stress networks analyzed were based on the existing low stress infrastructure, proposed improvements in the area, and select key improvements based on the LTS analysis. The final entire LTS bike network included the entire bike network and represented the network available to the most stress-tolerant bicyclists. For each of these analyses, the LTS network was converted into a Network Dataset in ESRI ArcMap. The service area tool identified the streets that were within a network distance of 3 miles from each of the study area MARTA stations.

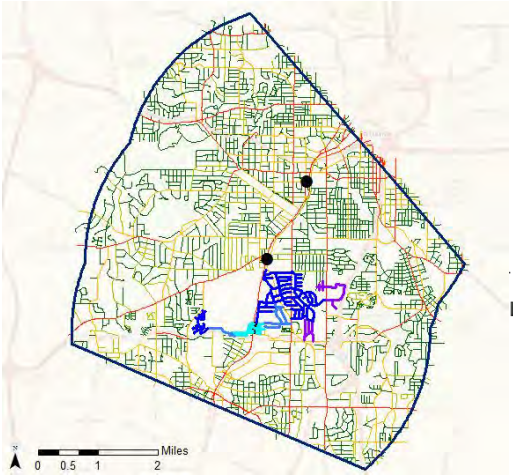
Figure 19 shows the LTS 1-2 area accessible to each of the study area MARTA stations by network distance.



West End



Oakland City



Lakewood / Ft. McPherson

3 miles
miles

Figure 19: Service Area Analysis based on Existing Conditions LTS 1-2 (Blue) and LTS 1-4 (orange) Network

Proposed Improvements – Low Stress Network

Figure 20 highlights the location and LTS classifications for the proposed improvements (the thick line shows the improved LTS and the superimposed thin line shows the original LTS for the same link). The specific improvements are concentrated in the around West End MARTA station. The addition of the Southwest portion of the beltline trail and the proposed multi-use trail along Peters Street and Lee Street are the most impactful improvements. Figure 21 shows the bike-able network based this proposed network, restricted to a 3 mile network distance from each of the study area MARTA stations.

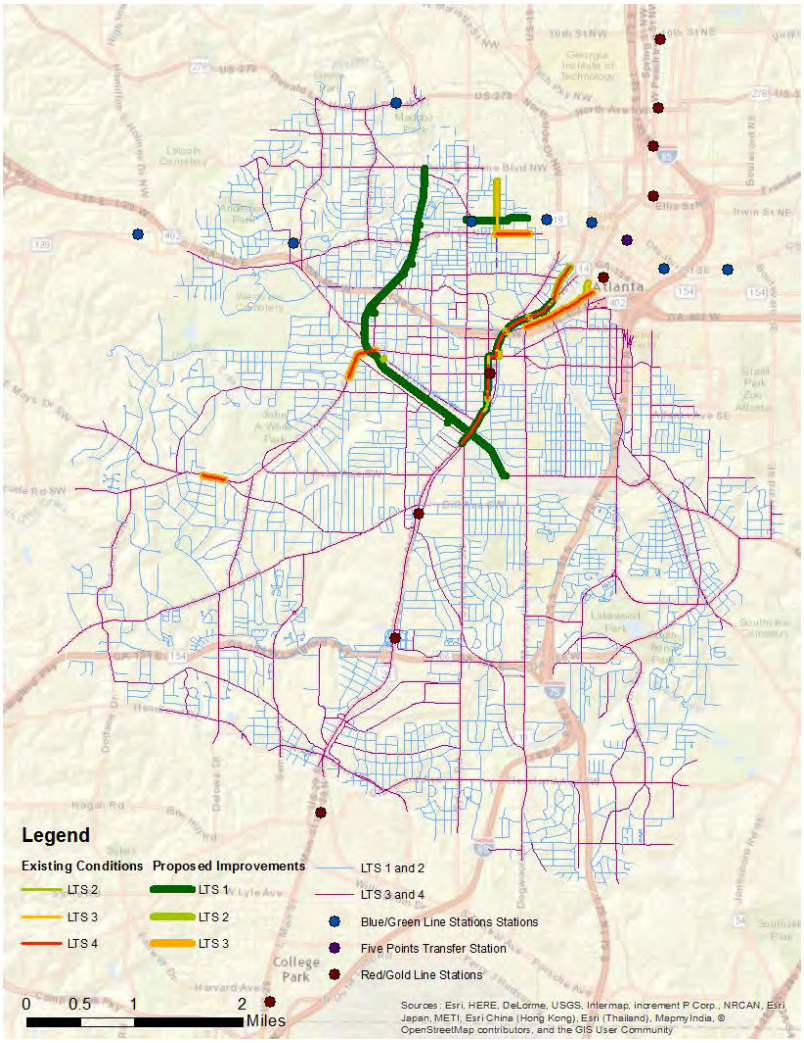
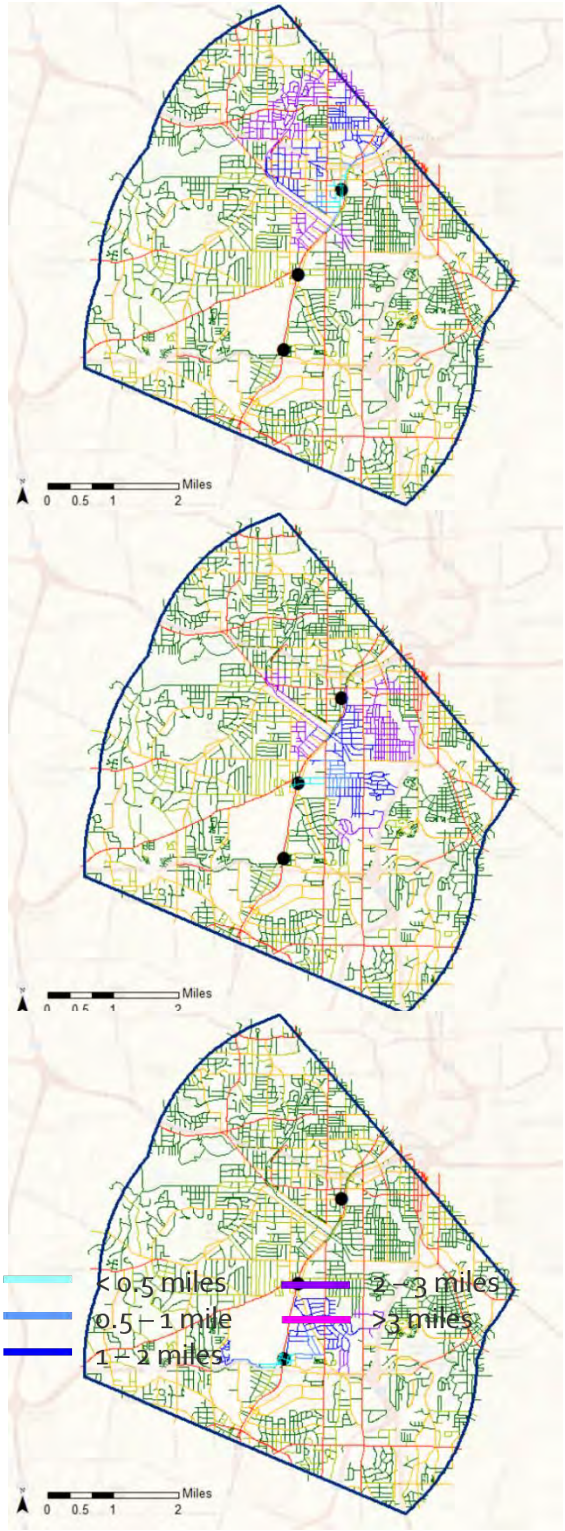


Figure 20: LTS for Links with Proposed Improvements (thick line) and Previous LTS (thin line)



West End

Oakland City

Lakewood / Ft. McPherson

Figure 21: Service Area Analysis based on Proposed Conditions based on Cycle Atlanta Phase 1.0 Plan, Infrastructure Bond, and Southwest Beltline Access Points

CHAPTER 4: MODELING CYCLING ROUTE SEGMENT & PATH CHOICE

This chapter uses the results from the previous chapter to identify the factors that have the most significant impact on cyclists' route choices in both suburban and urban areas.

SUBURBAN APPLICATION ANALYSIS

Modeling Methodology and Dataset Generation

An ordinal logistic regression was utilized to determine how likely each link in the network would be used as part of a cycling route. Ordinal logistic regression is a discrete choice model, which means that the dependent variable being estimated (in this work, a cycling route likelihood level) is categorical. A multinomial logit regression, another discrete choice model option, was not selected, as the dependent variable used in this work had an ordered nature to it: links could fall into one of five categories: from low, low-average, average, high-average, to high. Roadway characteristics, regional characteristics, and accessibility measures were included as potential independent variables in the estimation. The model assumes that alternatives are independent and identically distributed (IID), with a normally distributed error term estimated.

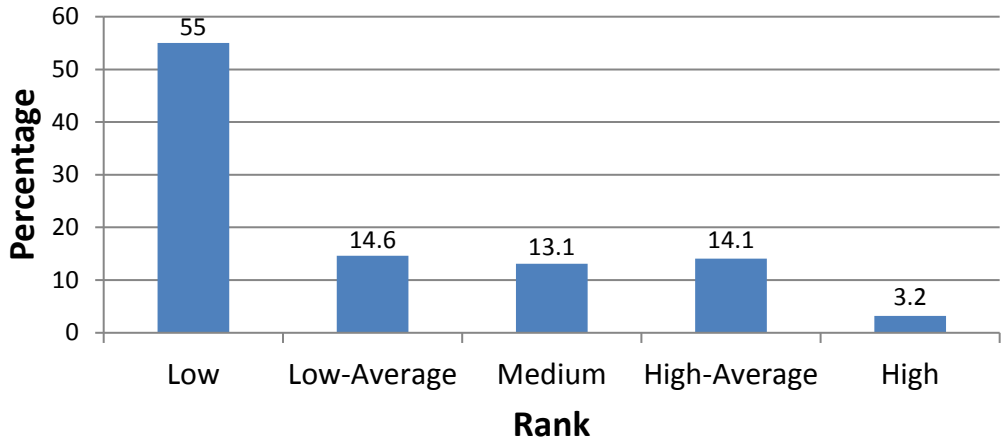


Figure 22: Percentage of Strava Users per Usage Rank

The ordered logit model is estimated by assuming that the series of dependent cycling route likelihood categories are all related to an underlying continuous utility value. This is a logical connection to make with this research, since the groupings are naturally progressive. The categorical version of the dependent variable is derived from setting thresholds in this continuous utility value. The benefit of this method is that the thresholds do not need to be evenly spaced and can reflect that there are nonlinear jumps in the factors when assigning cycling likelihood categories. The equation for the utility function is:

$$U_i = \sum_{n=1}^{N} \beta x_{ni} + \varepsilon$$

where i is the segment number, n is the variable number, x_{ni} is the value of variable n on segment i, β is the coefficient weight on variable x_{ni}, U_i is the utility of segment i , and ε is the normally-distributed error term. The independent variables considered for inclusion on the model can be seen in Table 7.

Table 7: Independent Variables Considered for Model

Roadway Characteristics	**Socio-Demographic Accessibility**	
Peak Hour Volume	Number of People Aged…	Household Size…
Number of Driveways	…5 to 17	… 1 member
Width of Outside Lane	…18 to 24	… 2 members
Width of Paved Shoulder	…25 to 39	… 3 members
Number of Lanes	…40 to 64	… 4 members
Pavement Condition	…65 and Up	… 5 or more
Total Volume		
Speed Limit	Number of Households with Income…	Vehicle Ownership…
Bike Facility Present	…10k to 29k	… none
Median Present	…30k to 59k	… 1 vehicles
	…60k to 99k	… 2 vehicles
Access Groups	…100k and Up	… 3 vehicles
Residential		… 4 vehicles
Shopping	Household Commute Times…	… 5 or more
Restaurants	… less than 10 minutes	
Mixed Development	…10 to 19 minutes	
Government	…20 to 29 minutes	
Community Spaces	…30 to 44 minutes	
Educational	…45 to 59 minutes	
Health Care	…60 minutes and up	
Office Space		
Parking		

The model parameters, including the coefficients and threshold limits were estimated using the Maximum Likelihood Estimation (MLE), which is an iterative process that determines the set of parameter values that achieves the observed set of outcomes. In this work, the MLE process tried to match the observed category assigned to each road segment.

The Pearson Chi-Square Goodness-of-fit measure was used to determine whether the model developed was significantly better than a constants only function. The Chi-Square value for the model was found to be 2,986.92, which is significantly greater than the critical value at 21 degrees of freedom for 99.5% level of confidence, χ^2= 41.401, which indicates a strong model.

To test the individual variables to determine whether they were significant, a student t-test was utilized. Each variable was tested against a mean of zero, representing a model that did not contain the variable of interest. Using the variable estimate and standard error, the t-test could be performed with the resulting t-statistic showing the confidence level. All of the variables and their coefficients resulted in a confidence level of 90% with all but one resulting in a confidence level of 95%. The student t-test formula used can be seen below with $\bar{x}$ being the variable mean, μ the hypothesized mean (in this case 0), and SE the standard error.

$$t = \frac{\bar{x}-\mu}{SE}$$

Analysis and Results

This section discusses the results of the ordinal logistic regression model that was developed. The variables that were included in the model, including the coefficients and t-stat, can be seen in Table 8. The final model developed including variables pertaining to the physical characteristics of the roadway, Access groups to different land uses, and Socio-Demographic access, with the variables having a positive coefficient increasing the likelihood of roadway segment use, and those having a negative coefficient decreasing the likelihood of roadway segment use. The variables were evaluated at the 90% confidence level, with those being insignificant dropped from the model.

Table 8: Ordinal Logistic Regression Variables

Explanatory	*Coeff*	*t-stat*
Threshold		
γ_1	1.701	2.19
γ_2	2.498	3.20
γ_3	3.410	4.34
γ_4	5.634	6.90
Roadway Characteristics		
Peak Hour Volume	0.003	6.13
Number of Driveways	-0.094	-3.29
Width of Outside Lane	-0.130	-2.63
Width of Paved Shoulder	0.100	2.48
Access Groups		
Residential	0.305	6.29
Shopping	2.373	3.52
Restaurants	-21.369	-3.23
Mixed Development	19.270	3.84
Government	-1.098	-3.23
Socio-Demographic Accessibility		
Number of People Aged…		
…5 to 17	-1.598	-5.23
…65 and Up	2.034	5.38
Number of Households with Income…		
…10k to 29k	-0.286	-2.73
…30k to 59k	-1.115	-3.39
…100k and Up	-0.871	-1.94
Household Commute Times…		
...30 to 44 minutes	2.047	4.38
…45 to 59 minutes	2.002	3.07
…60 minutes and up	1.140	2.30

Roadway Characteristics

The first major set of variables used in the model was based on the characteristics of the roadway, such as roadway width, number of lanes, etc. From the ordinal regression that was performed it was seen that paved shoulder width had a positive impact on how well the link performed as part of the cyclists chosen route. This positive impact shows that as the width of the paved shoulder increases, that link has a higher likelihood of being chosen as part of cyclist's route. The positive impact that shoulder width had on the likelihood of choosing that link makes sense in that, the more space that cyclists have on the shoulder, the further away from traffic the cyclists can travel and maintain more of a buffer space between the traffic and themselves. The majority of roadway segments in the city of auburn do not include a paved shoulder. However, when comparing where cyclists are traveling with this figure it there are a few spots where there is an increased amount of cycling activity in areas that have a paved shoulder, no matter how wide, suggesting that some shoulder width is better than not having a shoulder.

At the same time that shoulder width has a positive impact, the peak hour volume also was shown to have a positive impact on the likelihood of a link being used as part of a cyclist's route. The positive coefficient in Table 8 shows that a roadway with higher Peak Hour volumes is more likely to be used as part of a bicyclist's route. This positive impact with increased peak hour volume is interesting since common thought would be that as the volume of a road increases, it would be less desirable for cyclists to ride on that stretch of roadway. While this roadway characteristic is having an opposite impact on route choice than would be expected, it could be that those links that have higher road volumes also are better connected to where the cyclists want to go and are being chosen due to their connectivity, even if the traffic levels are higher than other links that aren't as well connected. When looking at Table 8, it can be seen that the routes that,

while not using the highest peak hour volume roadway segments, the average to high-average volume roadways are being used over those with lower peak hour volumes.

