Modeling Factors Influencing Commuter Cycling Routes:
A Study of GPS Cycling Records in Auburn, Alabama

by

Michael Alexander Moore

A Thesis submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Master of Science

Auburn, Alabama
December 12, 2015

Key Words: Bicycle, STRAVA, Regression, Ordinal-Logistic, ArcGIS

Approved by

Jeffrey J. LaMondia, Chair, Assistant Professor of Civil Engineering
Rod Turochy, Associate Professor of Civil Engineering
Sweta Byahut, Assistant Professor of Community Planning
Chandana Mitra, Assistant Professor of Geosciences

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 thesis uses data collected using the Strava cycling smartphone application to determine factors that influence route choice. An ordinal logistic regression model was developed in order to determine the influencing factors and the level of influence that they had on a cyclist's decision of what route to take. Along with the model developed, GIS was utilized in order to perform a qualitative analysis, looking at specific areas and facilities to see what caused them to differ from surrounding facilities. From the analyses it was found that roadway characteristics and surrounding land-use had a significant impact on whether a particular street segment would be used.

ACKNOWLEDGMENTS

I would first like to thank my professor Dr. Jeffrey LaMondia. It is with his many hours of guidance and his constant enthusiasm that has allowed this research to be possible. I would also like to give thanks to Dr. Rod Turochy, Dr. Sweta Byahut, and Dr. Chandana Mitra for serving on my thesis committee. Last, but certainly not least, I would like to thank my parents. It is with their encouragement and continuing support that helped me make this possible. Words cannot express how much I thank you for all you have done.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AASHTO   American Association of State Highway and Transportation Officials

ACS   American Community Survey

BLOS   Bicycle Level of Service

BCI   Bicycle Compatibility Index

DOT   Department of Transportation

FHWA   Federal Highway Administration

GIS   Geographic Information Systems

GPS   Global Positioning System

LOS   Level of Service

MATSim   Multi-Agent Transport Simulation Toolkit

PLUM   Polytomous Universal Model

NHTS   National Household Travel Survey

CHAPTER 1

INTRODUCTION

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 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.

The rest of this thesis is structured as follows: Chapter 2 discusses the previous work done, including the various statistical models used, and defines and discusses factors which influence cycling route choice. Chapter 3 discusses the data used in this research, including the source of the data and what information was contained in the data. The creation of the final dataset used in the route choice model built in this thesis is also discussed in Chapter 3. Chapter 4 presents the methods used for analysis of the cycling routes, including the defining and discussion of the ordinal logistic regression, and the use of Geographic Information Systems (GIS) that were used in the analysis of cycling route choice in the City of Auburn. The results from the ordinal logistic regression model and the GIS analysis are presented and discussed in Chapter 5. A summary and conclusion of the work conducted is presented in Chapter 6.

CHAPTER 2

LITERATURE REVIEW

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 thesis.

2.1 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.

2.2 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).

2.3 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 in order to obtain information about the cyclists and information about the roadway.

2.4 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 in order 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. In order 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 in order 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 in order to find the suitable alternative routes for the use in the choice set that they generated. In order 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

12

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.

2.5 Summary vis-à-vis This Research

From a review of the past work conducted on route choice, it was seen that many of the models focused on route choice by looking into the route as a whole. The main limitation of modeling route choice in this manner is the need to develop alternative routes by generating a route choice set. To overcome this limitation, the research presented in this thesis models route choice as a choice of links, focusing on modeling the route on the link level. The connectivity of these links to various land-use areas and socio-demographic groups throughout the whole city was also used in the modeling, whereas many of the past research focused on which land-use was adjacent to the route being used and not how well the route was connected to land-use throughout the city.

CHAPTER 3

DATA

Data from a variety of sources was collected and compiled to analyze the factors affecting where cyclists are likely to choose routes. These include STRAVA cyclist route records, roadway characteristics, land use patterns, and regional socio-demographics.

3.1 STRAVA Cyclist Route Data

The routing data used in this project was obtained 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 3.1.1, 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.

The accuracy of the GPS data 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. While Strava is open to anyone for use, due to it having to be

downloaded onto a device, it appeared that serious cyclists would be more apt to download the

smartphone application and use the route tracker, than the average recreational cyclist, although

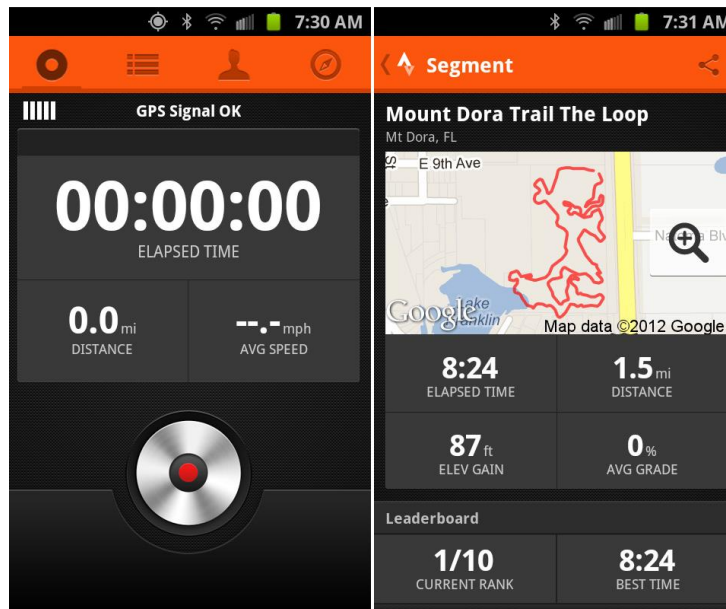no user demographic data was given by Strava.

Figure 3.1.1: Strava App Screenshot   (Source: John Stone 2012)

The research conducted in this thesis is the first to utilize the route data collected by

Strava. While the data that was utilized provided counts on the number of Strava users selecting

a particular roadway segment, and the number of cycling trips taken on a segment, it was not

possible to determine a specific route from the data that was provided. Also due to protecting the

privacy of its users, it was not possible to know background demographics on those using the

Strava app.

3.2 Roadway Characteristics

In order 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.

3.3 Land Use Characteristics & 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.

In order 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.

$$A_i = \sum_{n=1}^{z} x_z d_{iz}^{-1.5} \qquad\qquad (2)$$

In order 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.

3.4 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 in order 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 in order 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. In order to be able 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.

3.5 Final Dataset

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 in order to be able to use the data. Screening was performed in order to verify that there were no abnormalities in the data provided, for example checking the roadway segments in order 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 in order 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 3.5.1 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 3.5.1 showing the Strava usage groupings, Figure 3.5.1 shows on which segments these groups chose to travel.

Table 3.5.1: Strava Usage Groups

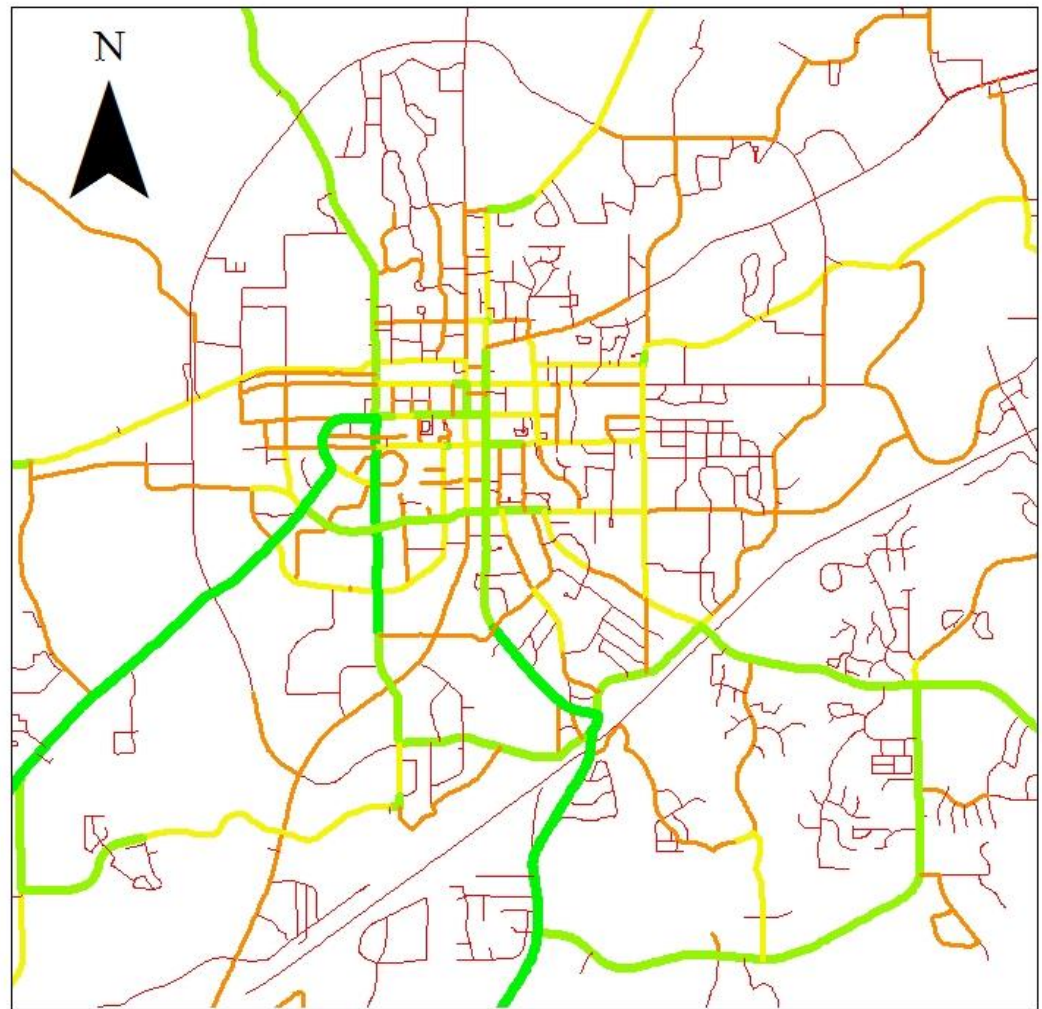| Usage Group	Number of Cyclists
Low	0-13
Low-Average	14-34
Average	35-58
High-Average	59-93
High	94-157

