How The Weather, Land Use, and Infrastructure Influence Non-Motorized Mode Choices

…………………………………………………………………………………………………

Matthew Dunlap

…………………………………………….…………………………………………………

Submitted in partial fulfillment of the
Requirements for the degree of
Master of Science in Civil Engineering

University of Washington
2015

Committee:
Yinhai Wang
Anne Moudon
Cynthia Chen
Zhibin Li

Program Authorized to Offer Degree:
Department of Civil and Environmental Engineering

©Copyright 2015

Matthew Dunlap

University of Washington

Abstract

How The Weather, Land Use, and Infrastructure Influence Non-Motorized Mode Choices

Matthew Dunlap

Chair of the Supervisory Committee:

Professor Yinhai Wang

Department of Civil and Environmental Engineering

Compared to many other developed nations, the United States is largely reliant on automobiles for satisfying the daily transportation needs of its citizens. This large demand of auto trips produces excessive amounts of congestion, environmentally harmful vehicle emissions, and does little to help our ongoing obesity epidemic. Galvanized by the problems facing our nation's current non-sustainable transportation system, government stakeholders are now actively promoting non-motorized forms of transportation as viable, and healthy alternatives to the auto for both individuals and our society as a whole. However, in order for the limited resources in non-motorized transportation to be spent wisely on infrastructure improvements, which will have the long term effect of significantly increasing the non-motorized mode share, a more thorough understanding of this understudied mode group is needed. By identifying the factors which influence non-motorized users, transportation engineers and planners will be more capable of eliminating specific deterrents of non-motorized transportation, as well as improving the relative attractiveness of modes such as biking and walking. In this thesis, factors affecting non-motorized mode choice will be examined. Firstly, the effect of weather variables on a population of bicyclists in the arid city of Albuquerque, NM will be analyzed in detail. It will be confirmed that weather conditions are strongly tied to daily bicycle demand and that the hourly bicycle distribution, indicative of the times of the day in which bicyclists choose to ride, shifts with seasonality. These findings have implications as we strive to build and manage livable communities to that are conducive to a full spectrum of modal alternatives. Secondly, the impact of infrastructure and land use on non-motorized mode choices will be investigated for the population of Seattle, WA. Utilizing individual trip information from a randomly administered household survey as well as rich spatial information pertaining the relevant built environment factors such as land density and urban non-motorized shared trails, binary discrete mode choice models are developed which help explain the relationship between the built environment and non-motorized activities. Significant contributing factors are identified and their impacts are quantified in terms of travelers' preferences on non-motorized travel modes. The research findings are helpful for developing appropriate policies and infrastructure deployments for enhancing transportation system sustainability.

Acknowledgements

……………………………………………………………………………………………….

I would personally like to thank Professor Yinhai Wang for his constant support and guidance during my graduate studies. Also, I would like to thank Guohui Zhang, and Shen Zhang for providing much needed assistance, and encouragement through the course of my research studies.

To my parents, my little sister, Megan, my friends, and the many caring mentors that have touched my life over the years: without your strong love and support, I would not be where I am today.

……………………………………………………………………………………..

I. Table of Contents

II. List of Tables & Figures

Chapter 1: Introduction

1.1 Driving Towards Apathy

Compared to many developed countries around the world, the United States is heavily dependent on personal automobiles for satisfying its transportation demands, taking 84% of all trips by car compared to Germany and the Netherlands, for which auto modes constitute less than half of all trips taken (Pucher and Buehler, 2006). Furthermore, U.S. growth in the number of vehicles has far outpaced the growth in households and persons. This is evident from an observation that US households owned an average of 1.86 motor vehicles in 2009 compared to about 1.77 vehicles in 1990. This increase in the number of autos has contributed to our growing traffic congestion, air quality degradation, and increased energy consumption (Sener et al., 2009). In this context, non-motorized travel modes have drawn considerable attention due to their economic and environmental benefits, such as congestion alleviation, environmental protection and energy conservation. Evidence is also mounting on the health-enhancing potential of non-motorized forms of transportation, since walking and cycling are more sustainable and effective means of being active for currently sedentary people (Dunn et al., 1999).