The other two roadway characteristics that were found to be significant were number of driveways along the link, and width of the outside lane. While the first two variables had a positive impact on the significance of the route link, these two variables had a negative impact. This negative impact shows that as the number of driveways along a stretch of road increase, the likelihood that cyclists will choose that as part of their route will decrease. This result is intuitive since as the number of driveways increases, the number of possible interactions with vehicles increases, causing the cyclists to feel less comfortable on the road as they have an increased possibility of collision with a vehicle. Segments containing the highest amount of driveways having little use by cyclists.

Similarly as the outside lane width increases, the likelihood that a cyclist will use that road as part of their route also decreases. Unlike a few of the other variables, this result is counterintuitive since a wider lane would seem to be more appealing by allowing the cyclists more space on the roadway. With the increase in the outside lane often being done as a way to provide space for cyclists to ride, without having to add a bike lane to the roadway, cyclists still have to ride within the flow of vehicular traffic, increasing the odds of a collision with the vehicle traffic, than if the cyclists were provided with their own dedicated lane.

With the City of Auburn using the standard 11 to 12 foot lane widths for the main there is not much variation found within the City. What can be observed is that the not as well connected roadway segments have the wider outside lane widths. Although these roadway segments have the wider lane widths, because they are not as well connected as other roadway segments they have a lower usage amongst the Strava users.

Access Groups

The next set of variables used in the model looked into how well connected the individual links were to the whole area and how well the links were connected to certain land use groups, such as residential areas, shopping areas, governmental areas, etc. The first variable tested from this group was looking into the effect of a link being well connected to Residential areas. From the regression analysis run, it can be seen that this variable had a positive coefficient suggesting that cyclists choose links that are well connected to residential areas. It can be seen that the areas that have the higher number of residential land-use also correspond to where the cyclists are traveling. This is makes sense in that the residential area of the city are going to be the larger trip generators for cycling and that an increase in accessibility to these locations has an increasing effect on the likelihood of usage for a roadway segment.

Links with higher accessibility to shopping also had a positive impact on the likelihood of a link being chosen for a cyclist's route. While shopping had a positive coefficient from the model, areas with higher accessibility to restaurants have a negative coefficient. With the way that shopping and restaurants are located in the City of Auburn, these two access groups should be discussed together. Since shopping and restaurant areas in the City of Auburn are located in the same areas, it would make sense that these two land-uses would both have coefficients with the same sign in front. When looking at the coefficients, if they were to be combined, the overall coefficient would be negative suggesting that because these land-uses are typically found on the main roads within the city that the segments most accessible to them are being avoided.

Another land-use accessibility variable that had a positive impact on the likelihood of a link being used was the access to mixed development. This positive impact is most likely due to the fact that in mixed development areas, there are not only shopping and restaurants, but also

residential areas in which those choosing to bike could be living, with the area being designed for not only vehicles but also pedestrians and bicyclists. Mixed Development areas also provides cyclists the ability to bike to one location and then to be able to walk around and enjoy multiple kinds of activities, i.e. shopping, restaurants, and entertainment, without having to commute from one location to another. Another factor resulting in a positive coefficient for mixed development is that these areas are found toward the center areas of the City of Auburn. With mixed use developments being in the center of the city, they are equidistant to the outer edges of the city.

Another variable that had a negative coefficient associated with it was the access to governmental areas. This also makes sense in that the governmental facilities are on the periphery of the City of Auburn. Since these facilities are located on the periphery of the city, there are not as many roadway segments with access to these areas, resulting in cyclists avoiding these areas since there is not adequate access to them.

Socio-Demographic Access

The next set of variables dealt with how links accessible to areas with different socio-demographics were likely to be selected as part of a cyclist's route. It was found that links that had higher accessibility to areas with people aged 5 to 17 were negatively impacted in likelihood of being chosen as part of a route. Being highly connected to areas with large numbers of children is negative on the likelihood of that link being used as part of a route because it means that link is most likely located in a neighborhood, which tend to be toward the edges of cities and not in the center where all the activities of a city are taking place. On the other hand, it was found that links being highly accessible to areas with people aged 65 and up were positively impacted with respect to likelihood of being chosen as part of a cyclist's route.

The area with the lower age groups and those with the higher age groups can be identified. What is interesting is that by mapping the average age of the census block groups, it is straightforward to determine which areas have the higher student populations. While not the main cause for their increased use, Wire Rd and portions of Donahue Dr. traverse the areas that appear to have a higher student population. The highest age groups can also be seen to be located in area closer to downtown, giving them better accessibility to various parts of the city.

Along with the accessibility to different age groups, how well a link was accessible to different income groups was also analyzed. It was found that roadway links with higher connectivity to lower income areas, $10k to $29k, $30k to $59k, and $100 and up were less likely to be selected as part of a cyclist's route. The median income for each block group within the City if Auburn can be seen. While the lowest income group tends to be more toward the center of the city, matching up to where the younger population reside, the higher and low-average income groups tend to be in the outer edges of the city. By the higher and low-average income groups being on the outer edge of the city, these groups are more sectioned off from the rest of the city resulting in less roadway connections and lower accessibility to these groups. With fewer roadway segments to choose from and less accessibility, these roadway segments found in these areas are not being used by Strava users as much as other more connected roadway segments in the city.

The final variables considered in the model are looking into how well connected links are to areas with respect to the areas' commute times. From the regression analysis, it was seen that only the variables dealing with accessibility to areas with a commute time of 30 minutes or greater were significant. For the links that are well connected to areas with a commute time of 30 to 44, and 45 to 59 minutes, and 60 minutes and up, as the accessibility of a link to these areas increases, so does the likelihood that the link will be used as part of a cyclists' route. Areas with the higher

commute times are on the periphery of the city. Since the periphery of the city has the lower access and connectivity it seems counterintuitive that the high commute time areas would be the areas that increase the likelihood of using a link as part of a cyclist's route. Because these areas are often away from the shopping centers and other major areas of cities that attract traffic, the amount of congestion and traffic are lower, giving rise to easier conditions on the roadway for cyclists.

Qualitative Review

In order to fully understand where cyclists are choosing to ride in the City of Auburn, a qualitative analysis was also performed using GIS. To perform the visual analysis, the road network of Auburn was input into GIS, and then color coded based on the number of cyclists using a roadway. The roads were coded into four groups, which can be seen in Table 9 below. The numbers used for each of the color groupings were based on the percentiles of the highest number of users on a road segment, with the 0, 25, 75, and 100 percentiles represented.

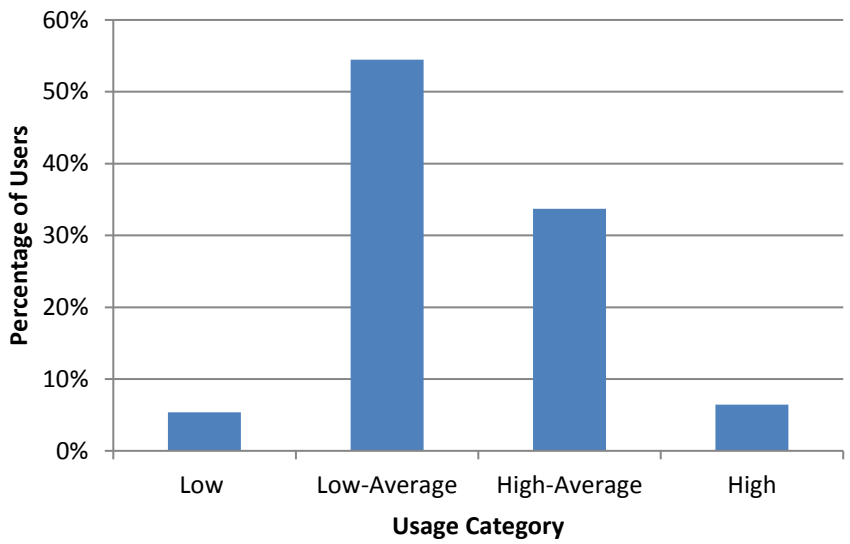
Table 9: Qualitative Analysis Street Colors

Number of Users	Street Color
0… Low	Red
1-40… Low-Average	Orange
41-118… High-Average	Green
199+… High	Blue

Along with color coding the streets, which allowed for easy and quick recognition of the heavily and little used streets, the streets were also given a ranking (1-4). The rankings were assigned to the roadways with a ranking of 1 corresponding to the roads with no use,2 with 1 to 40 users (25th percentile), 3 with 41 to 118 users (75th percentile), and 4 being the roads that had a

high amount of cyclist use, 119+ users. These rankings were placed into the attribute table for the roads layer in GIS, with the table then being exported to SPSS.

Figure 23: Qualitative Roadway Usage



With the attribute table entered in SPSS, the roadway characteristics could then be merged with the attribute table to form a final dataset with both the roadway usage rank and the characteristics of each roadway. To perform the final piece of the qualitative analysis, each of the rankings was selected, one at a time. The average, minimum, and maximum of roadway characteristics (volume, speed limit, lane width, etc.) were then determined for each of the roadway rankings. Along with the roadway characteristics, the LOS of the roadways was also evaluated, using the Bicycle Compatibility Index (BCI), to see if there was a significant difference between the four different usage groups. A total of 837 segments were included in the final dataset for this analysis, with the most segments being in usage groups 2 and 3. This process of selecting a ranking and then determining roadway characteristic averages allowed for the evaluation of the roads to see which of the physical characteristics of the roadway might have had an influence on whether a cyclists used them or not.

The City of Auburn has a growing bicycle path network with around 40 miles worth of bicycle lanes and paths located across the city. The city's bicycle network consists of not only on road bicycle lanes but also a mixture of off-road paths and multi-use pathways. As can be seen in the Table 10, bicycle lanes are the most common bicycle facility found in the City of Auburn, accounting for over half of the city's bicycle facilities. The next largest percentage is concrete multi-use paths, which allow for the use by both cyclists and pedestrians. The mileage of each facility type and percentage of the total bicycle network can be seen below in Table 10. The City of Auburn's bicycle path network is expected to grow with almost 114 miles worth of bicycle path and lanes proposed, with the proposed routes also being mapped in Figure 24. While information was available about the type of facility that is currently built within the City of Auburn, facility type was not available for the proposed bicycle facility routes, which were gathered from the City of Auburn Bicycle Plan.

Table 10: Auburn Bicycle Facilities

Facility Type	Mileage	% of Network
Bike Lane	22.03	56%
Off-Road Bike Path (Paved)	6.36	16%
Off-Road Bike Path (Unpaved)	1.87	5%
Concrete Multi-Use Path	8.60	22%
Multi-Use Lane	0.63	2%
Total	39.49	100%

The City of Auburn, AL Bicycle Facilities

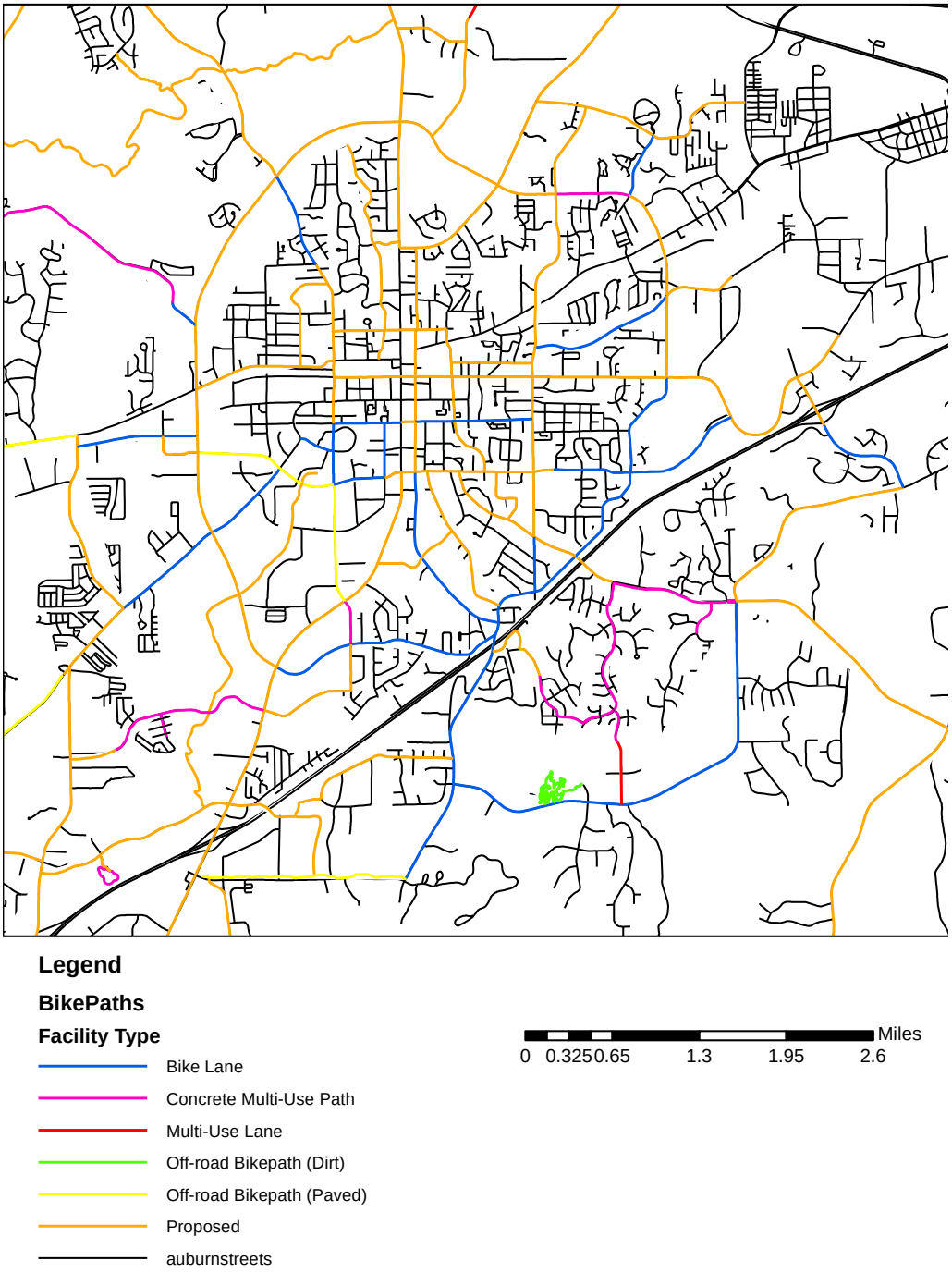


Figure 24: City of Auburn, AL Bicycle Facilities

Since only about 40 miles worth of bike facilities exist in the City of Auburn, that means the majority of cycling trips are being taken within the same stream of traffic as motor vehicles. In order to see which routes were the most suitable for cyclists, LaMondia and Moore (2015) looked into determining the suitability of collectors and arterials in the City of Auburn. The collection of surveys asked individuals how often they cycled and how they would classify themselves (Strong and Fearless, Enthused and Confident, etc.). Finally, the surveys asked the individuals to mark the location of routes within the city that they felt were suitable for cycling. The resulting suitability that LaMondia and Moore (2015) found of the roadways can be seen in Figure 25 below.

Figure 25: Auburn Suitability Map

From the above map, it can be seen that the most suitable areas, according to those who completed the survey, for cyclists are those that are near or on the property of Auburn University,

the shaded cross-hatched area. Along with the roads through the university property, those roads that traverse downtown Auburn, just to the northeast of the university area, are also deemed to be more suitable for cyclists. It can also be seen that the roads that are farthest from the city center are seen as less suitable for cyclists than those that are closer to the city center.