Figure 3.5.1: 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 3.5.2 below shows the number of trips and the frequency of cyclists who cycled that many trips.
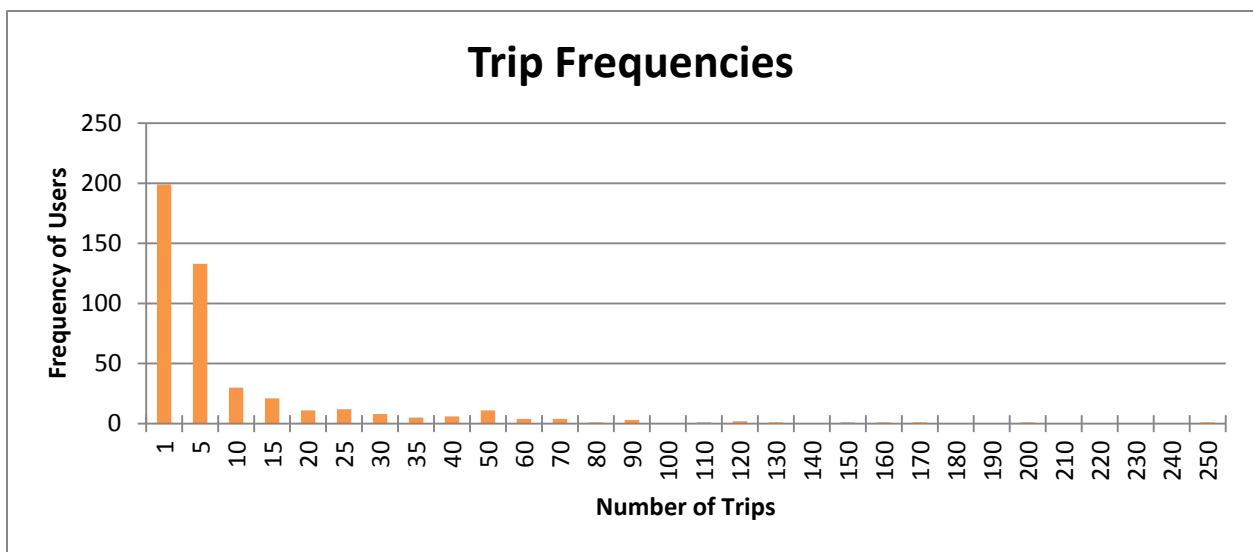

Figure 3.5.2: Trip Frequencies

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 3.5.3. This percentage breakdown suggests that cyclists are more concerned about tracking their recreational trips and not their commute trips.
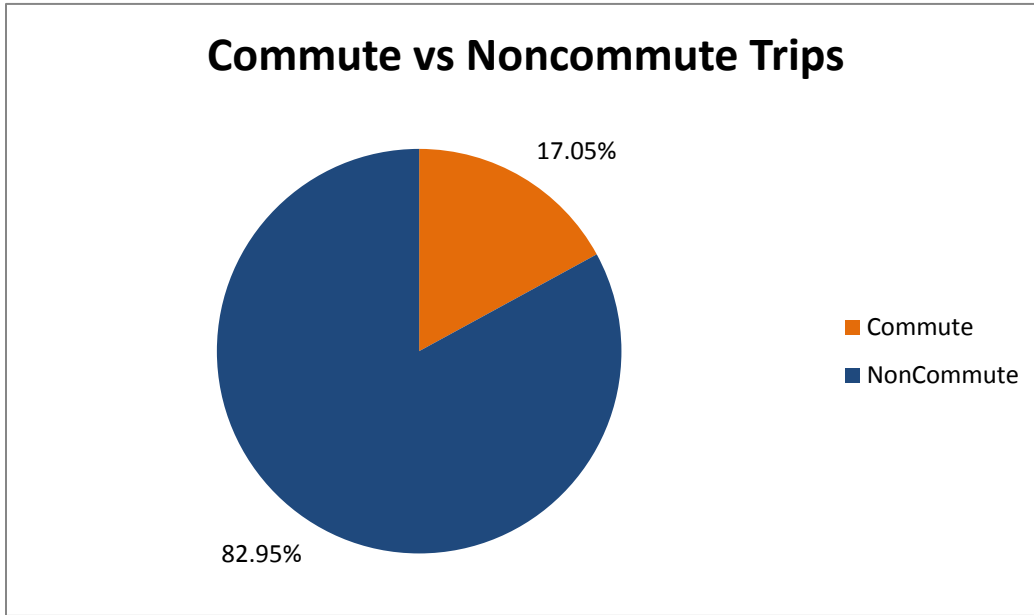
Figure 3.5.3: Percentage of Commute vs Non-Commute Trips

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.

CHAPTER 4

METHODOLOGY

4.1 Ordinal Logistic Regression

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.
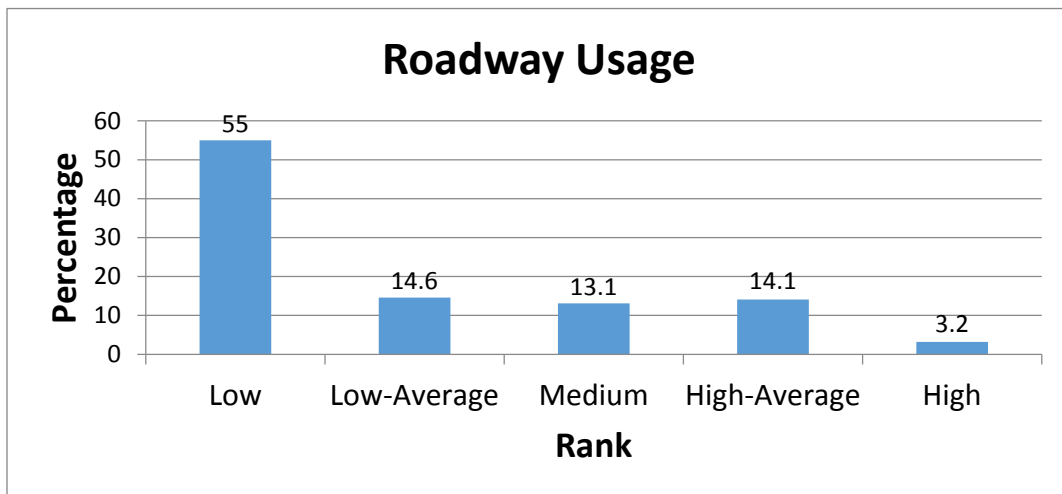

Figure 4.1.1: Percentage of Strava Users per Usage Rank

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 4.1.1.

Table 4.1.1: Independent Variables Considered for Model

Roadway Characteristics	Socio-Demographic Accessibility	
Peak Hour Volume	Number of People Aged…	Houhsehold 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 memebers
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, $\chi^2 = 41.401$, 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.

$$t = \frac{\bar{x} - \mu}{SE} \qquad\qquad (2)$$

4.2 Geographic Information Systems (GIS) Qualitative Analysis

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 4.2.1 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 4.2.1- 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.
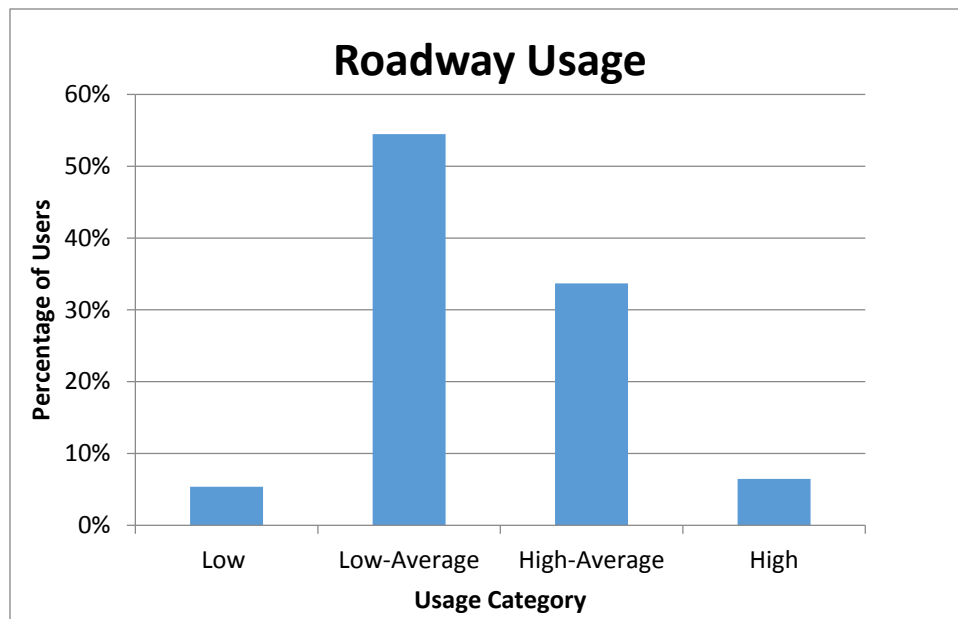

Figure 4.2.1: Qualitative Roadway Usage

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.