However, in spite of the benefits non-motorized transportation can provide, and the efforts of planning agencies to encourage non-motorized modes, 66% percent of the driving age public never rode a bicycle during the summer months in 2012 compared to 57% in 2002, while the statistical data for the walking frequency also remains unchanged from 2002 to 2012 (Santos et al., 2011). According to the 2009 National Household Travel Survey (NHTS 2009), only 9.7% of all person-trips made in the U.S. relied on non-motorized travel modes, though 40% of total trips are shorter than 2 miles (USDOT, 2009). Actually, the percentage of the population using non-motorized transportation (NMT) has even decreased in the past few years. It is evident that the current strategies aimed at decreasing the mode share of the automobile have been, overall, an unsuccessful endeavor and many planners and stakeholders are looking towards progressive European nations for answers. However, looking up to more progressive nations like the Netherlands can be a little intimidating and frustrating. It begs the question of whether large scale promotion of non-motorized transportation in the U.S. could ever come close. However, is the U.S truly fundamentally different than our less auto-oriented European neighbors, or is it possible for our transportation system and society to evolve.

John Pucher, an urban planning professor at Rudgers University in New Jersey, has devoted much of his career to answering this question and his conclusions are quite insightful. In his paper, "Urban Travel Behavior an outcome of public policy: The Example of Model-split in Western Europe and North America" (Pucher and Buehler, 2006), Pucher compares the mode splits of 12 different countries in North America and Western Europe. As one might expect, Pucher identifies that there lies a large discrepancy between the mode share of auto trips in the U.S., and the mode share of auto trips in Europe and Canada. The auto mode dominates U.S. trips (82% of all trips are auto) compared to Canada (<74% of trips are auto), and for the Swiss, Swedes, Italians, and Austrians, for which auto trips constitute less than 40% of the trips taken. Furthermore, transit and public forms of transportation, for which the vast majority of these trips begin and end with at least some amount of walking, are the least popular mode in the U.S., generating a mode share of less than 3% of all trips. The primary differences identified that Pucher attributes to the disparity of NMT in the U.S. are rooted in land use and transportation policies. One prominent example of this is employer-supplied parking in the U.S. This is a common convenience afforded to many employees and it greatly increases the relative convenience of the auto-mobile by eliminating the need to find and pay for parking. A healthier and more sustainable habit would be for employers to pay for bus passes for employers to encourage bussing, or installing changing facilities and bicycle racks for those individuals who commute by bicycle. In addition, taxes on Gasoline and the purchasing of automobiles are much higher in Europe, than in North America which makes auto travel much less attractive.

Drawn from observations of transportation in the U.S., Pucher concludes that the U.S. has done very little to discourage urban sprawl and encourage denser developments. He states that, "For the most part, development in the U.S. arises haphazardly, as private developers and builders try to maximize their profits with little regard for social and environmental consequences." In a society allegedly governed by and for the people, Pucher states that there is little reason to believe that American policies have always reflected the citizen preferences or long term interests of society as a whole. Thus the main message here is that the future transportation policies and land development in America must be modified to reflect the transportation needs and preferences of society as a whole.

Tradeoffs will need to be made on many sides of the political spectrum in order to create a transportation system on-par with many progressive European cities. A big difficulty facing the U.S. transportation system is the large funding shortage. This shortage of funding stems from an inherent aversion of taxes in the U.S. However, increasing the gasoline and automobile taxes could greatly benefit the current transportation system. Not only would this allow for the repair of current roads, bridges and infrastructure vital to the movement of both people and goods, but it would generate funds for new potential investments in sustainable and active forms of transportation such as transit, walking and bicycling. Infrastructure developments for non-motorized facilities such as bicycle paths, and cycle tracks have been shown to have immense benefits which far exceed the costs of installation; this is not to mention, that the addition of well- designed new facilities have the potential to attract more bicyclists and walkers.

1.2 Biking for Health

As previously mentioned, walking and cycling are considered to be sustainable and effective means of being active for sedentary people (Cervero et al., 2009). In continuing this discussion of non-motorized transportation in the U.S., it is necessary address the issue of public health problems in the U.S. and how active non-motorized forms of transportation can help assuage some of these issues. More specifically, this section will first quantify public health problems related to obesity and the largely sedentary lifestyles of the general American population. Next, a review of the link between exercise and personal health will be presented, and finally, it will be discussed how actively commuting to work or travelling by bicycle can help improve individual health in terms of reducing the risk of heart disease, diabetes, cancer, and a myriad of chronic diseases.