Along with the above analysis of the bikeability of the City of Auburn, AL, an analysis performed using GIS and SPSS saw that the four different usage groups had differences in roadway characteristics and bike facilities present. Looking at Table 11, the characteristics for usage group one, low usage, match up well with what is commonly seen on local neighborhood roads, with respect to peak hour volumes, lane widths, speed limits, and the presence of bicycle facilities. On local roadways typically there are lower traffic volumes, lower speed limits, wider lane widths, and few to no bicycle facilities. On the other hand, the higher usage groups match up well with roads of higher classification, such as a collector or an arterial. On these roads, speeds are higher than those found on local roads, lanes are the standard 11 to 12 feet wide, peak hour volumes are higher, and the presence of bicycle facilities. It is interesting to note that in Auburn, based on the table below, the higher used roadway segments correspond to some of the busier roads within the city.

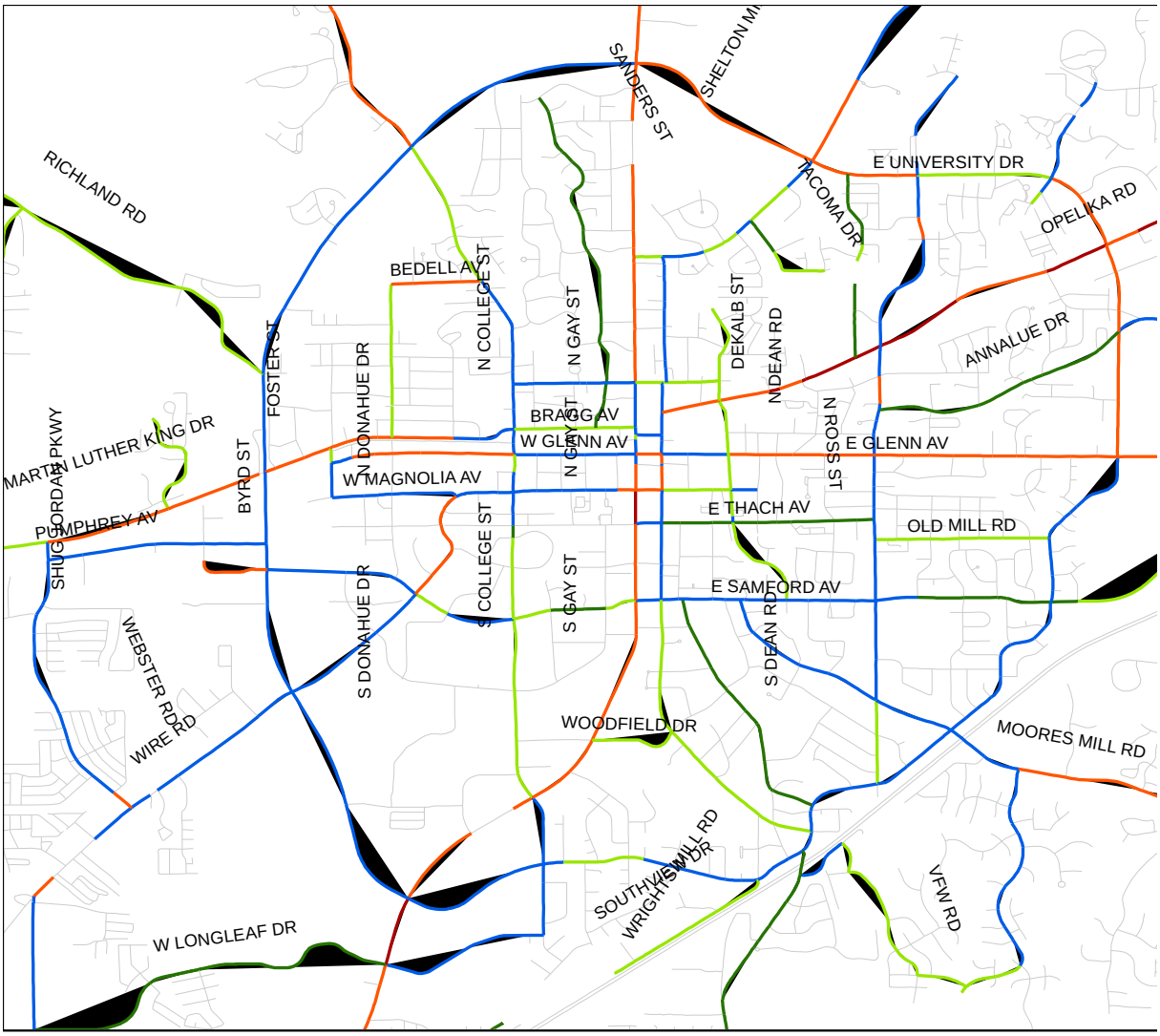
Table 11: Roadway Characteristics for Usage Groups

	Usage Group			
	Low	Low-Average	High-Average	High
	(n=45)	(n=456)	(n=282)	(n=54)
Street Characteristics				
Curb Lane Volume (vph)	350	2335	3102	2857
Number of Driveways	2.20	2.91	2.57	2.50
Pavement Condition	3.84	3.91	3.82	3.87
peak hour volume	35.00	233.46	310.17	285.67
Speed Limit (MPH)	28.78	33.75	35.66	40.00
Total # of Lanes	2.09	2.48	2.24	2.33
Total Volume	1500	6621	6818	6592
Width of outside lane	13.18	12.23	11.12	10.83
Width of paved shoulder	0.00	0.48	0.34	0.11
Bike Facilities Present				
Bike Lane	0%	11%	24%	44%
Multi-Use	0%	1%	4%	17%
Concrete Multi-Use	0%	0%	1%	0%
Off-Road	0%	1%	0%	0%
None	100%	88%	72%	39%

The biggest influencing factor for this increase use on the busier roads is that the percentage of segments in those groups that contain a bike facility also increases. With the roads that are not used, in Usage Group 1, the amount of bike facilities present is zero. This suggests that even if minimal, cyclists want to use roads that have some sort of bike facility. On the other hand, the group that had the highest use by cyclists, usage group 4, had a remarkable 61% of road segments with a bike facility present, even with smaller outside lanes, and higher traffic volumes.

As well as the presence of bike facilities and street physical characteristics, the bicycle level-of-service (BLOS) was also considered when looking into the four different usage groups. For this project, the Bicycle Compatibility Index (BCI) was chosen as the level of service measure to show the compatibility of streets to cycling. The BCI level of service measure uses geometric and operational conditions, such as presence of bike lane, speeds, and traffic volumes, to reflect the comfort levels of bicyclists that could potentially use the roadway. The Bicycle Compatibility Index has a level of service range of A through F with A being the best and F being the worst (hrsc.unc.edu).

From the map below of the City of Auburn, it can be seen that the roads in the city widely range from LOS B to LOS F, with the majority being a LOS C or D. Not surprisingly, the roads that have the most traffic, and go toward the shopping areas in town, Opelika Rd and College St., have lower LOS of E/F. The road segments that have the higher LOS of B and C tend to be in more residential areas where speed limits and traffic volumes are lower, along with areas that have bike facilities present.



Legend

<all other values>

BCI

- B
- C
- D
- E
- F
- N/A
- Auburn Streets

Miles

0 0.2250.45 0.9 1.35 1.8

Figure 26: Auburn Streets BLOS: Bicycle Compatibility Index

Seen in Table 12 below, the highest usage group experienced the highest percent of roadways with an LOS of B and C, with the other LOS levels being about the same as the other usage groups. This shows that as the level of service of a road improves, with respect to bicycle compatibility, the more use the road will see from cyclists. While this holds true for usage group four, usage group three saw a significant number of roads fall into a level of service D, meaning that less experienced cyclists may not feel as comfortable using these roads as compared to more experienced riders.

Table 12: Bicycle Level of Service by Usage Group

	Usage Group			
	Low	Low-Average	High-Average	High
	(n=45)	(n=456)	(n=282)	(n=54)
Bicycle Compatibility Index…				
…A	0%	0%	0%	0%
…B	0%	14%	8%	24%
…C	38%	31%	21%	24%
…D	9%	27%	51%	31%
…E	0%	20%	18%	20%
…F	0%	3%	0%	0%
N/A	53%	5%	2%	0%

The presence of parallel roads also presented another unique aspect to consider when looking into the routes that cyclists chose. Using GIS, it can be seen in Figure 27 that S College St., and S Gay St. are in parallel and W Magnolia Ave. parallels W Glenn Ave. While spaced further away, S Donahue Dr. also provides another choice for the north/south route options. In the set of Donahue Dr., College St., and Gay St, the most used of these three can be seen to be Donahue Dr., followed by Gay St., and then College St. This is not surprising in that Donahue Dr. contains an off-road paved bike path that allows for an additional separation from traffic that both College

St. and Gay St. do not provide. While Gay St. does not have an off-road path, it does contain a bike lane for some of its segments, and it has a lower amount of traffic as compared to Donahue Dr., and College St.

With the higher amounts of vehicular traffic and higher speed limits, College St. is not used as often as its two neighboring north/south route options. As the main roadway from the interstate to downtown Auburn, College St. receives the majority of the traffic attempting to traverse through the city, compared to Gay St., and Donahue Dr. The lack of bicycle facilities along College St. is also a deterrent factor to cyclists as they have to mix in with the vehicle traffic along a roadway with higher speed and higher volumes. While College St is used as a major thoroughfare through Auburn, Donahue Dr. and Gay St. have more connections to residential areas, more specifically the student populations that live toward the southern parts of the City of Auburn. This connection to those populations is important as it gives those users a more directly connected route that does not involve as many detours to avoid less desirable streets.

Along with the parallel north/south routes, there are a couple of parallel streets that run east/west through the City of Auburn. Most notably there are Samford Ave., Magnolia Ave., Glenn Ave., and Thach Ave. These four roads serve a number of student residential areas as well as provide routes that traverse the heart of the city. As can be seen in Figure 27, Samford Ave. and W. Magnolia Ave. are the two routes that receive the most use, with Samford Ave. seeing a steady amount of traffic over the segments present in this map. These two streets are the most used as they represent the two streets that can be used to move from one side of the campus if Auburn University to the other side. While the two streets are on the high end of the usage rank, there are differences between them. Unlike Magnolia Ave., Samford Ave. does have bike lanes present from the intersection near College St., all the way to Shug Jordan Pkwy. on the west side of town. This

allows for the cyclists using this stretch of roadway to have a dedicated lane outside the main vehicular traffic flow, unlike on Magnolia Ave. where the cyclists have no dedicated space and ride in the lane with the vehicles present. Because of these dedicated spaces for the cyclists on Donahue Dr., and portions of Gay St., the BLOS for these roadways is also higher, with both averaging a C/D LOS whereas College St is around a E/F, indicating that the comfort level for the cyclists on those roadways is higher. This increased comfort level on Donahue Dr. and Gay St. can also be leading cyclists to use them over College St., and other similar routes in the area.

It is also interesting to note that for the portions of the roads east of Gay St., the lower volume roads, such as Samford Ave., and Thach Ave. are being selected over the higher volume and higher speed roads such as Glenn Ave. Since Magnolia Ave. on the eastern side of Gay St. only continues for a couple of blocks before terminating, cyclists are choosing other road options that provide the necessary connections, such as Thach Ave. which continues to Dean Rd. This switching of roads is interesting and shows that while Magnolia Ave. on the east side of Gay St. has a comparatively high LOS, due to it not being connected to where cyclists want to travel, the cyclists are choosing to switch to other roads that can provide that connection. Along with the connection to Dean Rd., Thach Ave. also provides a route with a high LOS for cyclists, an LOS level of B.

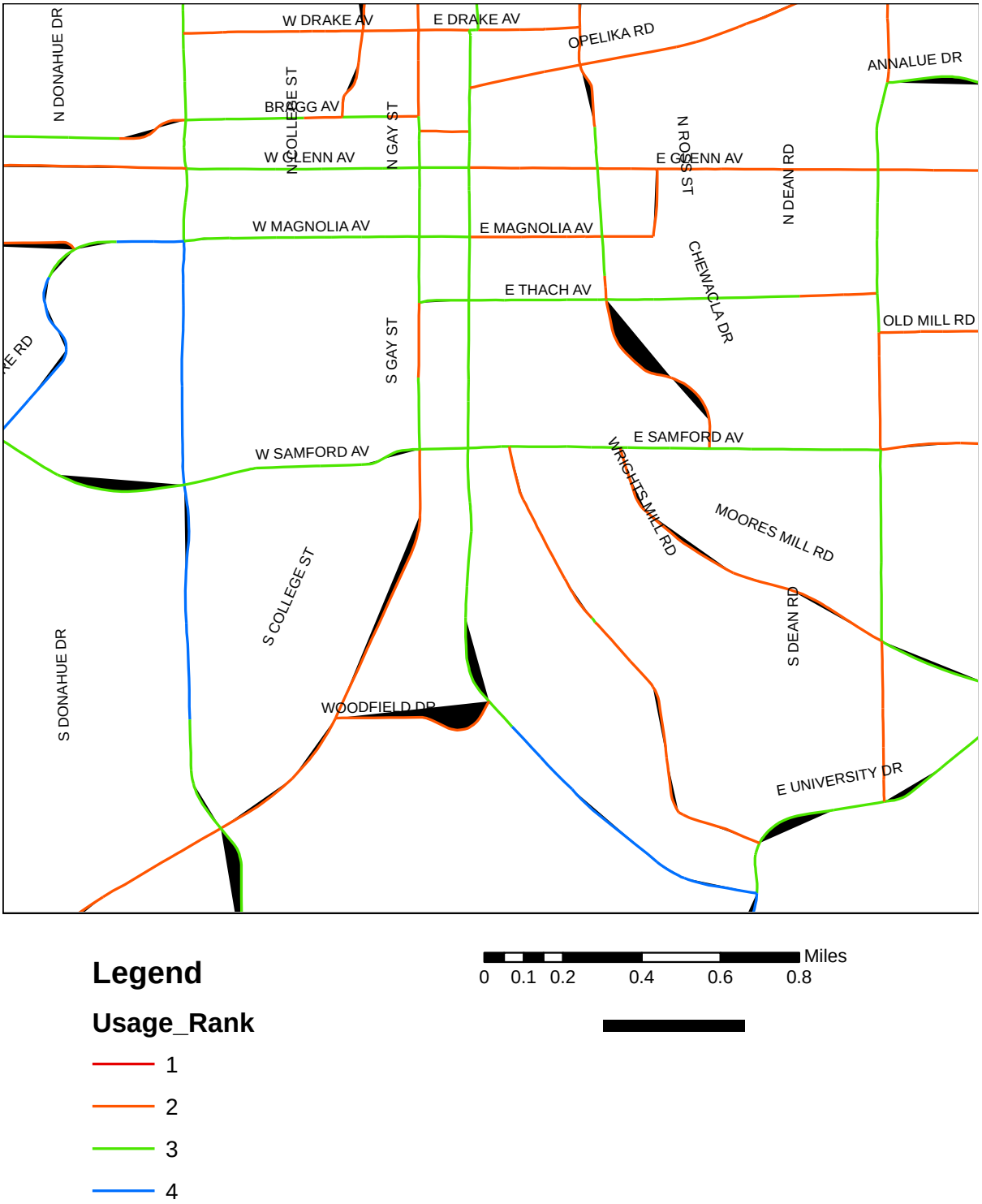


Figure 27: Auburn Street Usage- Parallel Facilities

URBAN APPLICATION ANALYSIS

Modeling Methodology

For the purpose of this research, the alternative to the chosen route was taken to be the shortest route path generated by the A-star algorithm. The binary logistic choice that was modelled was whether the riders chose the shorter of the two routes depending on their age, gender, and what type of rider they are. Additional regression models were also constructed to understand (i) the relationship between trip length and rider characteristics and (ii) the percent deviation of the chosen route from the predicted shortest route based on rider characteristics.