CHAPTER 5

RESULTS

5.1 Ordinal Regression Model 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 5.1. 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 5.1: 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

5.1.1 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. Looking at Figure A.1 in the Appendix, it can be seen that the majority of roadway segments in the city of auburn do not include a paved shoulder. However, when comparing where cyclists are traveling in Figure 3.5.1 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 5.1 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 A.2, 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. The avoidance of roadway segments with higher number of driveways can be seen in Figure A.3, with the 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.

The roadway segment outside lane widths can be seen in Figure A.4. 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.

5.1.2 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. Looking at Figure A.5 in the Appendix, 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, Figures A.6 and A.7, 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, Figure A.8. 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, Figure A.9 in the appendix. 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.

5.1.3 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.

By looking Figure A.10, 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. From Figure A.11, 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. Looking at Figure A.12 in the appendix, it can be seen that the 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.

5.2 City of Auburn, AL Bikeability and Qualitative Analysis

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 5.2.1, 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 5.2.1. 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 5.2.1. 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 5.2.1 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

Legend

BikePaths

Facility Type

————	Bike Lane
————	Concrete Multi-Use Path
————	Multi-Use Lane
————	Off-road Bikepath (Dirt)
————	Off-road Bikepath (Paved)
————	Proposed
————	auburnstreets

Miles

0   0.325 0.65   1.3   1.95   2.6
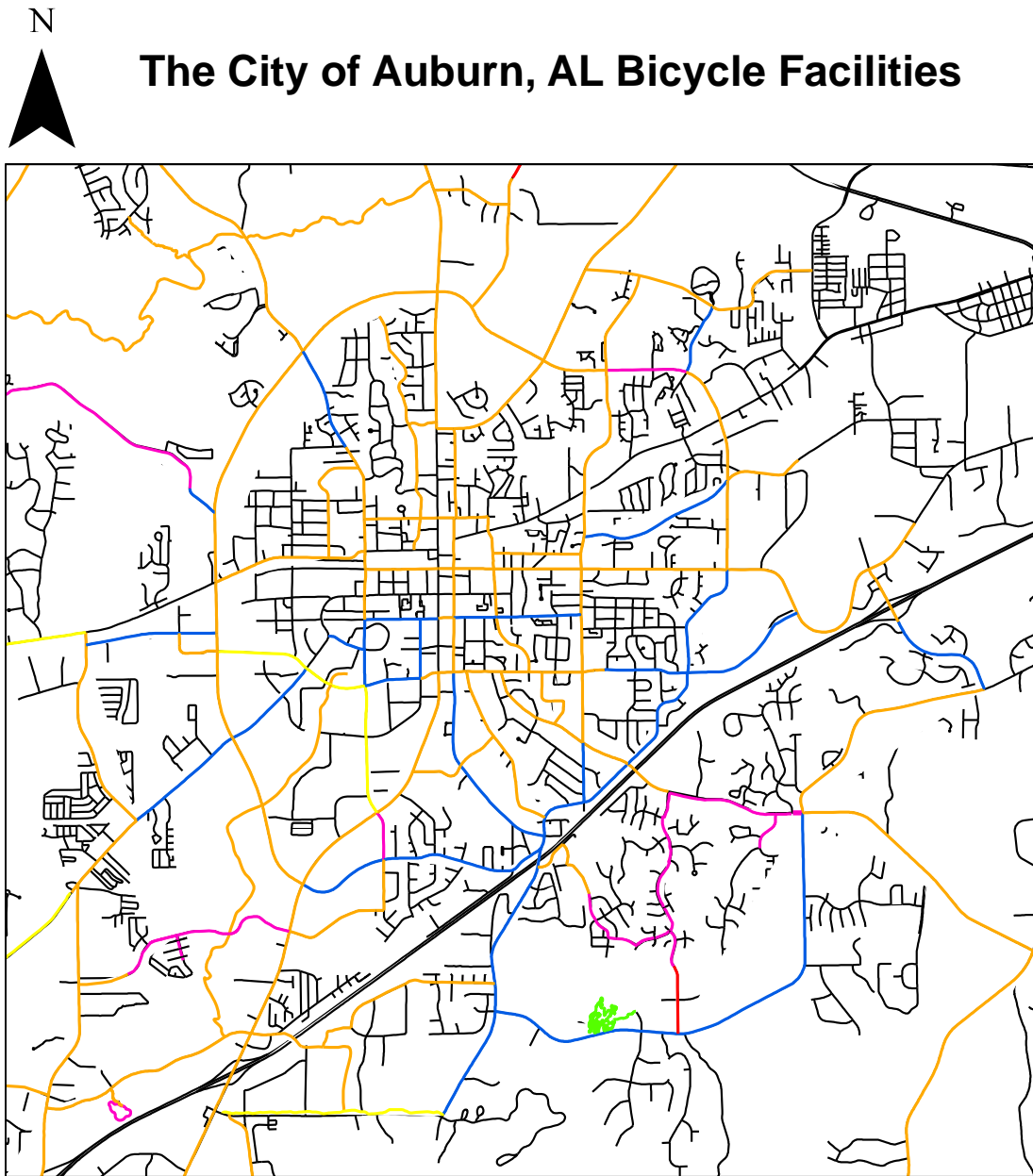
Figure 5.2.1: City of Auburn, AL Bicycle Facilities

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 5.2.2 below.

Figure 5.2.2: 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 5.2.2, 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 5.2.2: 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.
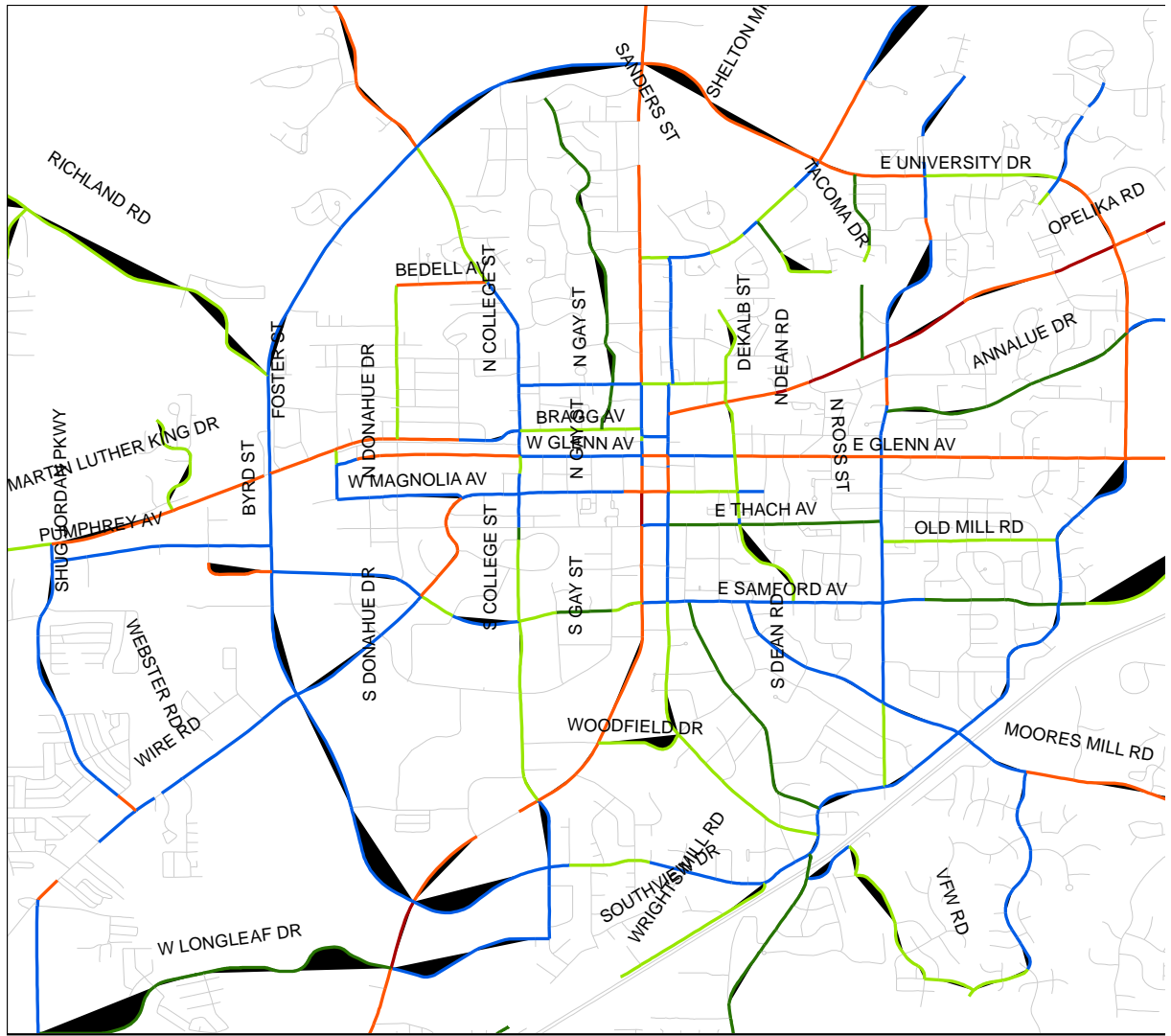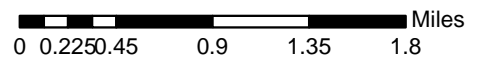
Figure 5.2.3: Auburn Streets BLOS: Bicycle Compatibility Index