Each year, in the U.S., about 720,000 people have a heart attack. Furthermore, 600,000 people die of heart disease each year which amounts to 25% of the nation's deaths (CDC, 2015).

Factors contributing to this lethal killer include physical inactivity, obesity, and a poor diet. The total cost for treating people with heart disease in the U.S. is currently estimated to be 108.9 Billion dollars per year, which is bad news. The good news however, is that heart disease, and similar chronic disorders such as diabetes are quite preventable. Studies have verified that regular exercise is associated with a decrease in premature death (Lee and Skerrett, 2001, Warburton et al., 2006). Furthermore, evidence from a well cited notable literature on the subject, revealed that exercise resulted in a greater decrease in death than any other cause. Being fit or active was associated with more than a 50% percent reduction in the risk of mortality (Lee and Skerrett, 2001). However, how much exercise is necessary to reap these benefits?

Exercise is commonly defined by medical practitioners in terms of three key components: duration, frequency, and intensity.

These three components are defined below:

Duration: length of time of physical exertion

Frequency: How often a person exercises (times per week).

Intensity: Rate of energy expended during exercise.

All three of these factors are used in determining whether the recommended level of exercise has been met. For example, to find the total energy expended during a week of exercise, the following equation may be used. It is upon this expended weekly energy that markers for physical activity are based.

$$Energy = Intensity \cdot Frequency \cdot Duration \tag{1-1}$$

Where,

$Energy(kcal / wk)$
$Intensity(kcal / hr)$
$Frequency(/\text{ week})$
$Duration(hr)$

The units of exercise *intensity* are commonly expressed in the medical world in MET's. MET stands for metabolic equivalent unit and 1 MET = 1000kcal/hr. According to recent standards published in the American College of Sports Medicine and the American Heart Association, to maintain and promote health, adults require 30 minutes of moderate-intensity aerobic physical activity 5 days/wk. (Haskell et al., 2007). Furthermore, studies recommend that an individual get between 500-1000 MET-min of exercise per week (De Hartog et al., 2010). Furthermore an individual who increases there level of physical activity by 4200 KJ has a mortality benefit of 20% (Lee and Skerrett, 2001).

A person who bikes regularly to work or for personal errands gains the benefit and convenience of both transportation and exercise. A leisure bicycle commute travelling at a pace of 15km/hr. for 7.5 km, 5 days/wk is the equivalent of 600 MET-min/wk of expended energy (De Hartog et al., 2010), and meets the current recommended level of exercise. Thus it is quite attainable for an individual to obtain his/her recommended level of exercise to maintain a healthy lifestyle with a leisurely daily bicycle commute.

However, for many individuals, factors such as the perceived dangers associated with bicycling are enough to discourage individuals from bicycling (Noland and Kunreuther, 1995). Thus obtaining exercise via biking can be a complicated matter. However, one well cited study suggests that overall, the perceived risk of bicycling may be misconstrued. Researchers compared the benefits of cycling to the risks of cycling, to determine overall if cycling regularly both as a form of transportation and exercise does more good than bad (De Hartog et al., 2010). In this study, the benefits of individuals were looked at for the hypothetical scenario if 500,000 Dutch suddenly decided to commute to work every day by bicycle instead of by personal automobile. To assess the costs of bicycling, these researchers looked at both the risk of getting into an accident and also the increased exposure to pollutants. The personal benefits of bicycling included the health benefits of exercise. They concluded that for people who shift from car to bike, the benefits are 9 times greater than the costs by looking at the life years gained by cycling vs. not cycling. Thus the increased physical activity of bicycling outweigh the risks of being more exposed to fatal accidents and breathing in higher dosages of pollutants.

This previous study (De Hartog et al., 2010) shows that, not only is bicycling a healthy and convenient form of exercise, but it's safe to do as well. However, telling your neighbor this information will not necessarily change his mind about preferred daily commute mode; mode choice decisions are complicated processes (which will be discussed in greater detail further on). If one is to successfully understand the large disparity of bicycle and non-motorized trips in the U.S. compared our progressive European neighbors, a more thorough understanding of the factors influencing non-motorized activities is necessary.