Three data sources were used to create the road network map. The Atlanta Regional Commission's street network shapefile (RC_ROUTES) was obtained from the travel demand modeling group of Atlanta Regional Commission (ARC). It is a modified version of the roadway database maintained by the Georgia Department of Transportation (GDOT) and focuses on state managed roadways rather than locally managed roadways and bikeways. However, it contains the most comprehensive inventory of roadway characteristics like speed limit, annual average daily traffic (AADT), number of lanes, truck volume, etc. which are useful information for route choice modeling at a later stage. The second data source used was Open Street Map's (OSM) bicycle map for Atlanta. The OSM map has local roads and locally managed facilities which were not present in the RC_Routes map. The two maps were spatially joined based on a buffer distance to get a more complete map of the road network of Atlanta. The resulting map was then cleaned for non-bicycling facilities like freeways. The final data source was the Metro Atlanta Bicycle Facility Inventory. The location of on street parking on roadways with conventional bicycle lanes and buffered bicycle lanes was manually coded in ArcGIS using Google Earth imagery. The treatment of intersection approaches with right turn only motor vehicle lanes that connect to links with

conventional bicycle lanes, buffered bicycle lanes, or protected cycle tracks were also manually coded in ArcGIS using Google Earth Imagery. As a final measure, the trips were plotted on the map and checked for links traversed by cyclists but missing in the network. Such links were manually added where more than 2 bicycle trips were found to follow a path, but the path was not marked as a link in the network. This was assumed to be mainly because of tendency and ability of bicyclists to use cut-throughs and private alleys which are not marked in regional network maps. However, shortcuts through parking lots were not added as links although there were multiple such cases.

Figure 28(a) shows the number of trips recorded by each rider. Figure 28 (b) shows the trips by purpose. Figure 28(c) shows the trip purpose across age – riders in the age group > 45 years use cycling for exercise than any other group. Figure 28(d) shows trip purpose by gender, and since the data are heavily dominated by male cyclists, they are the dominating group in all trip purpose categories. However, female riders have almost a similar share of shopping trips in spite of being a small fraction of the riders. This calls for particular consideration in land use planning to allow women to do trip chaining comfortably and easily. Figure 28(e) shows trip purpose by rider type. The strong and fearless riders make more social and shopping trips by cycling than other types of riders, while enthused and confident riders using cycling for commute more than any other rider type. Comfortable but cautious riders use cycling for exercise more than other rider types. Figure 28(f) shows the frequency distribution of trip length. The mean trip length, marked by dashed red line was found to be about 3.75 miles (about 5.5 km). The majority of the trip lengths were within 4-6 miles which is a standard commute distance. Figure 28(g) shows the trip length by age. It should be noted that the highest frequency of trips for younger riders are at a shorter distance than that of senior riders which is initially counter intuitive. However, one of the reasons

may be that senior riders are less likely to choose shorter routes if that does not provide sufficient safety and comfort while younger riders may prefer shorter distance to a detour for a bike facility. The younger riders are also dominated by college students, and their commutes may be much shorter in length. Figure 28(h) shows trip length across gender, and we see a similar trend as age here – the highest frequency of trip lengths for women are longer than that of men. Figure 28(i) shows trip length across rider type, and enthused and confident riders are seen to have shorter trips than comfortable but cautious riders. Strong and fearless riders have slightly longer mean trip length than enthused and confident riders, but that may be because they bicycle longer distances.

For the purpose of this study, we considered only the "primary" trip of each user and therefore restricted the analysis to work or school trips (trip purpose = "work"/ "school"), thus reducing the number of trips to be considered from about 20,000 to about 12,000. The trips were further restricted to be greater than 1 mile and less than 8 miles, resulting in about 10,000 trips.