Seen in Table 5.2.2 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 5.2.3: 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 5.2.2 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 5.2.2, 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.
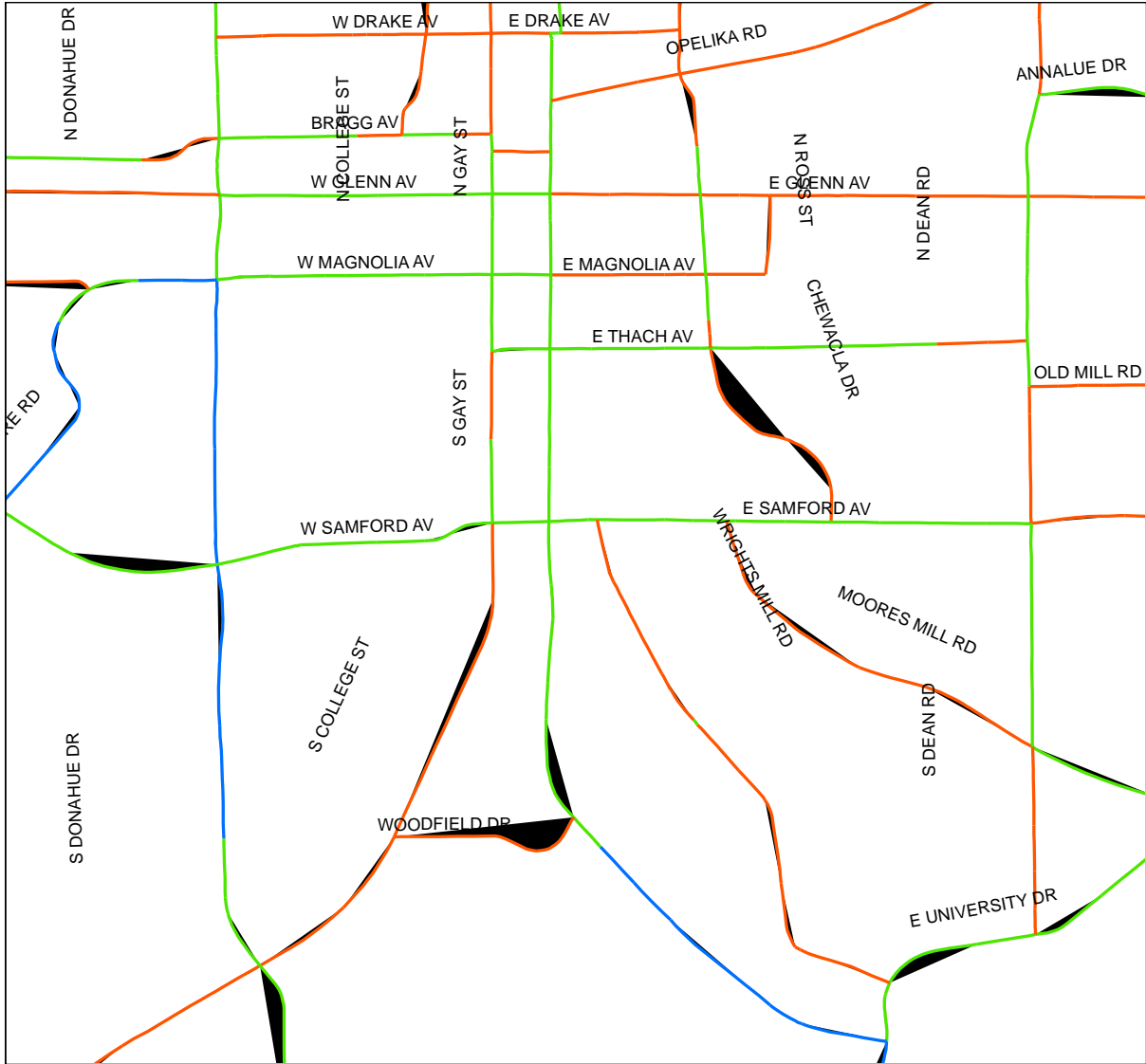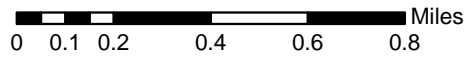
Figure 5.2.4: Auburn Street Usage- Parallel Facilities

CHAPTER 6

CONCLUSIONS

This analysis used unique data collected from the GPS cycling route tracking Strava which was restructured in order to model the likelihood of a link being chosen using an ordinal logistic regression model. The sample was comprised of individuals from the Auburn area and, while it is not certain what user level they were, it can be assumed that the sampled individuals were more on the experienced cyclists side of the bicycle user scale than the inexperienced. Despite the less than ideal mix of experienced and inexperienced users, a number of variables were found to be significant including factors that describe the roadway characteristics, connectivity to various groups, and connectivity to various socio-demographic groups.

The model 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 model 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.

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.

The research presented in this thesis help further the research into bicycle route choice by providing a method of evaluation on the likelihood of a segment to be selected as part of a cyclist's route. The model created in during this project is also important in that it does not rely on the creation of an alternative route choice set in order to evaluate the likelihood of a roadway segment to be selected. This research is also significant in the fact that it is the first to utilize data collected from the bicycle route tracker Strava.

Information presented in this thesis 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.

To further understand the roadway segments that cyclists in the City of Auburn choose to incorporate into their routes, future work could be conducted into how students and the regular residential population of auburn differ in the roadway segments that they prefer. To accomplish this further work, GIS could be used to look into areas with large rental populations compared to those areas with higher amounts of home ownership. Looking into differing land-use types, such as multi-family residential zoning, can give further insights into how different population groups cycle.

CHAPTER 7

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.

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.

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. "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.

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>.

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.

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.

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.

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.

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.

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.

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

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.

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

"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.

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.