Hindrances of non-motorized transportation (NMT) are numerous and include an array of both objective factors such as: weather, presence of daylight, the need to carry bulky goods, the need to arrive well-groomed, age, health, and less quantifiable subjective factors, such as: cultural norms, perceived convenience and dangers, and the built environment. For the purpose of this thesis, it would prove both exhaustive and infeasible to try and capture the effect of all these factors on non-motorized road users. Therefore, the scope of work has been narrowed. This thesis will investigate the response of non-motorized road users to two integral factors of the built environment: land use and infrastructure. In addition, the effect of weather on non-motorized users will also be analyzed.

The remaining content of this thesis is structured as follows: A literature review (Chapter 2) will discuss past research in this area, upon which I have derived my own research and analytical methods. Next, Chapter 3 will discuss how I attempted to quantify the effect of weather on a population of bicyclists in the arid region of Albuquerque, NM, the data I used, the analysis done, and my results. Chapter 4 will discuss how I sought to understand the influence of the built environment on non-motorized mode choices by developing mode choice models with built environment variables. Finally, Chapter 5 a conclusion of the overall research results will be presented and suggestions for future work will be discussed. For the sake of brevity, some details of the analyses for this research have been summarized and simplified, however and appendix been added which document some of the fine details and are referenced in the text for the reader's convenience. Also, for brevity's sake, the term "non-motorized transportation" will commonly be abbreviated as NMT

Chapter 2: Literature Review

2.1 Bicyclists and the Weather

Bicycling is inherently different from other motorized forms of transportation in that its users are readily exposed to the outdoor elements. As a result of this, something as mundane as a shift in the wind or slight temperature changes may sway the volume of bicycle riders seen on any given day. There have been many studies to attempt to quantify the influence of various weather factors on bicyclists. A pioneer of this area was Hanson and Hanson, who wrote a paper studying the impact of weather on bicyclists in 1977 (Hanson and Hanson, 1977). Stemming from this research, others have also conducted duplicate studies for the effect of weather on bicyclists for an array of different climates and unique bicycle populations (Ahmed et al., 2012; Gallop et al., 2012; Hanson and Hanson, 1977; Lewin, 2011; Lindsey et al., 2007; Mirada-Moreno and Nosal, 2011; Nankervis, 1999; Niemeier, 1996). In order to study the “weather effect”, these researchers utilized bicycle count and weather data, and analyzed the data using regression techniques. This research has almost without exception concluded that bicyclists are significantly affected by the weather. In comparative studies between different types of bicyclists, recreational bicyclists were found to be more significantly impacted by inclement weather than commuter cyclists riding to work (Gallop et al., 2012). One prominent study compared bicyclists in Portland, OR to those of Melbourne, Australia and concluded that bicyclists were sensitive to weather changes and that the magnitude of influence was relative to each city’s base climatic conditions (Ahmed et al., 2012).

Another branch of research relevant to understanding the effect of weather on bicyclists is rooted in transportation safety. A fair quantity of research has been done to understand the interaction between weather and other environmental variables on bicycle-vehicle collisions (Brezina and Kramer, 1970; Schroeder and Wilbur, 2013). Although this research is important, it is out of this study’s scope and is acknowledged solely for academic purposes.

As was previously mentioned, there have been several researchers who analyzed this “weather effect” using count and weather data through regression techniques (Ahmed et al., 2012; Gallop et al., 2012; Hanson and Hanson, 1977; Lewin, 2011; Lindsey et al., 2007; Mirada-Moreno and Nosal, 2011; Nankervis, 1999; Niemeier, 1996). I will review six of these studies most relevant to this paper: Ahmed et al., 2012; Gallop et al., 2012; Lewin, 2011; Mirada-Moreno and Nosal, 2011; Nankervis, 1999; Niemeier, 1996.

Table 1, below, shows the different locations and the climates for each of the 6 most relevant studies that were reviewed. As can be noted from this table, much of the research in this area has been done in coastal cities like Portland, Seattle, and Vancouver that receive relatively high amounts of precipitation. The study conducted by Lewin in Boulder, Co., presents the first study of its kind in an arid climate (Lewin, 2011).