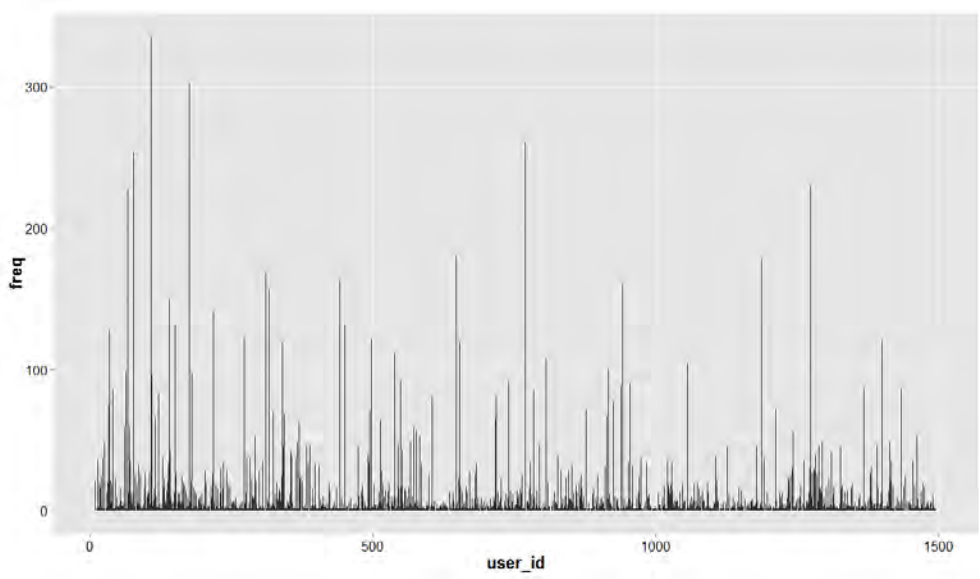


Figure 28(a) Number of Trips Recorded by Users

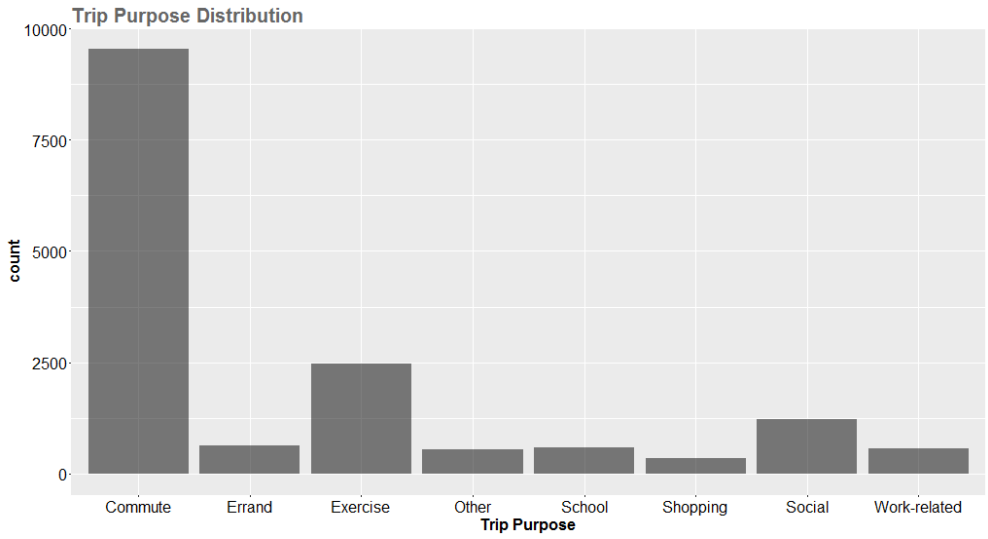


Figure 28(b) Trip Purpose Distribution

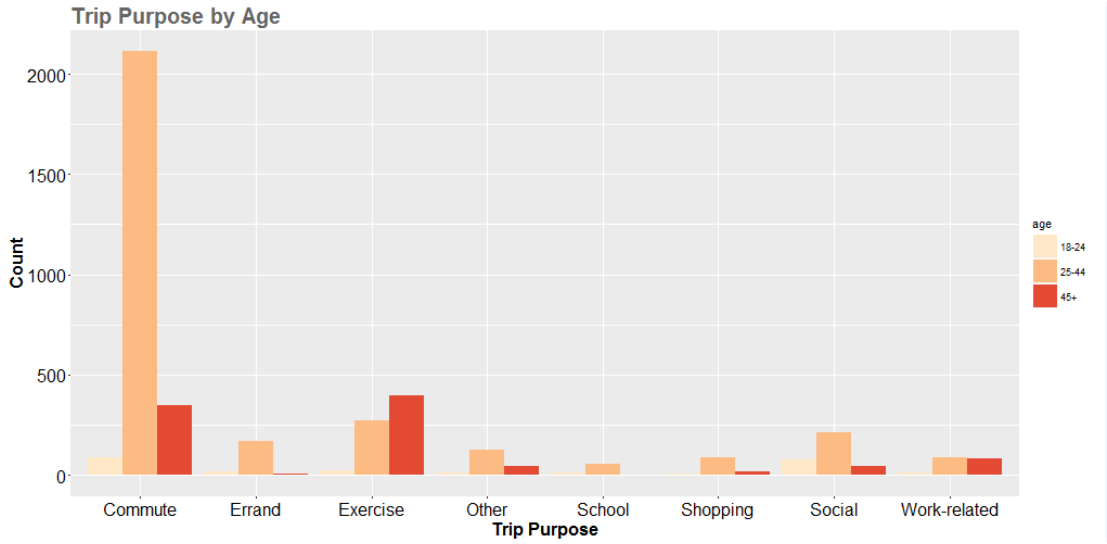


Figure 28(c) Trip Purpose Distribution across Age

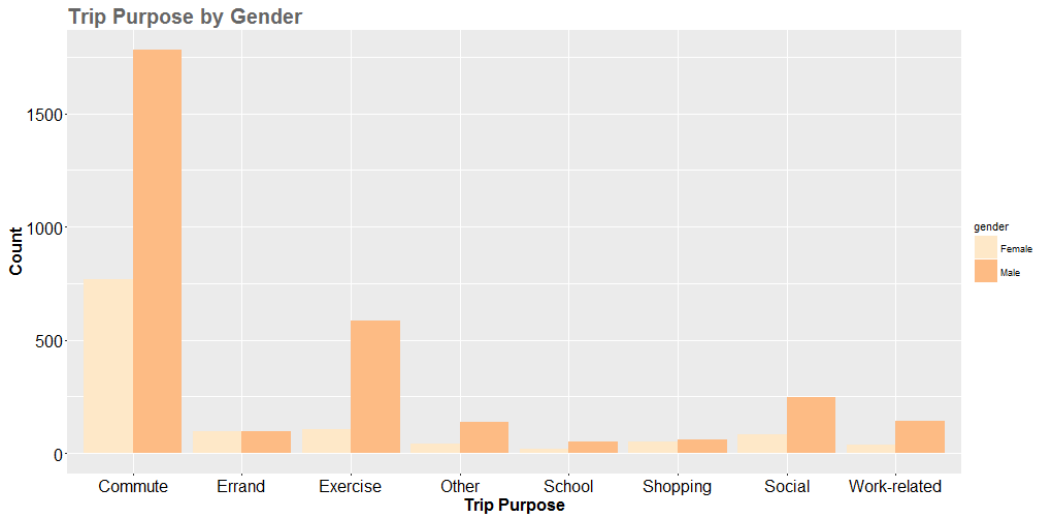


Figure 28(d) Trip Purpose Distribution across Gender

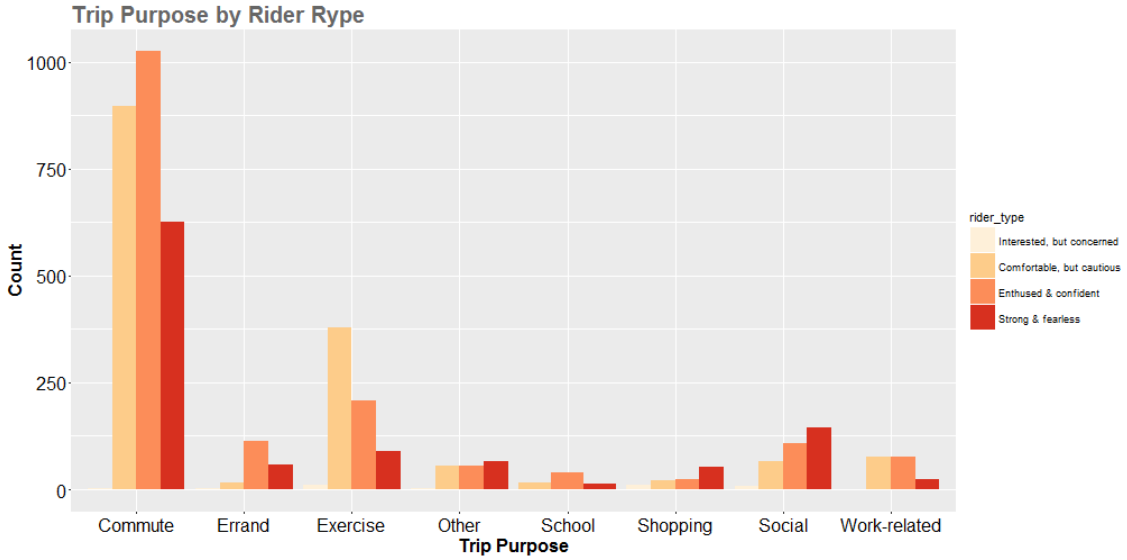


Figure 28(e) Trip Purpose Distribution across Rider Type

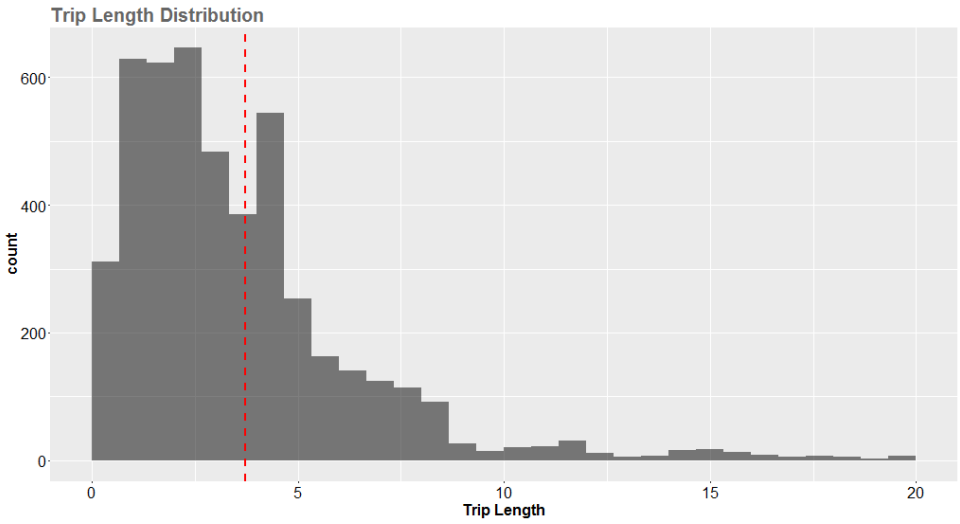


Figure 28(f) Trip Length Distribution

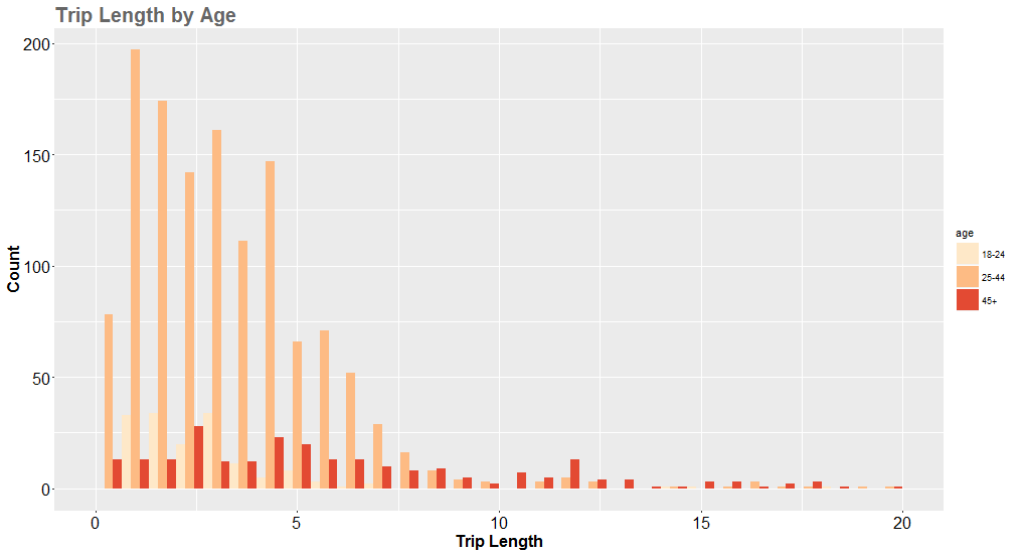


Figure 28(g) Trip Length across Age

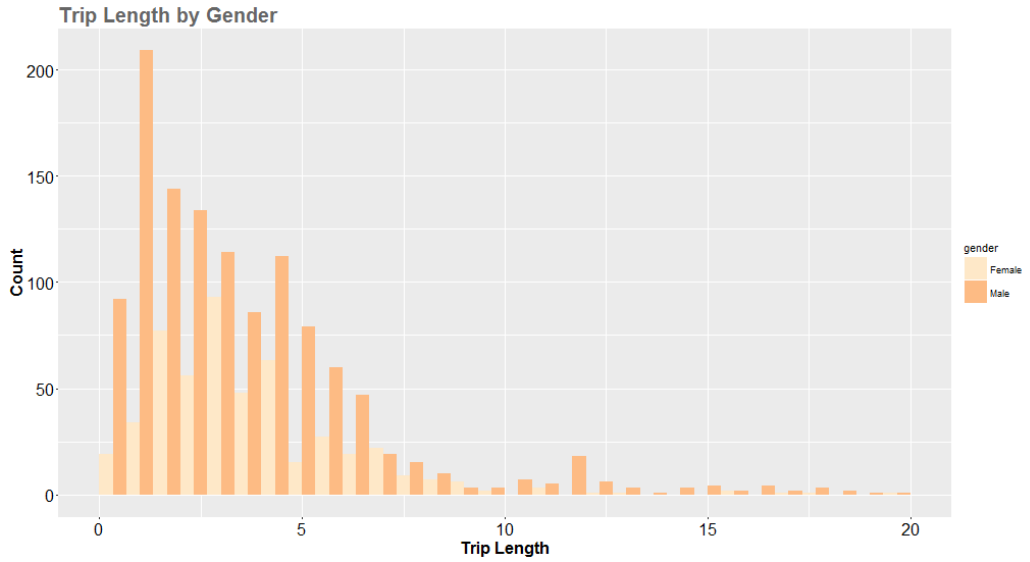


Figure 28(h) Trip Length across Gender

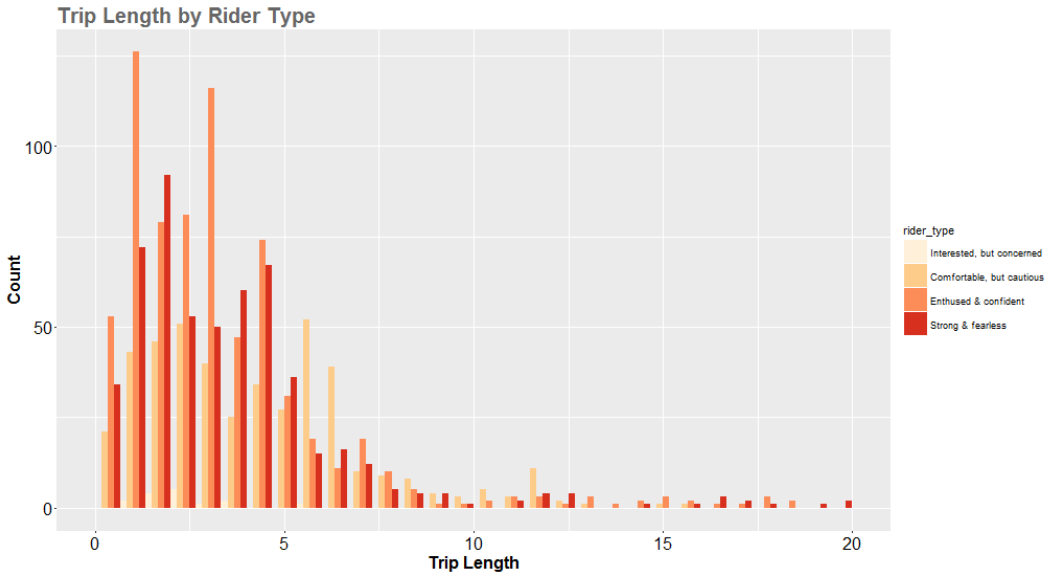


Figure 28(i) Trip Length across Rider Type

Figure 28: Cycle Atlanta Trips (a) Number of Trips Recorded by Users (b) Trip Distribution by Purpose (c) Trip Purpose Distribution across Age (d) Trip Purpose Distribution across Gender (e) Trip Purpose Distribution across Rider Type (f) Trip length Distribution

Analysis and Results

One random trip was then chosen per user to perform the shortest route analysis. The first model estimated is a linear regression model to understand the relationship between trip length and age, gender, and rider types. For all the models, the *comfortable but cautious & interested but concerned* group was chosen as the base group as was age 18-24 implying that all results should be interpreted in a comparison to that category. Table 13 presents the results of the regression model on trip length as function of sociodemographic characteristics of the riders. Age has a positive relationship with trip length and male riders are also more likely to ride longer distances. *Enthused and confident* riders are less likely to take longer trips than *comfortable but cautious* riders, but *strong and fearless* riders are more likely to take longer trips. This may be because *enthused and confident* riders are more inclined to use shortest routes even if there are no bicycle facilities which renders their trip short compared to *comfortable but cautious* riders. On the other hand, *strong and fearless* riders are more likely to naturally undertake longer trips than any other categories.

Table 13: Trip Length as Function of Socio-demographic Characteristics

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)	Significance
(Intercept)	2.4334	0.17233	14.12	< 2e-16	***
age25-44	0.96803	0.15798	6.128	9.67E-10	***
age45+	2.37601	0.1906	12.466	< 2e-16	***
genderMale	0.10798	0.10187	1.06	0.289	
Enthused and confident	-0.52289	0.08236	-6.349	2.38E-10	***
Strong and fearless	0.04501	0.07	0.643	0.52	

*Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*

The second model was used to understand the relationship between deviations from the network based on the shortest route depending on their socio-demographic characteristics. For the majority of the trips, the network based shortest route is shorter than the actual trip length, therefore this model may serve as a proxy to understand if any rider group is systematically choosing a longer route possibly because of factors not yet known to us. Table 14 presents the results of the model estimate. Gender is only significant in this model and male riders are less likely to deviate from shortest route as compared to female riders. Similarly *enthused and confident* riders and *strong and fearless* riders are also less likely to choose longer routes over shortest routes as compared to *comfortable but cautious* riders, with *strong and fearless* riders more likely to choose shortest routes than *enthused and confident* riders. People in the age group >45 are less likely to choose the shortest route than riders in the age group of 18-24 while people in the age group of 25-44 are more likely.

Table 14: Deviation from Network based Shortest Route as Function of Socio-Demographic Characteristics

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)	Significance
(Intercept)	6.7953	8.9253	0.761	0.4465	
age25-44	-0.1981	8.1843	-0.024	0.9807	
age45+	2.3757	9.8707	0.241	0.8098	
genderMale	-9.3924	5.273	-1.781	0.0749	.
Enthused and confident	-2.4476	4.2633	-0.574	0.5659	
Strong and fearless	-5.2741	3.6233	-1.456	0.1456	

*Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*

Finally, a binary logistic choice model was used to understand whether a rider chose a shorter of the two available alternatives depending on the socio-demographic characteristics. The model estimates are presented in Table 15.

Table 15: Choice of Shorter Route Based on Socio-demographic Characteristics

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)	Significance
(Intercept)	0.618	0.0303	20.407	< 2e-16	***
age25-44	-0.072	0.0278	-2.593	0.00956	**
age45+	-0.150	0.0335	-4.486	7.42E-06	***
genderMale	-0.012	0.0179	-0.681	0.4956	
Enthused and confident	0.058	0.0145	3.989	6.73E-05	***
Strong and fearless	0.007	0.0123	0.595	0.55207	

*Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*

The results show that senior riders are more likely to not choose shortest route while *enthused and confident* riders as well as *strong and fearless* riders are more likely to choose shortest routes. However, male riders are also more likely not to choose shortest routes, which is counterintuitive. This may be because in general, male riders undertake longer trips, and hence, there is not much difference from the shortest route and the actual route. Another model was estimated by including trip distance to see if distance is a reason for the counterintuitive sign of this model. The results are presented in Table16. With the introduction of trip length, age loses its significance indicating that trip length is related to age of a rider. However, trip length is significant and has a negative sign indicating that longer the trip is, riders are less likely to choose shortest routes possibly because either the difference is not significant enough or because longer trips require being comfortable for a longer time and people are more likely to choose facilities that maximize that perceived comfort.

Table 16: Choice of Shorter Route Based on Socio-demographic Characteristics and Trip Distance

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)	Significance
(Intercept)	0.7549	0.0293	25.7620	<2e-16	***
age25-44	-0.0179	0.0264	-0.6770	0.4986	
age45+	-0.0171	0.0322	-0.5300	0.5959	
genderMale	-0.0061	0.0169	-0.3630	0.7169	
Enthused and confident	0.0283	0.0138	2.0590	0.0396	*
Strong and fearless	0.0098	0.0116	0.8430	0.3995	
Trip length	-0.0562	0.0025	-22.7850	<2e-16	***

*Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*

CHAPTER 5: DEVELOPING ROUTE SUITABILITY SCORE AND CYCLIST SELF-IDENTIFICATION MODELS

This chapter combines the findings from the previous chapters to develop two tools for assisting in network improvement prioritization: a route suitability score model and a cyclist self-identification model. Both models will assist planners and decision-makers in selecting which roadway segments should be improved.

ROUTE SUITABILITY SCORE MODEL

Dataset Generation

During prior research to examining alternate bicycle LOS measures, the research team introduced the idea of a Suitability LOS measure. The research team conducted a survey of perceived bike route suitability in Auburn, AL. The goal of bicycle LOS measures is simply the determination of whether a roadway is suitable for cycling. The goal of this survey was the same, but by substituting roadway characteristics for public opinion. The essential details of the survey will be outlined here. For a thorough description of the survey and its results, see (0).

Surveys were conducted in the two locations in Auburn that had the busiest weekday foot-traffic, as well as distributed to participants of the Auburn Bicycle Committee and the East Alabama Cycling Club. The public surveys were conducted over multiple weeks on multiple weekdays between the hours of 9am and 3pm. Respondents were comprised mostly of students, faculty, staff, and other city residents. A total of 565 complete and useable surveys were collected and used for analysis.

Along with indicating on a map which roadways they thought were suitable for cycling, participants in the survey were asked to indicate both how often they cycled as well as how

comfortable they considered themselves in cycling. These questions were categorical, giving the cyclist one of five options to pick from, mimicking a Likert-like scale, and were included with the intention of better understanding what cyclists of all experience levels consider important suitability factors.

Respondents indicated on a map by highlighting or circling what cycling routes they considered suitable for cycling. These results of the survey were summed in ArcGIS by segment and cyclist type. Roadway segments were then given a LOS score based on the number of respondents that named that segment suitable. The results were scored using natural breaks to separate the segments into five naturally-forming categories. To mimic LOS measures, these categories were labelled "A" through "E/F." The lowest scoring segments, i.e., segments that the fewest number of respondents deemed suitable, were given a LOS score of "E/F," while the highest scoring segments received a LOS score of "A." The natural breaks occurred at 0-25, 26-55, 60-107, 108-168, and 169-247 (the max).

Modeling Methodology

In order to determine what factors beyond physical roadway characteristics were affecting suitability LOS score, accessibility measures, including access to specific land use and different sociodemographic factors were collected. Land use and sociodemographic data were obtained from the US Census database and formatted in an ArcGIS spatial database. All types of land use were considered, including educational areas, government areas, office space, parking, commercial space, and restaurants. Sociodemographic data included information such as household makeup, household income, commuting distances, and number of vehicles in a household.

In order to use this information in a LOS analysis, all of this information on the surrounding land needed to be condensed down to a per roadway segment basis. From a LOS standpoint, the access a roadway provides to these surrounding factors is what is important. To accomplish this, a gravity model function was applied. This function considers how well a specific roadway segment was connected to a specific access measure based on its distance to areas that fulfill that measure. The function used to obtain the accessibility of a specific link can thus be written as

$$A_n = \sum_i x_i d_{in}^{-1.5}$$

Where

n = roadway segment ID
A_n = access measure for link n
i = zone ID
x_i = land use factor in zone i (Household Members: Age 18-24, etc.)
d_{in} = average distance from link n to zone i

As is standard in gravity model functions, distance is raised to a power of -1.5 so that land use factors that are further away from a the segment count for less. Consider, for example, a roadway segment was very close to multiple areas with a high population of household members age 18-24, but most restaurants are located miles away on the other side of town. This segment would receive a high accessibility score for household members age 18-24 because this link is well connected to those areas, but a very low accessibility score for restaurants, because it provides a poor connection to areas with restaurants. These scores are unitless and only have meaning for comparison purposes. However, these access measurement scores provide a means to evaluate the impact of land use and sociodemographic characteristics on bike LOS when used in a model.

Analysis and Results

The final ordered probit model identified three key variable types: Roadway Characteristics, Land Use/Activity Accessibility, and Sociodemographic Accessibility. This model creates an index using data from Auburn, AL that can be used to determine the suitability of a roadway for cycling, giving that roadway score that translates into a BLOS grade from A to F. The results of the ordered probit model can be seen below in Table 17.

Table 17: LOS Suitability Index

Variables	β	t-stat
Suitability LOS Thresholds		
Between D/F (γ_1)	1.637	3.01
Between C/D (γ_1)	2.993	5.43
Between B/C (γ_1)	4.432	7.86
Between A/B (γ_1)	5.379	9.25
Roadway Characteristics		
Speed Limit	-0.081	-7.21
Bike Facility Identifier	1.385	7.47
Barriers Present in Bike Lane	-1.722	-2.80
Total Number of Lanes	0.705	6.58
Land Use/Activity Accessibility		
Access to…		
Community Areas	0.086	1.91
Governmental Areas	0.670	2.03
Health Care Facilities	-10.727	-4.49
Mixed Development	-40.192	-6.24
Residential Areas	0.071	1.75
Restaurants	38.073	4.66
Sociodemographic Accessibility		
Household Members Age 25-39	-4.28E-04	-4.55
Typical Commute:		
Less than 10 miles	3.05E-04	3.30
20-29 miles	6.73E-04	1.84
30-44 miles	-2.00E-03	-3.09
More than 60 miles	-2.53E-03	-4.93
Annual Household Income:		
30k-59k	1.39E-03	2.79
More than 100k	-8.96E-04	-2.20

Roadway Characteristics

As shown in Table 17, speed limit is negatively correlated with cycling. This follows conventional wisdom; as the speed limit of a roadway increases, it becomes less desirable for cyclists to use because cyclists perceive fast vehicles as a great risk to cycling. The presence of bike facility identifiers increases the BLOS score of a roadway. People prefer to take routes that are already identified as cycling routes, likely because the presence of those facilities make the road seem safer for cycling. Roads with barriers to cycling, such as on-street parking, are less desirable to cyclists.

The final significant roadway characteristic in this model is number of lanes, which surprisingly shows a positive correlation with BLOS score. While one might think that an increased number of lanes would be a turn-off to potential cyclists due to the increased traffic and thus, increased risk, the model results say the opposite. Roads with more lanes are actually preferable to cyclists. This means that people prefer to cycle on the 'main' routes, which supports the idea that cyclists are cycling to get somewhere; they are cycling in order to reach a destination, not just for recreation. Roads with more lanes are usually the main roads, which typically are highly connected and accessible.