CHAPTER 8
APPENDIX

8.1 Roadway Characteristics
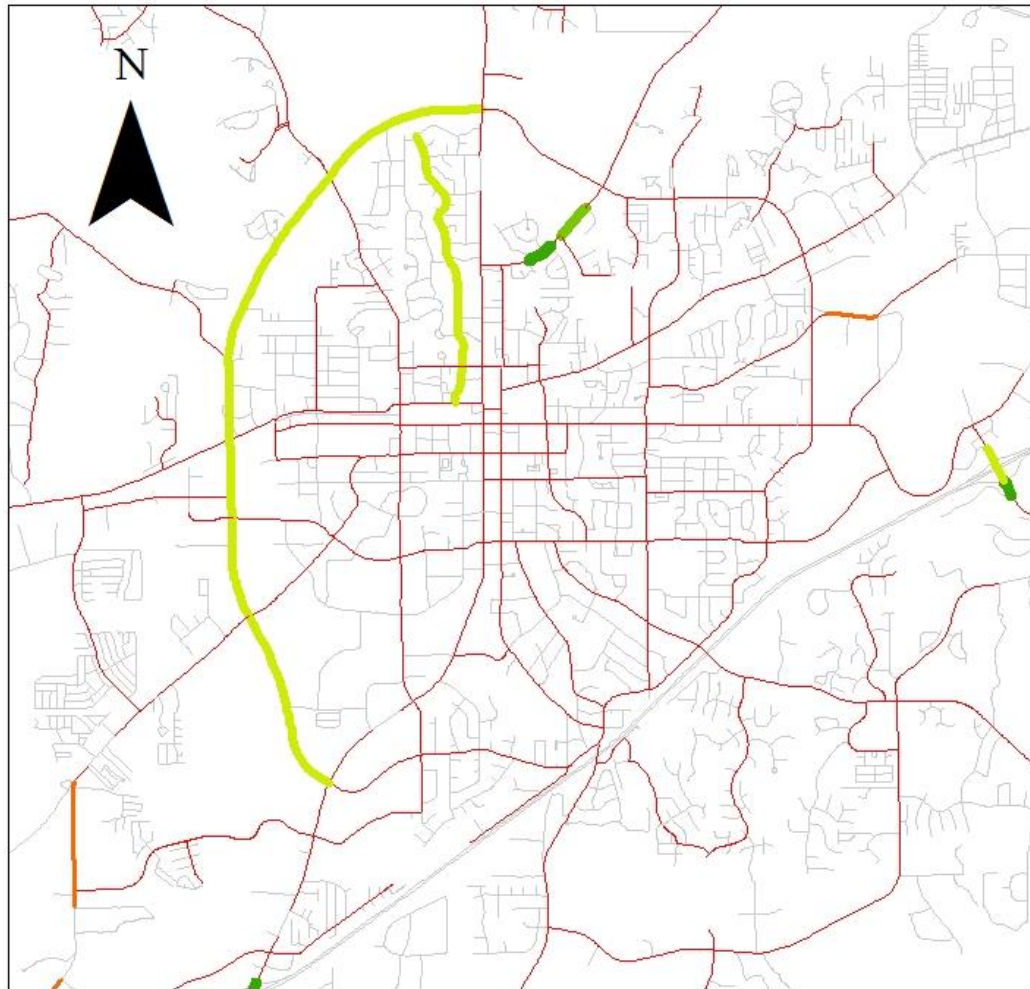
Legend

Auburn Streets

Width of paved shoulder

— 0.00
— 0.01 - 1.00
— 1.01 - 3.00
— 3.01 - 6.00
— 6.01 - 9.00
— 9.01 - 11.00
— Auburn Streets
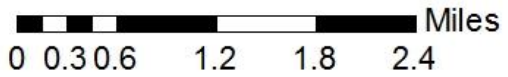
Figure A.1: Width of Paved Shoulder per Roadway Segment
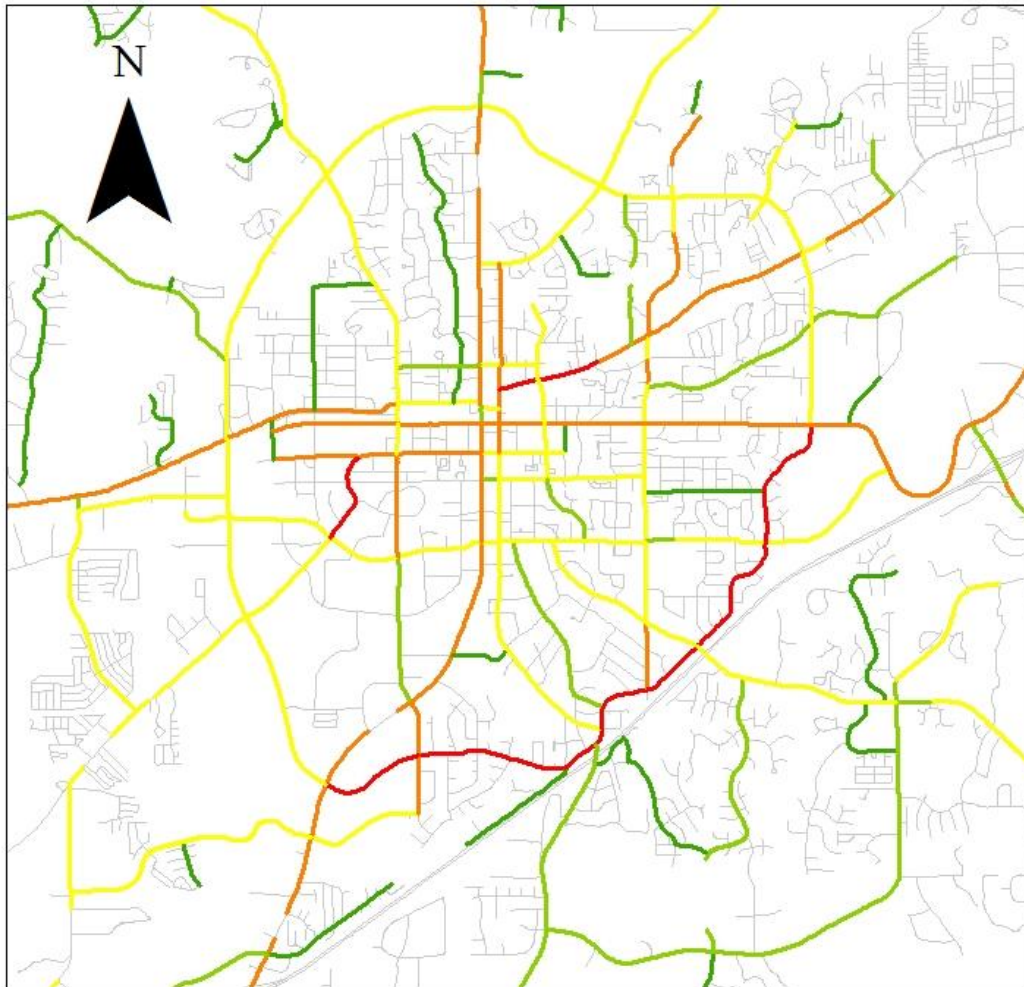
Legend

Auburn Streets

peak hour
- 0.00 - 85.00
- 85.01 - 208.00
- 208.01 - 352.00
- 352.01 - 500.00
- 500.01 - 915.00
- Auburn Streets
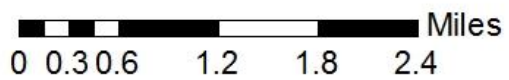
Figure A.2: Peak Hour Volume (VPH)
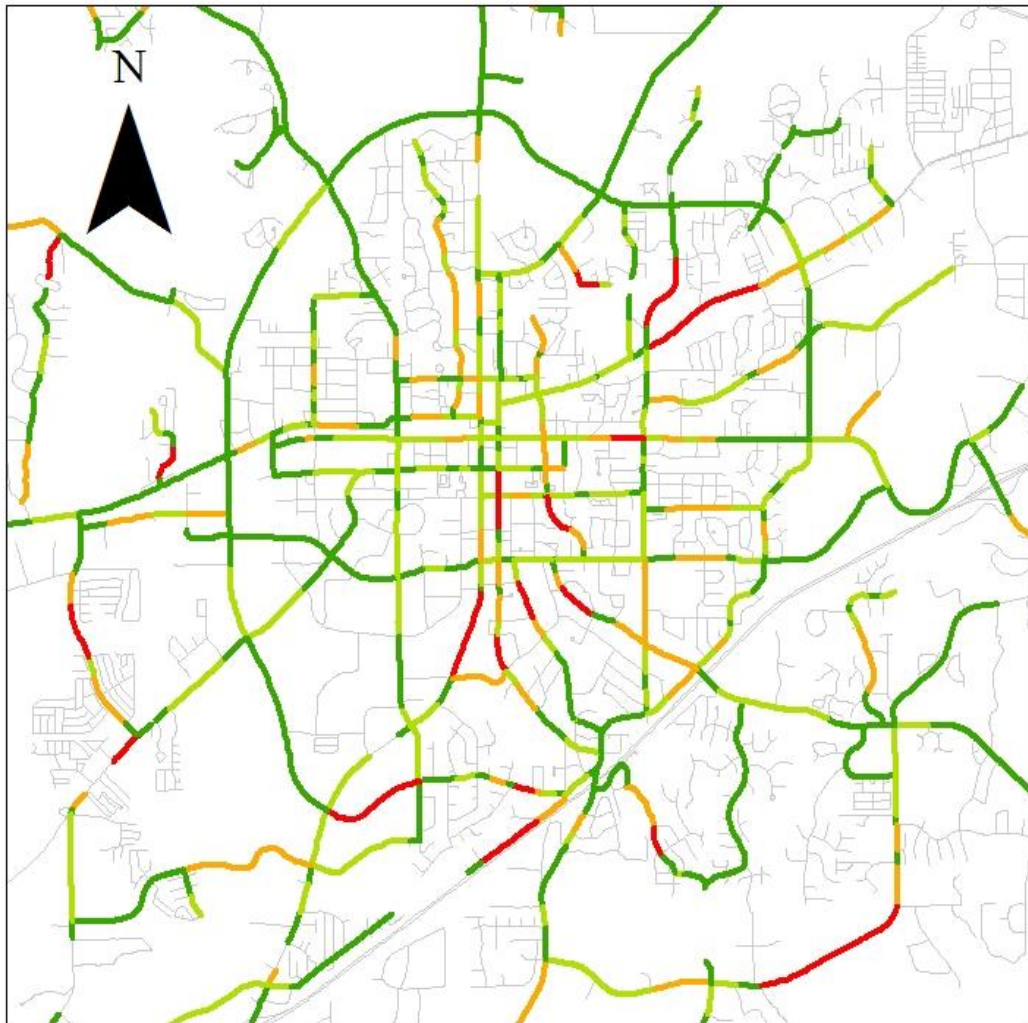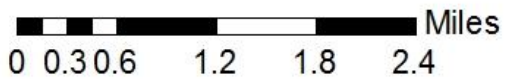
Figure A.3: Number of Driveways per Roadway Segment