Reference	**Location**	**Climate**
Ahmed et al., 2012	Portland, Oregon	Cool-Summer Mediterranean Climate
Lewin 2011	Boulder, Colorado	Cold Semi-Arid Climate
Ahmed et al., 2012	Brisbane, Australia	Humid Subtropical Climate
Niemeier, 1996	Seattle, Washington	Oceanic Climate
Gallop et al., 2012	Vancouver, Canada	Oceanic Climate
Moreno, 2011	Montreal, Canada	Humid Continental Climate
Nankerivis, 1999	Melbourne, Australia	Moderate Oceanic Climate

Table 1-1: Locations and Climates of Previous research. (Note: all climates are based on the Koopean classification system)

In order to quantify this “weather effect”, the researchers listed above utilized daily/hourly aggregated bicycle count data collected for a period of at least one year, and also comprehensive weather data that was obtained through meteorology agencies. With the exception of Nankervis (Nankervis, 1999), who obtained count data manually by counting the number of bicycles on stationary bicycle racks at several college campuses, all of the bicycle volume data was obtained through automatic counting methods such as loop detectors and pneumatic tubes. (Video-based detection methods were not utilized in any of the studies reviewed.)

Various regression techniques were used to develop bicycle demand models. In these models bicycle demand is explained as a function several independent weather variables or predictors. Linear regression was the most popularly used regression technique of those reviewed. Other regression techniques used to develop bicycle demand models include Poisson regression, and negative binomial regression, whose ability to predict non-negative bicycle count values is commendable. A less common ARIMA model (autoregressive integrated moving average model), popular for predicting future stock market values was also used (Gallop et al., 2012).

The vast majority of these literatures looked at the effect of weather on daily bicycle volumes. In their bicycle demand models, weather parameters for a specific day can be inputted in order to output a predicted bicycle volume for that day. The exception to this was Moreno (Mirada-Moreno and Nosal, 2011), whose model predicts hourly bicycle volumes based on hourly aggregated weather and count data. His model provides unique insight into how bicyclists are affected at the hourly level by changes in the weather. For this thesis, an analysis will be presented which determines how the hourly distribution of bicyclists changes with temperature throughout the year.

The findings of these six studies, listed in Table 1-1, vary based on the model used, climate of the study location, and the population of bicyclists. In Gallop et al., an autoregressive moving average model (ARIMA) was utilized. The advantage of using this model over the more conventional ordinary least squares model is that it is able to account for the complex autocorrelation patterns of the error terms. Gallop's ARIMA model results are compared against a base null model developed using OLS regression. The comparison clearly shows that OLS regression tends to greatly overestimate the effect of weather variables on bicycle demand, yielding coefficients that are significantly larger than those determined while accounting for the complex auto-correlation inherent to the data. According to the final model, a 1 degree Celsius increase in temperature corresponds to approximately a 1.65% increase in bicycle demand (all else remaining constant). Variables found to be statistically significant include: rain, rain in previous three hours, temperature, humidity, and clearness, although clearness only had a slight effect on overall bicycle demands and the coefficient magnitudes of humidity were relatively small. A survey distributed prior to this analysis suggests that inclement weather may have a greater effect for the going-trip compared to the return trip. This result makes sense since alternative mode choices of bicyclists are limited due to the constraint that the bicyclist will most likely want to transport their bicycle back home.

Moreno et al. conducted a similar study using data collected from automated bicycle loop detectors installed along several prominent bicycle paths in the Montreal area. Using this hourly count data, they developed hourly bicycle demand models using: precipitation, temperature and humidity and all had significant effect on bicycle ridership. Temperature was found to have a large significance on demand; when the temperature doubled, a 43-50 percent increase in bicycle volume resulted.

Perhaps when it comes to rain, there is no better place than the Emerald city of Seattle. Beddie A. Niereier from the University of Washington studied the effect of climate and weather on a utilitarian population of bicyclists in Seattle, WA (Niemeier, 1996). Using data collected from several on-trail bicycle counters, volume data was collected over a period of several months. The data was then analyzed against weather and temporal variables using Poisson regression. She determined that counts will decrease by 15-25% on rainy days or days with temperature less than 55 degrees Fahrenheit.

A similar study was done by Nankervis in 1999 using demand data from parked bicycle counts (Nankervis, 1999). The parked bicycle count data was collected from bicycle racks at a university in Melbourne Australia over a period of several years. Data was then graphed and fit with a straight-line to determine the general trend of bicycle demand and the time of year. As one would expect, bicycle ridership was shown to be significantly correlated with time, with bicycle demands reaching their peak in the fall the summer months and dropping to a low in July during the colder winter months. A categorical weather scale was created which ranks different types of weather into four main categories based on three principal weather variables: wind, rain and temperature. The author felt that a person's decision to ride will be based on the combined effect of several weather variables and not solely on one variable alone. For instance, a bicycle may choose to bike in hot weather, but avoid biking if it is hot and also windy. The results of this "combined weather construct" showed a significant statistical relationship, and yielded a Pearson's R-squared value of -0.38.