Land Use/Activity Accessibility

The Suitability LOS Index determined that roadways that are highly connected to important origins and destinations also play a significant role in determining the suitability of a roadway for cycling. Community areas have a positive correlation with BLOS because these areas are hot destination spots, especially ones that are typically associated with being active, such as parks and trails. Government areas are also positive, though this likely has more to do with their centralized

downtown location than being important destinations themselves. Roads with access to health care facilities or mixed use development, however, decrease BLOS score. This means that these areas are unimportant destinations for cyclists. Such areas may also be associated with heavy traffic volumes that are unattractive to cyclists. Residential areas have a positive correlation to BLOS. Many trips originate at home, so good access to such areas is important to cyclists and increases the suitability of such roads. Finally, roads with access to restaurants have a highly positive correlation to suitability due to these areas being attractive destinations.

Sociodemographics

The final variables included in the ordered probit model attempt to answer the question of *who* it is important to provide access to suitable cycling. Using data from the US Census, the research team determined what populations were highly connected to roads deemed suitable for cycling. The results determined that a relative few sociodemographic variables were important to consider.

The only significant age variable was household members aged 25-39 and had a negative correlation to BLOS. Access to cycling is unimportant for population centers in this age group. These young adults may live further from the town center, as this is primarily a college town.

Commute distance was also considered, but this commute is not necessarily associated with cycling to work. Connections to people who commute less than 10 miles to work have a positive correlation. Many of these people likely live closer to the downtown center, with higher access to destinations close to home. Roads highly connected to those who commute 20-29 miles to work are also positively correlated with cycling. These people may live in neighborhoods suitable for cycling. However, roadways connected to homes where people make longer commutes of 30-44

miles and more than 60 miles are negatively associated with BLOS scores. Access to these origins is unimportant to cyclists, as the people living in those areas are unlikely to be cyclists themselves.

The final variable considered by the model was annual household income. Most income brackets were unimportant to the model, leaving only two extremes. Access to areas with annual income between $30k-59k has a positive correlation with cycling suitability while access to areas with high incomes of more than $100k has a negative correlation. This means that areas in the lower income bracket are more likely to cycle than those in the high income bracket. Access for areas in the high income bracket is less important to cyclists than access for those with lower incomes.

CYCLIST SELF-IDENTIFICATION MODEL

Dataset Generation

The analysis uses the data collected through the Cycle Atlanta smartphone application, developed through a collaboration between the Georgia Institute of Technology and the City of Atlanta's planning office to promote cycling in Atlanta (The City of Atlanta, 2011).

The application was named Cycle Atlanta after the larger planning project for which the application was initiated, and was developed by an interdisciplinary team of researchers. The application was originally based on San Francisco's CycleTracks (Hood et al. 2011), although Cycle Atlanta was substantially updated to make better use of current features available in iOS and Android as well as to include features that the City and local bicycle advocacy groups wanted in the application. The basic feature is trip recording, where the application uses the GPS of the phone to record the location of the user once per second. In addition to tracking cyclists' trips, the app also provides options to enter personal information, including age, email address, gender, ethnicity,

home income, zip codes (home, work, and school), cycle frequency, rider type, and rider history (Misra et al. 2014).

The breakdown of age, gender, income, and ethnicity was kept similar to the breakdown as found in the household travel survey. The age and income intervals as well as the gender and ethnicity subcategories were adopted from the household travel survey conducted by Atlanta Regional Commission. The rider type and rider history categories are exclusive and unique to the design of Cycle Atlanta. The cycling experience field allowed users to specify how long they have been cycling and can choose from the categories 'since childhood', 'several years', 'one year or less' and 'just trying it/just started'.

As of June 2014, the Cycle Atlanta dataset consisted of 1529 unique users who could provide information on their age, gender, ethnicity, income, rider history and cycling frequency. Because there were only 6 cases in the age group of 65+, that group was merged with the age group of 55-64 years old and the new group is referred to as "age 55+" for the rest of the analysis. About 60% of the riders provided information on each of the socio-demographic categories. The users of Cycle Atlanta are predominantly male (about 75%), white (about 80%) and mostly from a high income group (>$75,000) (about 45%). The median age of the users is between 25-34 years, while the median income is between $60,000 and $74,999. The median rider type is an *enthused and confident* rider with median cycling frequency of several times per week and a median riding history of several years.

Two main types of variables were used in these models – the socio-demographics and the riding habit/pattern of the participant. The socio-demographic variables included age, gender, income and ethnicity while the riding pattern variables included cycling frequency and rider history. From the distribution of age and gender across rider type, it was evident that there were

very few participants in the age group above 45. So the age groups 45-54, 55-64 and 65+ were grouped into one category of 45+. The riding pattern was found to be distinctly similar across the age group of 25-34 years and 35-44 years and hence, these two groups were also merged to form a new group of 25-44 years. Similarly, different income categories were consolidated into 3 categories and different ethnicity types were consolidated into 4 categories. For rider history, the 'just started' category was merged with the 'less than a week' category, resulting in 3 categories instead of 4.

Of the total 989 users who provided data on rider type, only 26 users classified themselves as *interested but concerned*. Cross tabulation of rider type across other variables showed *interested but concerned* riders having zero cell values with cycling frequency 'less than once a month' and small valued cells for age group 45+ (2 users) and ethnicity 'African American'(1 user) and 'Other'(1 user) thereby presenting a problem of quasi separation. Within cycling frequency also, there are only 13 users who have cycling frequency less than once per month and none of them are *enthused and confident* riders (0 users) which again presented the issue of separation. Quasi/complete separation implies a perfect prediction scenario where the dependent variable Y can be completely predicted by variable X when the separation is complete. In case of quasi complete separation, perfect prediction happens only for a subset of observations (Albert and Anderson 1984). For example, in this dataset, it can be predicted with absolute certainty that none of the riders who bicycle less than once per month will classify themselves as *enthused and confident*, although the same cannot be said about whether riders with cycling frequency less than once a month will classify themselves as *strong and fearless* or *comfortable but cautious*. Models estimated under quasi/complete separation are more likely to either not converge or give high co-efficient estimates and infinite standard error as the log-likelihood will be presumably flat (Zorn

2005). The most common way of dealing with quasi separation is to remove the problematic covariate which again might give specification bias if the covariate is strongly correlated. We ran models both by removing observations and by aggregating the sparsely populated group with its nearest neighbor. In case of cycling frequency, the last group, cycling frequency less than once per month was merged with the group which bicycles a few times per month and the new group was named cycling frequency once or less per week. Models ran by removing the observations with cycling frequency less than once per month gave a much lower model fit than the aggregated models and hence, in this paper, models with aggregated data is presented. Similarly, for addressing the quasi separation problem related to rider type, two alternative model sets were designed – one where the *interested but concerned* group (26 users) was merged with its next higher group *comfortable but cautious* (333 users) and another where the *interested but concerned* users were removed from the sample space and models were estimated for the remaining three categories. The model estimates in either case were not significantly different and in keeping with our aggregation theme, in this paper, the aggregated models are presented.

Analysis and Results

Based on these rider type distributions, logistic regression models were estimated for each rider type to understand how the self-described confidence level is affected by socio-economic variables as well as riding patterns of the cyclists. Several logistic regression models were explored to find the best way to represent the pertinent relationships. Since cycling frequency and rider type may have bi-directional causality, they were tested for explanatory power and likely association. A single variable ordinal model for rider type with cycling frequency as the explanatory variable gives a McFadden's ρ^2 of 0.48 but an ordinal model for cycling frequency with rider type as the explanatory variable gives a McFadden's ρ^2 of 0.07 (both unadjusted for sample size difference).

Although it was found that cycling frequency has a greater explanatory power for rider type than rider type has for cycling frequency, in view of the simultaneity issue, models with cycling frequency and models without cycling frequency are both presented here.

Since the discrete observed rider type categories (y) were originally thought of as representing a latent continuous scale of confidence and comfort (y^*), two variations of the user's underlying decision process along that one dimensional scale were initially estimated. The first is where the self-classification process was thought of as representing a binary choice for each rider type (for example "Am I *strong or fearless* or not?"). This process was estimated using binary logistic regression models where the rider classifies himself/herself into a category ($y = 1$) if he/she perceives himself/herself above a certain confidence level threshold $(y^* > \tau)$; if the perceived confidence level is at or below the threshold $(y^* \leq \tau)$, the rider does not choose that rider type category ($y = 0$). Four different binary logistic models were estimated – one for each rider type. For each of these four choices, several models were run with different variable combinations to balance model fit and parsimony. Age group 45+, gender male, income less than $40,000, rider history since childhood, and cycling frequency of daily were chosen to be the base categories for age, gender, income, rider history and cycling frequency variables. Ethnicity was not included in the models due to its heavy bias towards white riders. Model fit statistics were calculated based off the corresponding equally likely model statistics (Mokhtarian 2016). In addition, even when not significant, variables with *t-statistic* >1 were kept in the models.

The first models were run with age and gender as explanatory variables which gave model fits in the range of 0.2 - 0.3 (with base equally likely). Age group 25-44 and gender were significant for *strong and fearless group* and for the group including *comfortable but cautious and interested but concerned*. At the second stage, income was added to age and gender. While income itself was

not significant, McFadden's ρ^2 for these models ranged between 0.3 – 0.45 although the sample size reduced to 932 from 742. Walden's t-test did not show significance of the variable income (p = 0.94, 0.32). Since the correlation between age and income was earlier found to be high (0.53), at the next step, an interaction term between age and income was introduced in the model. However, the model fit was not found to be significantly different from the previous model. In addition, introduction of interaction term led to perverse signs for the income variable. Therefore, age and income were included in the model as separate variables. Since models with age and income gave a better fit, we tested these models for multi-collinearity effect. The VIF (Variation Inflation Factor) test was performed on a linear version of the models and the VIF was found to be less than 5 for all variables including income.

Rider history was added to the model at the next step and was found to be significant across all the models. Wald's test as well shows that rider history is a significant variable (p = 3.2 e-09) for the model. At this stage, the ρ^2 values for the models range between 0.4 and 0.5 and both age groups and gender are significant across the *strong and fearless* and the *comfortable but cautious* and *interested but concerned* group. Rider history is the only significant variable for the *enthused and confident* group at this stage. Cycling frequency was added at the last step of model building and was found to be significant by Wald's test (p = 0.013). McFadden's ρ^2 values for the models with cycling frequency are ~0.7 (with base equally likely). Since the model fits were quite high, it was hypothesized that cycling frequency determines, to a large extent, the propensity of a cyclist to self-classify himself/herself into a particular category. However, at this stage model sample sizes were ~ 33% of the original sample sizes mainly because of missing data on income and cycling frequency. Since income was insignificant in all models, a final model was designed by removing income but leaving in cycling frequency which brought back the sample size ~ 50% of

the original. The ρ^2 for this model was found to be slightly lower than the earlier model but in absence of income, age group 25-44 was found to gain significance. Age, gender , rider history and cycling frequency was found to have significant influence on whether a cyclist classifies himself or herself into the categories of *strong and fearless* as well as *comfortable but cautious* and *interested but concerned*. The only significant predictor for the *enthused and confident* group was found to be cycling frequency and therefore, a model with only rider history and cycling frequency was built for this group and the ρ^2 was found to be ~ 0.6. Cycling frequency only model was found to provide a ρ^2 of 0.48 indicating that the propensity of cyclist classifying himself/herself into the *enthused and confident* category is well specified by his/her cycling frequency alone. It may therefore be suggested that cyclists who self-classify themselves into this category mostly do so because of their riding frequency rather than their self- perception on a confidence scale. As mentioned earlier, for all the categories, two final models are presented: one without cycling frequency and one with cycling frequency. Table 18 presents the model results for binary logistic models

The second variation on user's decision process was modeled using ordinal logistic models where the riders are thought of as classifying themselves into different categories *(y)* based on ordered partition of a latent continuous one dimensional confidence scale*(y*)* ($y = k$, *if* $\tau_{k-1} < y^* \leq \tau_k$ *where k = rider type categories in an ordered scale of 1 through 4, with 1 being least confident and 4 being most confident*). The model building exercise was the same as that for binary models and the results for the ordinal models are presented in Table 19.

Both the binary and ordinal logistic models are parsimonious and efficient as the choice is modeled on a single dimensional latent continuous variable. However, as mentioned by Bhat and Pulugurta (1997), it might be oversimplification of the actual decision process where the user is

actually choosing among many alternatives the one alternative that he/she feels best satisfied with. In this case, the user has a *k*-dimensional choice space where *k* represents the number of choices faced by the user and estimating an unordered response using an ordered response model can lead to biases in estimating probability of the choices (Bhat and Pulugurta 1997, Amemiya 1985). Therefore, the next set of models estimated were multinomial logistic regressions where the user was thought of as having to choose between the four rider type categories simultaneously ("Am I *strong and fearless* or *enthused and confident* or *comfortable but concerned*, etc."). The same model building exercise was followed in this case as with the binary logit models with the *comfortable but cautious* category treated as the base category. The first model included only age and gender and gave a McFadden's ρ^2 of 0.15. The final model, without income, included age group 18-24 and 25-44, gender, rider history and cycling frequency and gave a McFadden's ρ^2 of 0.6. The model with income and cycling frequency gave a model fit of 0.7 (unadjusted for model sample size). Age group 25-44 was found to be significant for the enthused and confident group when base group was changed to age group 18-24 indicating that cyclists in the age group of 25-44 behave significantly different in self- classifying themselves into enthused and confident group as compared to the age group 18-24. Chi-squared tests for model comparisons could not performed due to unequal sample sizes. Models with cycling frequency gave a higher McFadden's ρ^2 than the models without cycling frequency but were estimated on a much smaller sample size potentially removing a considerable amount of variation present in the dataset that was used for estimating the other models. Therefore, it cannot be definitively concluded that the models with cycling frequency are better models than their counterparts and hence, both types of models are presented in this paper. The multinomial logistic models are presented in Table 20. Table 21 presents the odds ratio for the multinomial and the ordinal models both with and without cycling frequency.