Legend

Auburn Streets

Width of outside lane

— 5.00 - 10.00

— 10.01 - 12.00

— 12.01 - 15.00

— 15.01 - 20.00
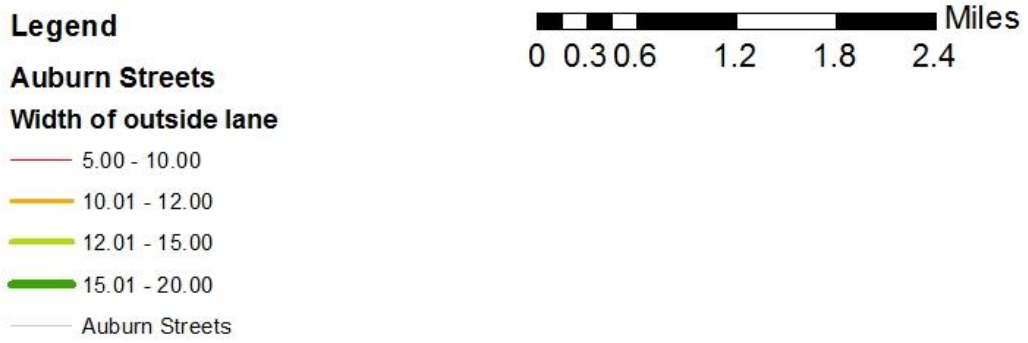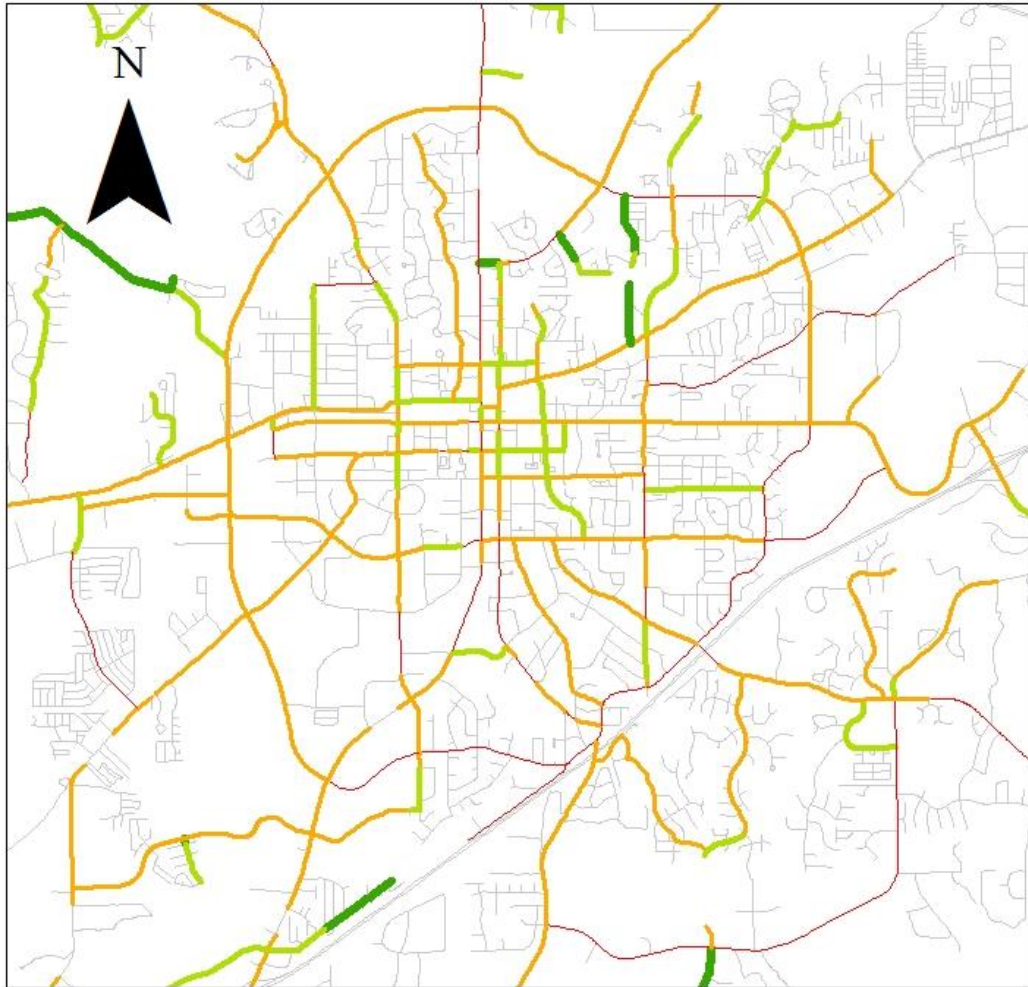
— Auburn Streets