A bicycle survey study was also conducted to determine bicyclists' reactions to different types of weather. A total of 64 riders were interviewed for the survey. The results indicate that there is a 25% deviation from bicycle mode-share in adverse weather. The survey also revealed that rain had the most significant influence on an individual's decision to ride, with 68.7% saying it has "some effect" on their decision to ride. The general consensus of riders who were surveyed is that neither climate, nor weather need to be serious barriers to bicycling. It is curious then, that the results from several quantitative studies show a strong correlation between weather and bicycle ridership indicating that weather heavily impacts a bicyclists' decision to ride. Research from a recreational bicycle population in Albuquerque, NM in a unique arid climate, to be presented in this thesis, will add to this blooming body of research and help corroborate previous findings.

2.2 Built Environment Factors and NMT

Perhaps equally as mundane as the weather, but even more relevant in terms of its influence on not only bicyclists, but all forms of non-motorized travel, is the built environment. In designing transportation systems and cities that cater towards healthy and environmentally friendly non-motorized activities, it is important for law makers and transportation planners consider how this group of road users will react to specific changes in their environment. The goal is to design better and more livable communities that are pedestrian and bike friendly. Planners often identify neighborhoods as pedestrian-oriented if they have a high density of development that include a variety of land uses, a street network with high connectivity, human scale streets and pleasing aesthetic features (Saelens and Frank, 2003). These attractive qualities of a neighborhood can be simplified into what urban planners refer to as the 3D's. The "3 D's" are essentially the "holy trinity" of urban planning and consist of: *density*, *diversity*, and *design*. The actual definitions of each of the 3D's are somewhat subjective and open to interpretation, but for the most part, they can be defined as follows:

- *density*: quantity of persons or businesses that occupy a given area of space
- *diversity*: how many different types of businesses exist in within a given area of space,

- *design*: the aesthetics of how that area looks, whether it includes bicycle and pedestrian facilities, etc.

Cities that have been successful in promoting non-motorized transportation, have likely thoughtfully incorporated each one of these D's into their transportation design plans.

Some have attempted to quantify the importance of various built-environment related factors on non-motorized road users. Mode choice models were commonly used for this task and output the likelihood, $P_i(j)$, individual j will choose the given mode i , given a set of variables containing a mix of built environment, demographic, attitudinal variables, and other factors such as weather and darkness. The generalized mode choice model form is shown:

$$P_j(i) = f(\text{built environment, demographics, attitude, other factors})$$
where
$i \in I$, a discrete mode choice set

In developing these mode choice models from disaggregate trip survey data , researchers have come to the general consensus that although denser, more diverse land types, with pedestrian and bicycle facilities (sidewalks, bike paths/ bike lanes), generally tend to increase one's willingness or propensity to choose non-motorized or transit transportation, demographic factors, and other exogenous factors, such as the topography of the landscape are much more influential on a persons' decision to choose non-motorized modes of transportation (Aker and Clifton, 2009; Cervero and Duncan, 200; De Hartog et al., 2010; Handy et al., 2002; Moudon et al., 2005; Rodriguez and Joo, 2004; Saelens and Handy, 2008; Saelens and Frank, 2003).

Chapter 3: Bicyclists and the Weather

3.1 Introduction

In urban areas, the bicycle often offers a convenient alternative to the personal auto-mobile, and in congested, and dense areas, bicycling may also be a more efficient form of transport. However, travelling by bicycle is not without its risks. Dangers posed by swift moving automobiles, as well as unfavorable weather conditions like rain and snow, (depending on the regional climate) often make bicycling not for the faint of heart. There exists a strong association between the weather and bicyclists. It has been noted in several previous studies that weather affects bicyclists, not only on a seasonal basis, but also a daily level as well (Ahmed et al., 2012; Gallop et al., 2012; Hanson and Hanson, 1977; Lewin, 2011; Lindsey et al., 2007; Mirada-Moreno and Nosal, 2011; Nankervis, 1999; Niemeier, 1996). Some key questions investigated in these studies include:

- What types of weather effect bicyclists the most and the least?
- What is the optimal riding temperature for bicyclists?
- How are bicyclists influenced by seasonality?
- How do bicyclists react to weather in different climates?