Table 18: Binary Logistic Regression Models

Co-efficients	Strong and Fearless Estimates (t-stat)		Enthused and Confident Estimates (t-stat)		Comfortable but Cautious & Interested, but concerned Estimates (t-stat)	
	Model 1 N= 740	*Model 2 N= 496*	*Model 1 N= 740*	*Model 2 N= 499*	*Model 1 N= 740*	*Model 2 N= 496*
Intercept	**0.239*** (5.832)**	**0.329 *** (6.846)**	**0.468 *** (8.929)**	**0.531 *** (11.406)**	**0.293 *** (6.077)**	**0.167 ** (2.832)**
Age			**Base: Age 45+**			
18-24	**0.102 . (1.763)**	0.0267 (0.429)	-0.082 (-1.112)		-0.02 (-0.286)	-0.05 (-0.654)
25-44	**0.11** (2.986)**	**0.066 . (1.665)**	-0.002 (-0.05)		**-0.108 * (-2.475)**	**-0.1 *** (-1.978)
Gender			**Base: Male**			
Female	**-0.155*** (-4.618)**	**-0.168 *** (-4.452)**	0.01 (0.243)		**0.145 *** (3.467)**	**0.158 *** (3.428)**
Income			**Base: Income < $75,000**			
Income>= $75,000	0.007 (0.248)		-0.045 (-1.142)		0.037 (1.03)	
Rider history			**Base: Since Childhood**			
One year or less	**-0.235*** (-5.779)**	**-0.169 *** (-3.644)**	**-0.14 ** (-2.685)**	-0.097 (-1.573)	**0.375 *** (7.81)**	**0.269 *** (4.733)**
Several years	**-0.143*** (-4.512)**	**-0.135 *** (-3.62)**	**0.081 * (1.982)**	0.063 (1.258)	**0.063 . (1.671)**	0.069 (1.503)
Cycling Frequency			**Base: Daily**			
Several times/week		**-0.071 . (-1.792)**		-0.08 (-1.517)		**0.144 ** (3.0)**
Once or less/week		**-0.181 *** (-3.88)**		**-0.183 ** (-2.965)**		**0.357 *** (6.252)**
Model Statistics						
Market Share of Group in the Model Dataset	*150 (20.27%)*	*90 (18.15%)*	*328 (44.32%)*	*226 (45.29%)*	*262 (35.44%)*	*180 (36.29%)*
Market Share of Other Groups in the Model Dataset	*590*	*406*	*412*	*273*	*478*	*316*
McFadden's ρ2 (Full model, base EL)	*0.368*	*0.630*	*0.460*	*0.636*	*0.473*	*0.660*
McFadden's ρ2 (MS model, base EL)	*0.177*	*0.177*	*0.292*	*0.292*	*0.251*	*0.251*
LL(0)	*-536.232*	*-536.232*	*-967.031*	*-967.031*	*-876.397*	*-876.397*
LL(MS)	*-441.357*	*-441.357*	*-684.42*	*-684.42*	*-656.533*	*-656.533*
LL(Full Model)	*-339.0786*	*-198.181*	*-522.546*	*-352.327*	*-462.058*	*-298.22*
G2=-[2(LL(Null)-LL(Full Model))]	*394.3068*	*676.102*	*888.97*	*1229.408*	*828.678*	*1156.354*

*Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*

Table 19: Ordinal Logistic Regression Models

Co-efficients	Estimates (t-stat)	
	Model 1 *N= 740*	*Model 2* *N= 496*
Intercepts	**Base: Comfortable,but cautious & Interested, but concerned**	
Comfortable, but cautious & Interested, but concerned| Enthused and confident	**-0.938** *** *(-4.53)*	**-1.755** *** *(-6.357)*
Enthused and confident|Strong and Fearless	**1.26** *** *(6.005)*	**0.692** ** *(2.62)*
Age	**Base: Age 18-24**	
Age 18-24	0.316 *(1.065)*	0.151 *(0.451)*
Age 25-44	**0.622** ** *(3.357)*	**0.448** * *(2.072)*
Gender	**Base: Male**	
Female	**-0.823** *** *(-4.847)*	**-0.939** *** *(-4.546)*
Income	**Base: Income < $75,000**	
Income >= $75,000	-0.058 *(-0.385)*	
Rider History	**Base: Since Childhood**	
One year or less	**-1.791** *** *(-8.127)*	**-1.388** *** *(-5.209)*
Several years	**-0.554** ** *(-3.532)*	**-0.596** ** *(-2.994)*
Cycling Frequncy	**Base: Daily**	
Several times per week		**-0.638** ** *(-3.045)*
Several times per month		**-1.68** *** *(-6.361)*
Model Statistics		
McFadden's ρ2 (MS model, base EL)	*0.093*	*0.093*
McFadden's ρ2 (Full model, base EL)	*0.33*	*0.58*
LL(Null Model)	*-1086.527*	*-1086.527*
LL(MS Model)	*-985.081*	*-985.081*
LL(Full Model)	-723.306	*-459.281*
G2(Full Model, base EL)	*726.442*	*1254.492*

*Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*

Table 20: Multinomial Logistic Models

Co-efficients	Enthused and confident Estimates (t-stat)		Strong and fearless Estimates (t-stat)	
	Model 1 N= 740	*Model 2 N= 496*	*Model 1 N= 740*	*Model 2 N= 496*
Intercepts	**Base: Comfortable, but cautious & Interested, but concerned**			
Enthused and Confident	**0.51*** *(2.09)*	**1.183 ** ** *(3.546)*		
Strong and Fearless			-0.274 *(-0.84)*	**0.829 *** *(1.933)*
Age	**Base: Age 45+**			
Age 18-24	-0.122 *(-0.355)*	0.202 *(0.513)*	0.363 *(1.392)*	0.297 *(0.519)*
Age 25-44	**0.348 .** *(1.6)*	**0.394 .** *(1.576)*	**0.945 ** ** *(3.149)*	**0.731*** *(2.041)*
Gender	**Base: Male**			
Female	**-0.413*** *(-2.14)*	**-0.478 *** *(-2.072)*	**-1.64 *** ** *(-4.833)*	**-2.199 *** ** *(-4.39)*
Income	**Base: Income < $75,000**			
Income >= $75,000	-0.221 *(-1.19)*		-0.1 *(-0.421)*	
Rider History	**Base: Since Childhood**			
One year or less	**-1.305 *** ** *(-5.38)*	**-1.053 ** ** *(-3.576)*	***-2.61 *** ** *(-6.09)*	**-2.07 *** ** *(-4.176)*
Several years	-0.061 *(-0.315)*	-0.156 *(-0.648)*	***-0.921*** ** *(-3.675)*	**-1.077 ** ** *(-3.135)*
Cycling Frequncy	**Base: Daily**			
Several times/week		**-0.771 ** ** *(-2.767)*		**-0.954 ** ** *(-2.789)*
Once or less/week		**-1.556 *** ** *(-4.949)*		**-2.547 *** ** *(-5.305)*
Model Statistics				
McFadden's ρ2 (MS model, base EL)	*0.09*			
McFadden's ρ2 (Full Model, base EL)	*0.34*	*0.58*		
LL(Null Model)	*-1080.56*	*-1080.56*		
LL(MS Model)	*-985.08*	*-985.08*		
LL(Full Model)	*-716.136*	*-452.958*		
G2(Full Model, base EL)	*537.89*	*1064.24*		

*Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*

Table 21: Odds Ratio for Multinomial and Ordinal Models with and without Cycling Frequency

	MNL Model 1		MNL Model 2		Ordinal Model 1	Ordinal Model 2
	Enthused and confident	Strong and fearless	Enthused and confident	Strong and fearless		
Age 18-24	0.884	1.89	1.223	1.345	1.372	1.163
Age 25-34	1.416	2.574	1.482	2.077	1.863	1.565
Gender	0.66	0.193	0.62	0.11	0.439	0.39
Income > = $75,000	0.8	0.904			0.943	
Rider history less than a year	0.27	0.0735	0.349	0.126	0.167	0.249
Rider history several years	0.94	0.398	0.859	0.341	0.575	0.551
Cycling frequency several times per week			0.462	0.385		0.529
Cycling frequency once or less per week			0.211	0.078		0.186

Overall, some distinct patterns were visible across all the models that we experimented with:

(1) Gender was significant in all the models with a negative sign implying that female cyclists are more likely to classify themselves into low comfort low confidence groups. The negative coefficients increase in value as we move from the *comfortable but cautious* and *interested but concerned* group to *strong and fearless* group which strengthens the previous inference. For the ordinal logit models, the odds ratio is ~ 0.4 which means that being female decreases the probability of being in higher confidence groups by about half. From the MNL models, being a female rider decreases the chance of being an *enthused and confident* rider as compared to *comfortable but cautious* rider by more than 30% while the chance of being a *strong and fearless* rider as compared to comfortable but cautious rider is decreased by about 80%

(2) Cyclists in the age group of 25-44 and 18-24 are more likely to be more confident riders than the cyclists in the age group of 45+. From the ordinal model without cycling frequency, cyclists in the age group of 25 to 44 are ~ 86% more likely to classify themselves into more confident categories as compared to the cyclists in the age group of 45+ while from the model with cycling frequency, riders in the age group of 25-44 are about 56% more likely to classify themselves into higher confidence groups; cyclists in the age group of 25-44 are also more likely to classify themselves into higher confidence groups than cyclists in the age group of 18-24. This may be due to the inherent construct of the dataset where most users in the age group of 18-24 are students and use bicycle because they do not have access to a car. Intuitively, they may be less bicycle enthusiasts than riders in the age group of 25-44, who, being in the higher income group (also a construct of this dataset), may have access to an automobile but still choose cycling as a mode of commute.

(3) Income is not significant but income greater than $75,000 is positively related to classifying oneself into *strong and fearless* and the *comfortable but cautious* and *interested but concerned* group and is negatively related to classifying oneself into *enthused and confident* group.

(4) Riders with more experience are likely to be more confident as is captured by the negative coefficients of rider history of several years and rider history of one year or less as compared to the riders riding from childhood. Riders in the several years category are ~50% less likely to be as confident as the riders riding from childhood while the new riders are ~80% less likely to be as confident as those riding from childhood.

(5) Cycling frequency is a significant determinant of rider type and higher frequency of cycling implies a more confident cyclist. Cyclists with cycling frequency several times per

week and cycling frequency once or less per week are both less likely to be more confident than cyclists with cycling frequency daily. However, the magnitude of the coefficient is higher in the once or less per week category than several times per week implying that cyclists in that category are even less likely than the cyclists in the several times per week category to be more confident riders. Cyclists who bicycle several times per week are about 50% less likely to rate themselves into higher confidence categories than riders who bicycle daily. Similarly, cyclists with cycling frequency once or less per week are about 80% less likely to classify themselves into higher confidence categories as compared to daily cyclists.

(6) Since the ρ^2 are similar across binary, multinomial and ordinal models, it is difficult to justify the use of any one particular type of model for the purpose of cyclist classification. However, ordinal models impose an inherent restriction on the estimation process by assuming that the effect of the explanatory variables are the same at different category levels, i.e., how gender influences in self-classifying someone as a *comfortable but cautious* rider rather than an *enthused and confident* rider is the same as the influence of gender on being *enthused and confident* rather than *strong and fearless.* This may not hold true if the perceived difference in confidence between being *strong and fearless* and *enthused and confident* is smaller than the difference between *comfortable but cautious* and *enthused and confident*. Gender may have a much more pronounced effect on choosing whether a rider is *comfortable but cautious* as compared to *enthused and confident* than in choosing between *strong and fearless* and *enthused and confident* rider type. Therefore, conceptually, MNL models seem to be more appropriate for the purpose of this research.

CHAPTER 6: CONCLUSIONS, RECOMMENDATIONS, & SUGGESTED RESEARCH

BICYCLE ROUTE CHOICE CONCLUSIONS

This report uses data collected using the Strava, CycleDixie and CycleAtlanta crowdsourced cycling smartphone applications to determine factors that influence route choice. Specifically, these factors are studied through a) modeling cycling facility prioritization preferences, b) modeling cycling route segment and path choices, and c) developing route suitability score and preference models. This comprehensive research uniquely includes work from both suburban areas, represented by Auburn, AL and urban cores, represented by Atlanta, GA. From the analyses it was found that demographics, roadway characteristics and surrounding land-use had a significant impact on whether a particular street segment would be used.

The models found that links well connected to residential areas, shopping, and mixed development are more likely to be selected as part of a route for a cyclist than other links that are not as well connected to those areas. At the same time, the model also found that links well connected to restaurants and government facilities less likely, maybe due to the increased amount of traffic that those areas attract. The model also looked into the connectivity of the links to various socio-demographic groups and found that those links well connected to areas with higher numbers of children and areas with an income of $10k to $29k, $30k to $59k and $100k and up are also less likely to be included as part of a cycling route, while links well connected to populations aged 65 and up are more likely to be selected.

The links that are well connected to areas with higher commute times, 30 minutes or greater are also favored more over those links that have shorter commutes. The most interesting finding

from this analysis was how the roadway characteristics affect the likelihood of being selected. The models also found that those links with higher peak hour volumes are more likely to be selected, along with links that have wider shoulder widths. Width of outside lane and number of driveways negatively impacted the likelihood of being selected as part of a route. Additional research could further explore the differences that commute trips and leisure cycling trips have in the decision of route choice.

The Cycle Atlanta app specifically compared stated route preference of cyclists with their actual revealed preference. The stated preference survey indicated that separate facilities are preferred by all cyclist types irrespective of how confident of a rider they are. However, actual trip analysis shows that more confident riders have shorter trip lengths and are more likely to choose shortest routes rather than detour for safer facilities. Similar trends are noted across age and gender. Therefore, to attract less confident riders and female or older riders, it is necessary to have low-stress physically separated infrastructure.

Along with the conclusions that can be made from the statistical model developed, a few conclusions can be made from the qualitative analysis. The first conclusion that can be made is that the roadway segments with the higher level of service results are being used more often over those segments that are close by that have a lower level of service. While a cyclists can not necessarily determine the LOS of a roadway from riding on it, the cyclist can determine how comfortable they feel on a particular roadway, which is what the LOS measured quantify.

Another interesting point to mention is that while a roadway may have a high LOS that does not mean that a cyclist will use it, if it is not well connected and does not allow them to get where they want to travel to. This shows that while cyclists value and safe and comfortable ride, they also place a high value on connectivity when choosing the route they are going to take.

Information presented in this report can be utilized be city planners in order to help highlight areas in which the incorporation of bicycle facilities can help support the cyclists in those areas. The model can be used to help identify roadway segments that have the highest potential for inclusion into a bicycle route. The qualitative analysis process and method can be used by city planners and engineers to identify areas that are the most connected and accessible. Applying both the model and qualitative analysis method simultaneously can give planners and engineers the information needed to identify roadway segments that are the most likely to be included in a route but also the segments that have the connectivity that is needed in order for cyclists to choose that segment over other potential segments.

It is important to recognize that while this research worked with local cycling communities to ensure that the crowdsourced data was representative, the results emphasize cyclists that are both comfortable with technology and interested in recording their travel. Therefore, the studies presented here should be repeated with different populations in different areas to confirm these conclusions.

BICYCLE LEVEL OF SERVICE AND LEVEL OF STRESS CONCLUSIONS

This study also uses roadway characteristic and cyclist survey data to compare the rankings of common bicycle level of service measures, perceived bicycle route suitability for different types of cyclists, and roadway characteristics. Specifically, four common types of bicycle LOS measures (e.g. Index, Interaction, point, and Scaled) were identified and a representative measure from each was calculated for all the major roadway segments within the city. Additionally, results from a survey of different cyclists (e.g. “Strong and Fearless”, “Enthused and Confident”, “Cautious but Comfortable”, “Interested but Concerned”, and “No Way No How”) on the perceived bicycle route suitability were collected and summarized. The comparisons highlight a surprising disconnect between level of service and suitability. Namely, suitability was perceived

the same across all different cyclist types but those segments ranked highly suitable did not correspond to those with high levels of service. Additionally, suitability was evaluated in terms of routes whereas level of service treats segments independently. The distribution of suitability and level of service were significantly different as well: suitability had a few highly suitable routes and an increasing number of less suitable locations, but the different level of service measures had varying distributions of what was acceptable or not.

Results from this study can be used by city and regional transportation planners to better inform their bicycle facility improvement decision-making in three significant ways:

1. Bicycle LOS measures are not transferrable, as each interprets the role of the similar sets of factors differently and provides different results. Planners must be thoughtful about deciding who their community perceives factors and select the appropriate measure.
2. Roadway characteristics (or combinations thereof) are not the only factors affecting cyclists' perception of the bicycle network. Access to activities and convenience influence cycling decisions and should be used in the planning process as well.
3. Selecting improvements to bicycle facilities should follow a two-step process. First, important routes should be identified (for example, by evaluating suitability). Second, LOS should be evaluated for roadway segments. Those poor segments on the highly suitable routes should be given priority as they will bolster those important connections and support the current community needs. As time progresses, additional improvements to the suitable routes can be improved.

There are many opportunities for future work within this field. First, the analysis should be repeated in different areas, including large metropolitan areas, to assess whether these trends remain the same in different areas. Second, further work must consider the relationship of level

of service and route choices, the minimum thresholds that cyclists are willing to accept and how single poor segments with a less safe design cause cyclists to reroute their trip. Third, LOS and suitability should be compared with other choices associated with cycling routes, including the propensity to bike and trip purposes.

REFERENCES

AASHTO. "Guide for the Development of Bicycle Facilities." 4th Edition. American Association of State Highway and Transportation Officials, Washington, D.C. 2012.

Akar G., K. Clifton. "The Influence of individual perceptions and bicycle infrastructure on the decision to bike." *88th Annual Meeting of the Transportation Research Board,* The Transportation Research Board of the National Academies, Washington, D.C. 2009.