Figure A.4: Width of Outside Lane per Roadway Segment

8.2 Land-Use Access

Legend

— Auburn Streets
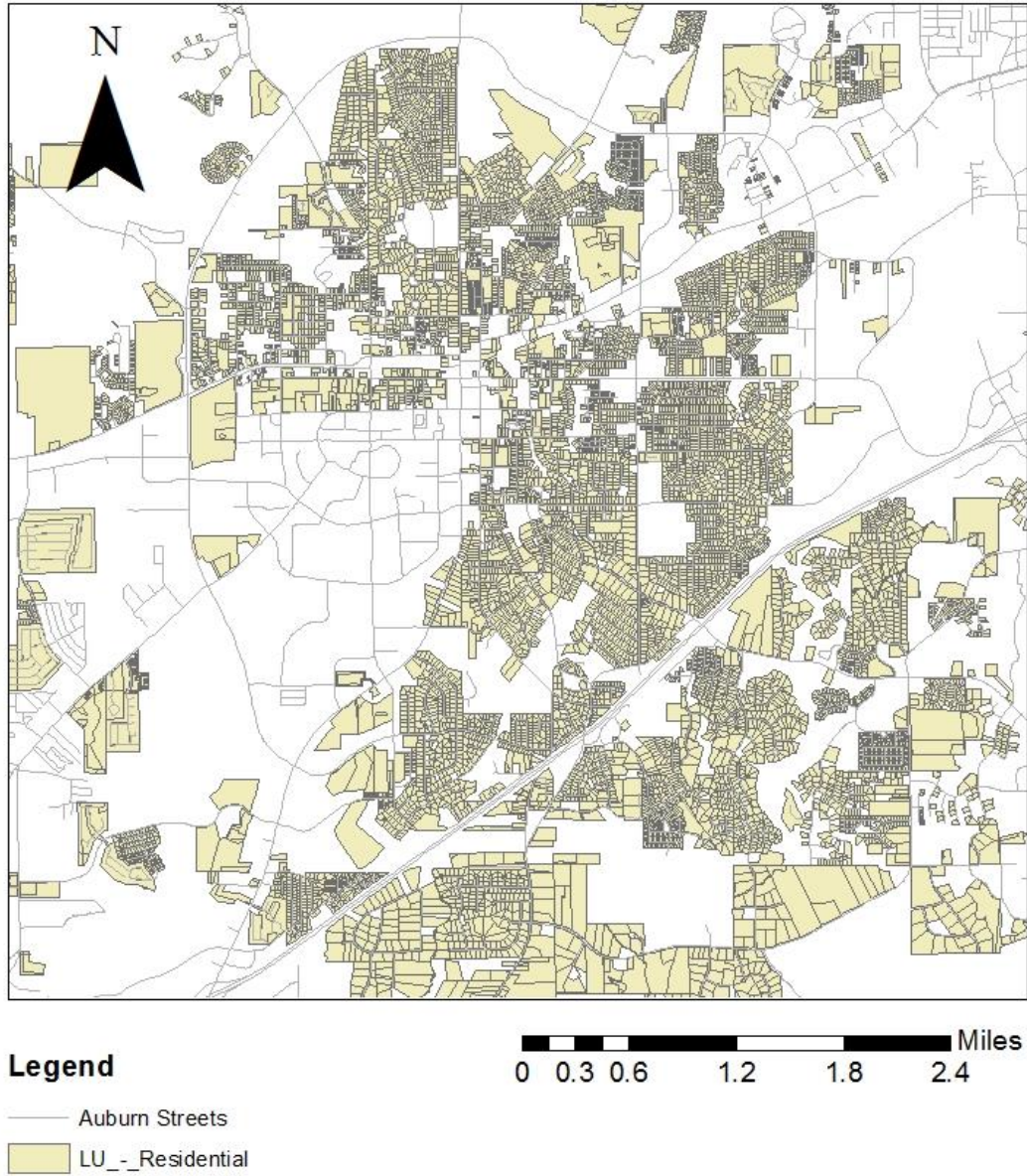
▨ LU_-_Residential

Figure A.5: Residential Land-Use
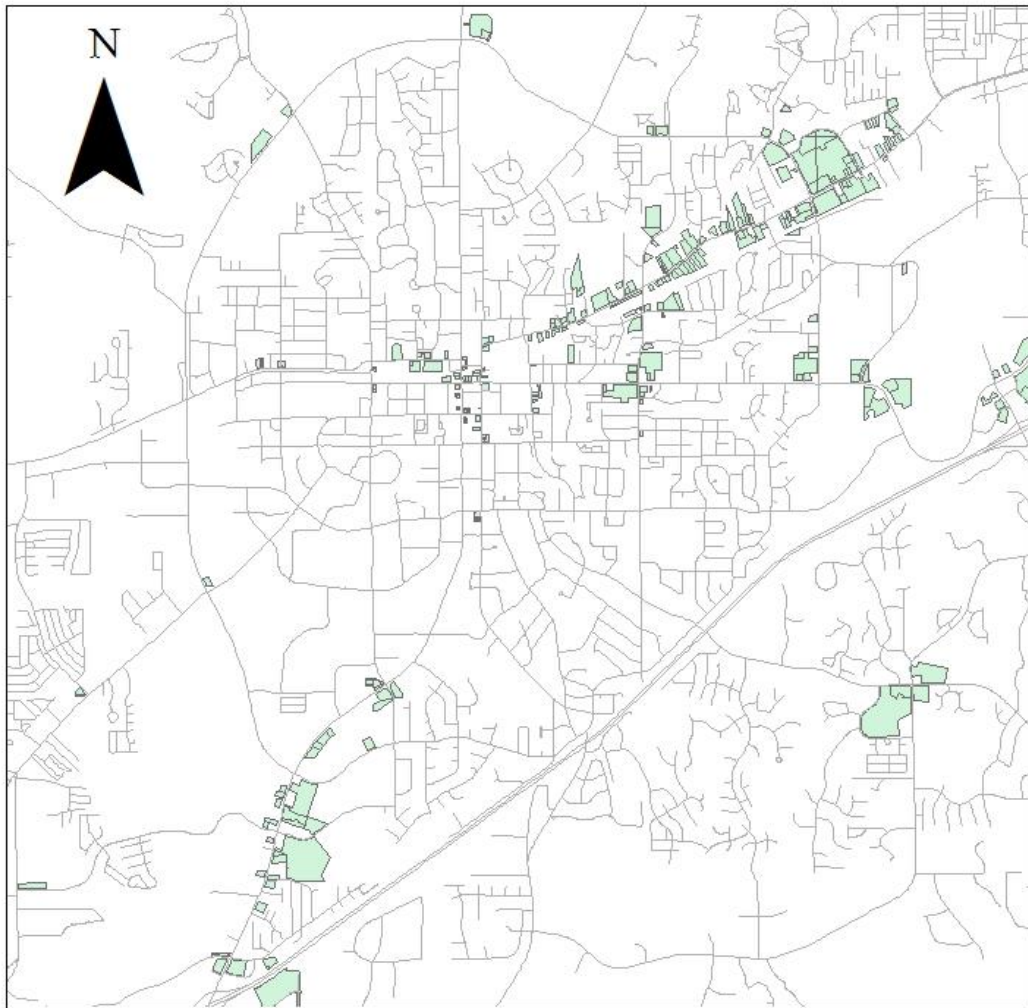
Legend

—— Auburn Streets

▢ LU_-_Commercial

Figure A.6: Commercial Land-Use
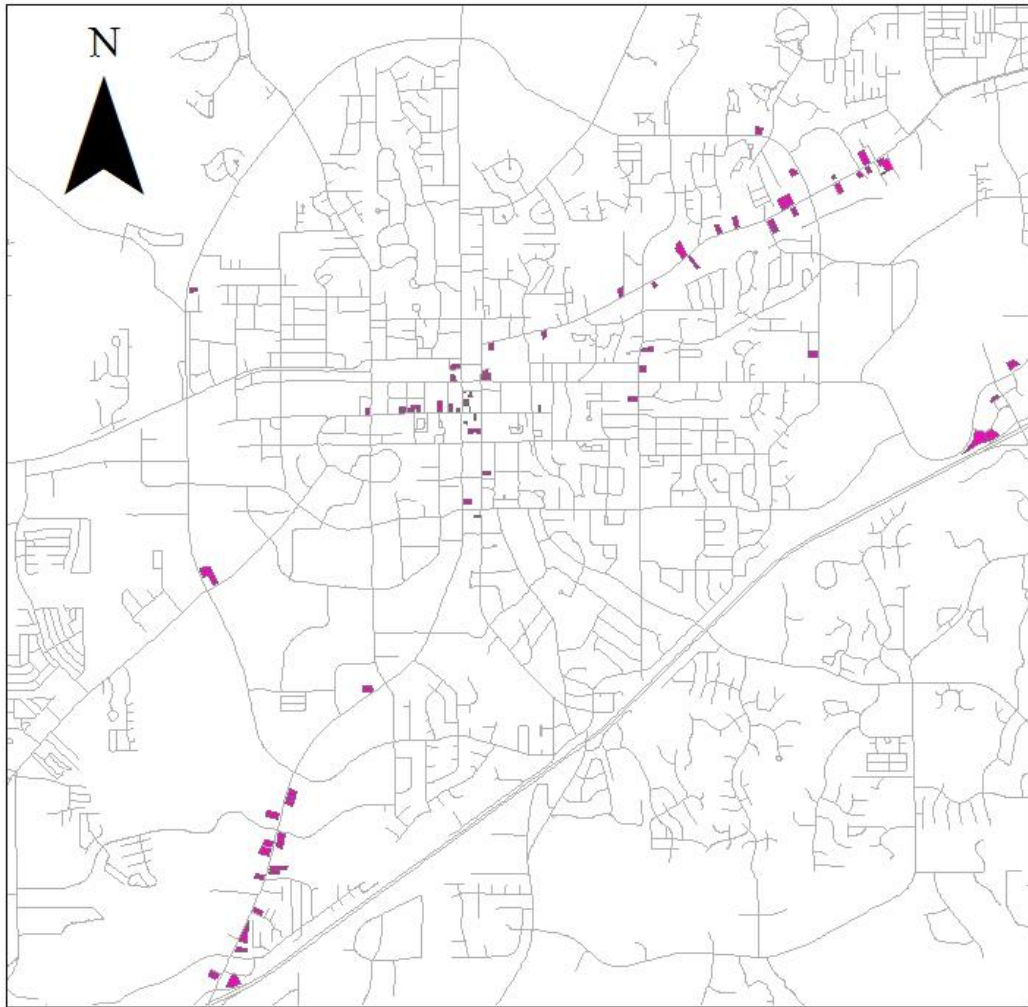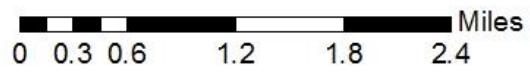
Figure A.7: Restaurant Developments

Legend

—— Auburn Streets

▮ LU_-_Mixed
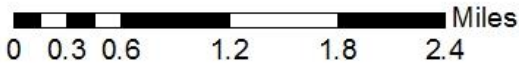
Figure A.8: Mixed-Use Developments
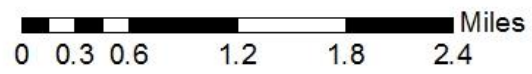
Figure A.9: Governmental Developments

8.3: Socio-Demographic Access

Legend

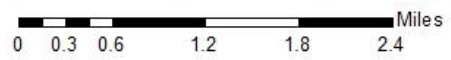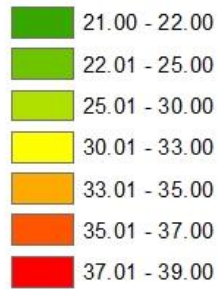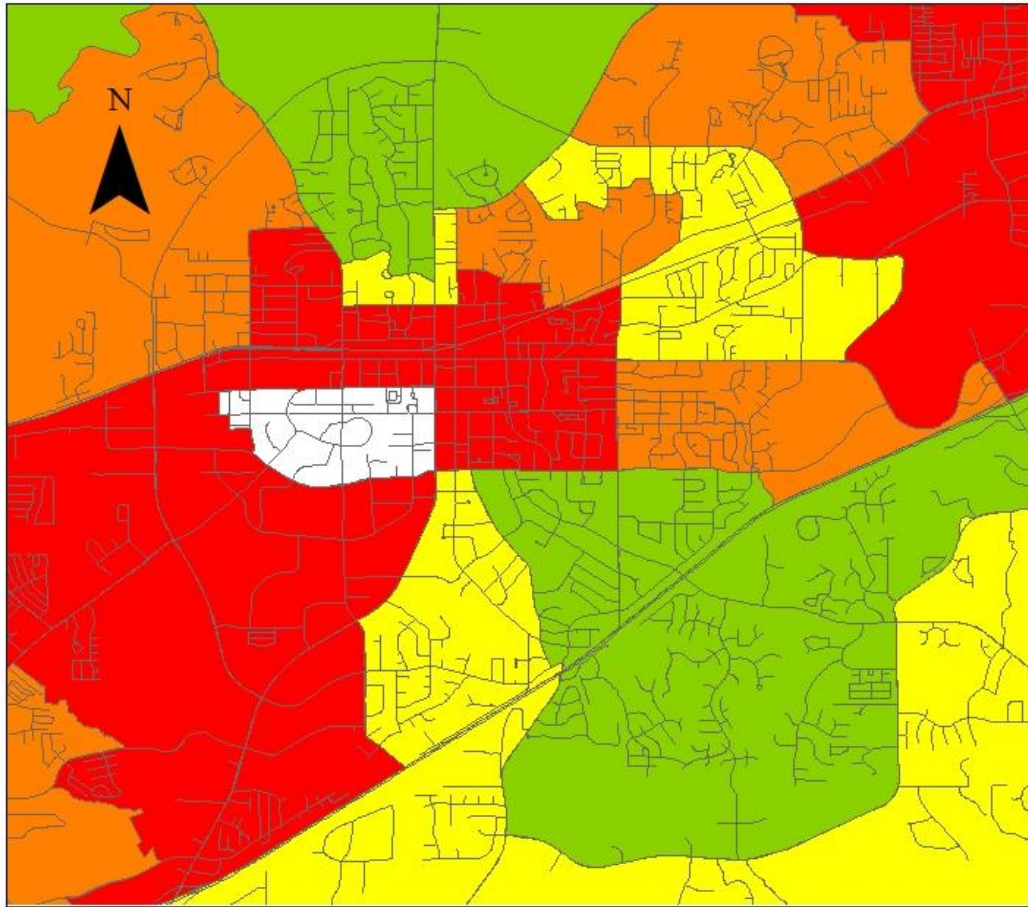——— Auburn Streets

Auburn Census Block Groups

Avg_Age

■	21.00 - 22.00
■	22.01 - 25.00
■	25.01 - 30.00
■	30.01 - 33.00
■	33.01 - 35.00
■	35.01 - 37.00
■	37.01 - 39.00

Miles
0   0.3 0.6   1.2   1.8   2.4
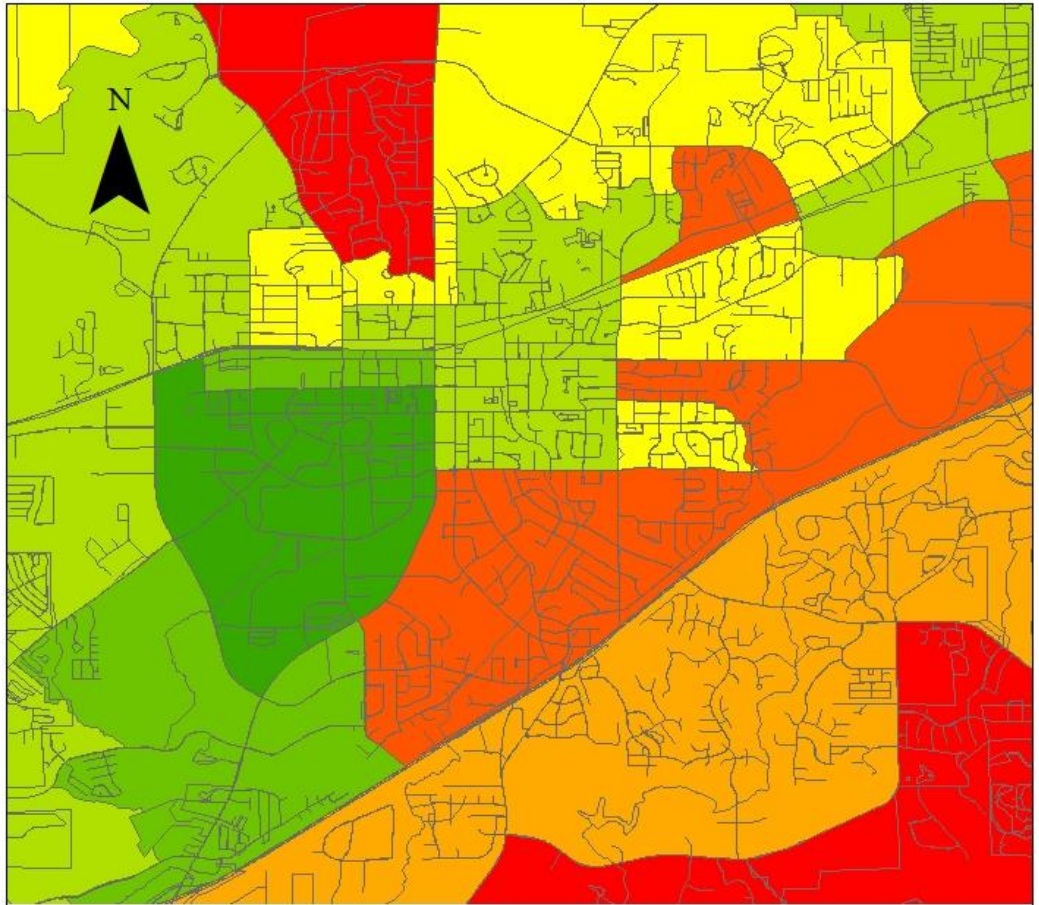
Figure A.10: Census Block Groups Average Age

Legend

⎯ Auburn Streets

Auburn Census Block Groups

Median Income

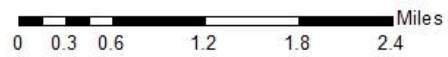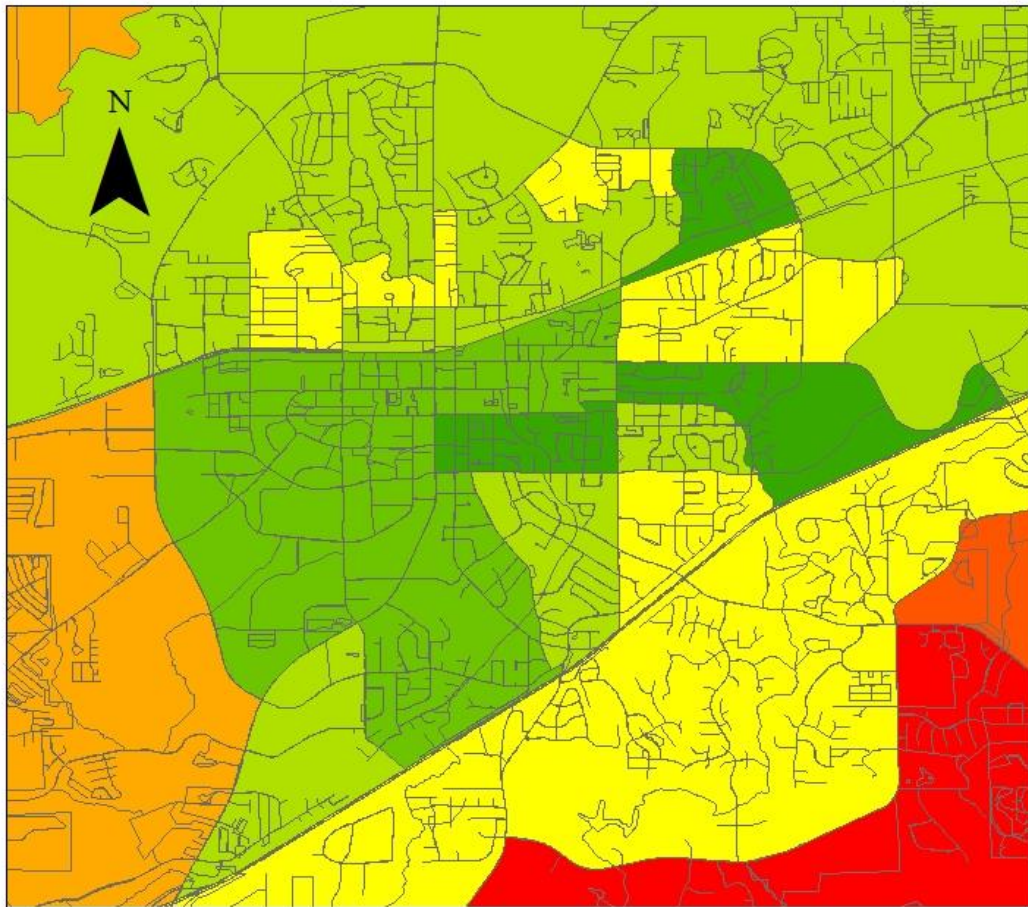|	
🟥	5292 - 30608
🟧	30609 - 47736
🟨	47737 - 70161
🟩	70162 - 109333
🟢	109334 - 206691

Figure A.11: Median Income Census Block Groups

Legend

—— Auburn Streets

Auburn Census Block Groups

Avg_Commute

- 13.50 - 14.00
- 14.01 - 16.00
- 16.01 - 18.00
- 18.01 - 20.00
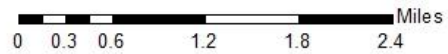- 20.01 - 23.00
- 23.01 - 27.00
- 27.01 - 33.25

Figure A.12: Average Commute Time (min) Census Block Groups