Allen-Munley, C., and J. Daniel. ,"Urban Bicycle Route Safety Rating Model Application in Jersey City, New Jersey." *Journal of Transportation Engineering © ASCE*, *Vol. 132 No.6, 2006*, pp. 499-507.

Aultman-Hall L., F. Hall, B. Baetz. "Analysis of Bicycle Commuter Routes using Geographic Information Systems: Implications for Bicycle Planning." *Transportation Research Record: Journal of the Transportation Research Board No. 1578.* The Transportation Research Board of the National Academies, Washington, D.C., 1997, pp. 102-110.

Beheshtitabar E., S. Rios, D. Konig-Hollerwoger, Z. Svaty. "Route Choice Modelling for Bicycle Trips." *International Journal for Traffic and Transport Engineering Vol. 4 No. 2,* 2014, pp. 194-209.

Ben-Akiva M., M. Bierlaire. "Discrete Choice Methods and their Applications to Short Term Travel Decisions." *Handbook of Transportation Science, International Series in Operations research and Management Science Vol 23,* 1999, pp. 5-33.

Bhat, C., and M. Stinson. "A Comparison of the Route Preferences of Experienced and Inexperienced Bicycle Commuters." *Annual Meeting of the Transportation Research Board,* The Transportation Research Board of the National Academies, Washington, D.C. 2005.

Bhat, C., I. Sener, N. Eluru. "An Analysis of Bicycle Route Choice Preferences in Texas, U.S." *Transportation Vol. 36 No. 5*, September 2009, pp. 511-539.

Bushell, M. A., B W Poole, C. V. Zegeer, and D. A. Rodriguez, "Costs for Pedestrian and Bicyclist Infrastructure Improvement." University of North Carolina Highway Safety Research Center, University of North Carolina, Chapel Hill, October 2013

Charlotte Department of Transportation, "Pedestrian & Bicycle Level of Service." February 2007, http://charmeck.org/city/charlotte/Transportation/PlansProjects/Documents/AppendixBCDOTPedestrianBicycleLOSMethodologyforSignalizedIntersectionsFEB2007.pdf, Accessed July 29, 2014

Damant-Sirois, G., M. Grimsrud, A. El-Geneidy. "What's Your Type: A multidimensional cyclist Typology" *93rd Annual Meeting of the Transportation Research Board*. Transportation Research Board of the National Academies, Washington, D.C. 2014.

Dill, J., "Bicycling for Transportation and Health: The Role of Infrastructure." In *Journal of Public Health Policy, Vol. 30, 2009*, pp. 95–110.

Dill J. "Measuring Network Connectivity for Bicycling and Walking." *Transportation Research Board Annual Meeting,* Washington, D.C. 2004.

Dill, J., N. McNeil. "Four Types of Cyclists? Examination of Typology for Better Understanding of Bicycling Behavior and Potential" *Transportation Research Record: Journal of the Transportation Research Board, No. 2387,* Transportation Research Board of the National Academies, Washington, D.C., 2013, pp. 129–138.

Dixon, L. B., "Bicycle and pedestrian level-of-service performance measures and standards for congestion management systems." *Transportation Research Record: Journal of the Transportation Research Board No. 1538*, Transportation Research Board of the National Academies, Washington, D.C., 1996 pp. 1-9.

FHWA. "Subject: Bicycle and Pedestrian Facility Design Flexibility." *Bicycle and Pedestrian Facility Design Flexibility*. Federal Highway Administration, 20 Aug. 2014. Web. 03 June 2015. <http://www.fhwa.dot.gov/environment/bicycle_pedestrian/guidance/design_guidance/design_flexibility.cfm>.

FHWA. "Federal-Aid Highway Program Funding for Pedestrian and Bicycle Facilities and Programs." *Funding*. Federal Highway Administration, 5 Feb. 2015. Web. 14 July 2015. <http://www.fhwa.dot.gov/environment/bicycle_pedestrian/funding/bipedfund.cfm>.

Flusche, D. *Bicycling Means Business:* "The Economic Benefits of Bicycle Infrastructure*."* *Advocacy Advance*, July 2012

Fosgerau M., E. Frjinger, A. Karlstrom. "Route Choice Modeling without Route Choice." Association for European Transport and Contributors, 2009.

Fosgerau M., E. Frejinger, A. Karlstrom. "A Link based network route choice model with unrestricted choice set." *Munich Personal RePEc Archive, Paper No. 48707,* July 2013.

Fricker J., L. Kang. "A Bicycle Route Chocie Model that Incorporates Distance and Perceived Risk." *92nd Annual Meeting of the Transportation Research Board,* The Transportation Research Board of the National Academies, Washington, D.C., 2013.

Geller, R. "Four Types of Cyclists." Portland Office of Transportation. 2009.

Gliebe, J., J Broach, J. Dill. "Development of a Multi-Class Bicyclist Route Choice Model using Revealed Preference Data." *12th International Conference on Travel Behavior*, Jaipur, India, 2009.

Handy, S., and B. McCann, "The Regional Response to Federal Funding for Bicycle and Pedestrian Projects: An Exploratory Study." *Journal of the American Planning Association, Winter 2001, Vol. 77, No. 1,* 2010, pp. 23-38.

Harkey, D. L., D.W. Reinfut, and M. Knuiman. "Development of the Bicycle Compatibility Index." *In Transportation Research Record: Journal of the Transportation Research Board, No. 1636,* Transportation Research Board of the National Academies, Washington, D.C., 1998, pp. 13-20

Highway Capacity Manual 2010. Transportation Research Board of the National Academies, Washington, D.C., 2010.

Hochmair, H. H., "GIS-based Identification of Need for Bicycle Level of Service Improvement in Street Networks." 12th AGILE International Conference on Geographic Information Science, Hannover, Germany, 2009.

Hood, J., E Sall, B. Charlton. "A GPS-based Bicycle Route Choice Model for San Francisco, California." *Transportation Letters: The International Journal of Transportation Research,* Vol. 3. 2011. pp. 63-75

Hou N., B. Popkin, D. Jacobs Jr., Y. Song, D. Guilkey, C. Lewis, P. Gordon-Lersen. "Longitudinal Associations between neighborhood-level street network with walking, bicycling, and jogging: The CARDIA study." *Health and Place Vol. 16,* 2010 pp. 1206-1215.

Jensen, S. U., "Pedestrian and Bicyclist Level of Service on Roadway Segments." In *Transportation Research Record: Journal of the Transportation Research Board, No. 2031*, Transportation Research Board of the National Academies, Washington, D. C., 2007, pp. 43-51.

Jones, E. G., and T. D. Carlson. "Development of Bicycle Compatibility Index for Rural Roads in Nebraska." In *Transportation Research Record: Journal of the Transportation Research Board, No. 1828,* Transportation Research Board of the National Academies, Washington, D.C., 2003, pp. 124-132

Kang, K., and K. Lee. "Development of a Bicycle Level of Service Model from the User's Perspective." *In Korea Society of Civil Engineers Journal of Civil Engineering, Vol. 16, Issue 6.* Korean Society of Civil Engineers, Seoul, South Korea., 2012, pp. 1032-1039

Kentucky Transportation Cabinet, "Bicycle Level of Service", Undated, http://transportation.ky.gov/Bike-Walk/Documents/Bicycle%20Level%20of%20Service.pdf, Accessed July 29, 2014.

Krenn P., P Oja, S. Titze. "Route Choices of Transport Bicyclists: A Comparison of Actually used and shortest Routes." *International Journal of Behavioral Nutrition and Physical Activity Vol. 11 No. 31,* 2014.

Krizek, K. J., S. L. Handy, and A. Forsyth, "Explaining Changes in Walking and Bicycling Behavior: Challenges for Transportation Research." *Environment and Planning B: Planning & Design*, *Vol. 36, No. 4*, 2009, pp. 725-740.

LaMondia, J., N. Moore. "Using Bicycle Level of Service for Decision Making: A comparison of Common Bicycle Level of services measures, Roadway Characteristics, and Perceived Bike Route Suitability." *94th Annual Meeting of the Transportation Research Board,* Transportation Research Board of the National Academies, Washington, D.C. 2015.

Landis, B. W., V. R. Vattikuti, and M. Brannick, "Real-Time Human Perceptions: Toward a Bicycle Level of Service." In *Transportation Research Record: Journal of the Transportation Research Board,* No. 1578, Transportation Research Board of the National Academies, Washington, D.C., 1997, pp. 119-126.

Li, Z., W. Wang, Y. Zhang, Jie Lu, and D. R. Ragland, "Exploring Factors Influencing Bicyclists' Perception of Comfort on Bicycle Facilities." Presented at the 91st Annual Meeting of the Transportation Research Board, Washington, D.C., 2012

Lowry, M. B., D. Callister, M. Gresham, and B. Moore. "Assessment of Communitywide Bikeability with Bicycle Level of Service." *Transportation Research Record: Journal of the Transportation Research Board, No. 2314*, Transportation Research Board of the National Academies, Washington, D.C., 2003, pp. 41-48.

Ma, L., J. Dill. "Effects of the Objective and Perceived Built Environment on Bicycling for Transportation." *93rd Annual Meeting of the Transportation Research Board*, The Transportation Research Board of the National Academies, Washington, D.C., 2014.

Mach, Paul. "What do 220,000,000,000 GPS data points look like?", Strava Labs. http://engineering.strava.com/global-heatmap/

Mekuria, M., P. Furth, H. Nixon. "Low-Stress Bicycling and Network Connectivity" Mineta Transportation Institute, Report No. 11-19. May 2012.

Menghini G., N. Carrasco, N. Shussler, K. Axhausen. "Route Choice of Cyclists in Zurich: GPS-based deiscrete choice models." *Institute for Transport Planning and Systems,* Swiss Federal Institute of Technology Zurich. 2009.

Mozer, D., "Calculating Multi-Modal Levels-Of-Service (abridged)", International Bike Fund, http://www.ibike.org/engineering/los.htm#Bicycles, Accessed July 29, 2014

Nashville Area Metropolitan Planning Organization, "Regional Bicycle and Pedestrian Study: Project Evaluation Methodology", November 2009, http://nashvillempo.org/docs/bikeped/Tech_Memo_4_final_113009.pdf, Accessed July 29, 2014.

Norušis, M. J. "Chapter 4 Ordinal Regression." *IBM SPSS Statistics 19 Advanced Statistical Procedures Companion*. Upper Saddle River, NJ: Prentice Hall, 2012. 69-89. Web. http://www.norusis.com/pdf/ASPC_v13.pdf

Ophartd, E., New York State Department of Transportation, "Developing a Bicycle Level of Service Map for New York State", From PowerPoint presentation April 2005, http://www.cdtcmpo.org/bike/apr05.pdf, Accessed July 29, 2014

Parkin, J., M. Wardman, and M. Page, "Models of Perceived Cycling Risk and Route Acceptability." *In Accident Analysis and Prevention,* Vol. 39, No 2*,* 2007, pp. 364-371.

PBIC and FHWA. "The National Bicycling and Walking Study: 15-Year Status Report." University of North Carolina Highway Safety Research Center, University of North Carolina, Chapel Hill, May 2010

Petritsch, T. A., B. W. Landis, and T. Scorsone, "Addressing Deficiencies HCM Bike Level of Service Model for Arterial Roadways." Presented at 93rd Annual Meeting of the Transportation Research Board, Washington, D.C., 2014

Poulos, R., J. Hatfield, C. Rissel, L. Flack, S. Murphy, R. Grzebieta. "Characteristics, cycling patterns, and Crash and Injury Experiences at Baseline of a Cohort of Transportation and Recreational Cyclists in New South Wales, Australia." *Accident Analysis and Prevention Vol. 78,* 2015, pp. 155-164.

Pucher, J.., J. Dill, S. Handy. "Infrastructure, Programs, and Policies to Increase Bicycling: an International Review." *Preventive Medicine*, *Vol. 50, 2010*, pp. S106–S125.

Qing Shen P., P. Chen, P Schmiedeskamp, A. Bassok, S. Childress. "Bicycle Route Choice: GPS Data Collection and Travel Model- Development Year 1 (2012-2013)- Final Project Report." *Pacific Northwest Transportation Consortium (PacTrans)*, USDOT University Transportation Center for Federal Region 10, University of Washington, Seattle, WA. 2014.

Robinson T., E. Kidando, D. Chimba. "Utilizing Bicycle Level of Service and Latent Demand Score for Small City Bicycle Route Planning." *2015 Transportation Research Board Annual Meeting,* The Transportation Research Board of the National Academies, Washington, D.C. 2014.

Segadilha A., S. Sanches. "Identification of factors that Influence cyclists' route choice." *XI Congreso de Ingenieria del Transporte (CIT 2014), Procedia- Social and Behavioral Sciences Vol. 160,* 2014, pp. 372-380.

Segadilha A., S. Sanches. "Analysis of bicycle Commuter Routes using GPSs and GIS." *XVIII Congreso Panamericano de Ingenieria de Transito, Transporte, y Logistica (PANAM 2014),* Procedia- Social and Behavioral Sciences Vol. 162, 2014, pp. 198-207.

Shinkle, Douglas, and Anne Teigen. "Encouraging Bicycling and Walking: The State Legislative Role." *NCI Standardized Surveys of Walking & Bicycling Database* (n.d.): n. pag. National Conference of State Legislatures, Nov. 2008. Web. 14 July 2015. <http://www.ncsl.org/documents/transportation/encouragingbicyclingwalking.pdf>.

Sousa, A., S. Sanches, M. Ferreira. "Perception of barriers for the use of Bicycles." *XI Congreso de Ingenieria del Transporte (CIT 2014), Procedia- Social and Behavioral Sciences Vol. 160,* 2014, pp. 304-313.

Stone, John. "Started Using Strava (It's Good to Be the King!)." *John Stone Fitness RSS*. N.p., 02 May 2012. Web. 08 June 2015. <http://www.johnstonefitness.com/2012/05/02/started-using-strava-its-good-to-be-the-king/>.

"Strava." *Running and Cycling GPS Tracker, Performance Analytics, Maps, Clubs and Competition*. Strava Inc., 2015. Web. 14 July 2015. <https://www.strava.com/>.

"United States Census Bureau." *Households*. American Community Survey, n.d. Web. 14 Aug. 2015. <http://quickfacts.census.gov/qfd/meta/long_HSD310213.htm>.

Urban, M., C. Porter, K. Proussaloglou, R. Calix, C. Chu. "Modeling the Impacts of Bicycle Facilities on Commute and Recreational Bicycling in Los Angeles County." *93rd Annual Meeting of the Transportation Research Board,* The Transportation Research Board of the National Academies, Washington, D.C. 2014

U.S. Census Bureau, Population Division.*2013 Population Estimate*. June 3, 2014. Retrieved June 3, 2014

USDOT, "Livability 101: Six Principles of Livability." September 9, 2013, http://www.dot.gov/livability/101, Accessed July 29, 2014.

"Vermont Pedestrian and Bicycle Facility Planning and Design Manual" Vermont Agency of Transportation. December 2002.

Wang J., A. Cheung, L. Mirza, and S. Moradi. "Understanding factors influencing Choices of Cyclists and Potential Cyclists: A case study at the University of Auckland." *Road and Transport Research: A Journal of Australian and New Zealand Research and Practice, Vol. 23 No. 4,* 2014, pp. 37-51.

Yamanaka, H., and S. Namerikawa, Measuring Level-Of-Service for Cycling of Urban Streets Using "Probe Bicycle System." *In Journal of the Eastern Asia Society for Transportation Studies. Vol. 7,* Eastern Asia Society for Transportation Studies, Tokyo, Japan, 2007, pp. 1614-1625

Yang, C., M. Mesbah. "Route Choice Behaviour of Cyclists by Stated Preference and Revealed Preference." *Australasian Transport Research Forum 2013 Proceedings*. 2013.

Zolnik, E., E. Cromley. "Poisson Multilevel Methodology of Bicycle Levels of Service for Road Networks." *Transportation Research Record: Journal of the Transportation Research Board No. 2031,* The Transportation Research Board of the National Academies, Washington, D.C., 2007, pp. 1-8.