This study summarizes an investigation of the effects of weather on a population of bicyclists in Albuquerque, NM. It should be noted that Albuquerque has an arid climate which makes it a unique setting for this study, as the vast majority of previous studies have been in wet coastal cities. This study borrows knowledge from past research in order to develop a daily bicycle demand model to investigate the "weather effect". A novel study on the hourly effect of temperature on bicyclists is also conducted and its significance is discussed in detail.

The rest of this chapter has been written as follows. First, the climate of the study location is discussed as well as the sources of data used for the analysis. Next, the significance of temperature and precipitation on bicyclists is confirmed using OLS regression. Finally the effect of seasonal temperature on the hourly distribution of bicyclists is examined in detail using a novel analysis method.

3.2 Climate of Albuquerque

In order to appreciate this study, one must first understand the climate of Albuquerque, New Mexico. The city of Albuquerque receives on average less than 10 inches of rain per year and is formally classified under the Koopean classification system as a cold arid climate. As is illustrated below in Figure 3-1, which shows the monthly distribution of rainfall, Albuquerque receives the most rain in June and July. This period is commonly referred to by New Mexicans as the monsoon season. These “monsoon” rain storms often occur in the form of hard thundershowers.

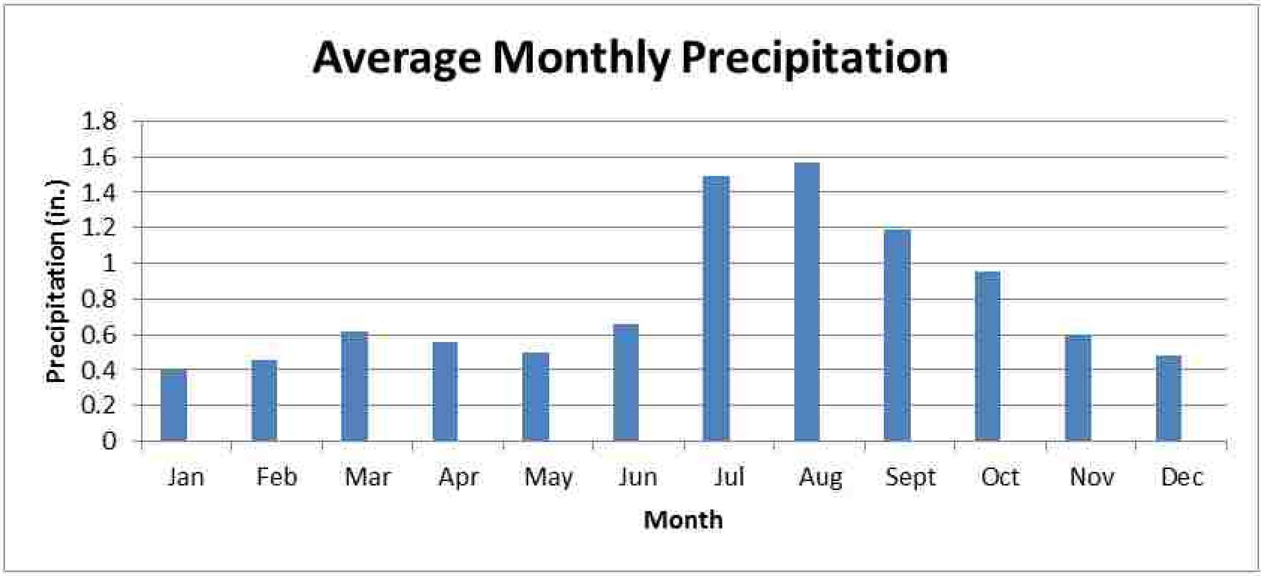

Figure 3-1: Average Monthly Precipitation for Albuquerque, NM.

One inherent characteristic common to many arid climates is the large difference between the daily maximum and minimum temperatures. The average annual daily temperature range for Albuquerque is 25.6 (ºF). For comparison purposes, the annual average daily temperature range and the average annual precipitation for several prominent US cities is shown below (Figure 3-2). Note that Seattle and Portland Oregon, previously studied by Ahmed, and Niemeier, have significantly smaller daily temperature ranges compared to Albuquerque and other dry cities. It can be noted from this graph that, overall, the cities receiving less annual precipitation have greater temperature ranges than those receiving larger annual precipitations.

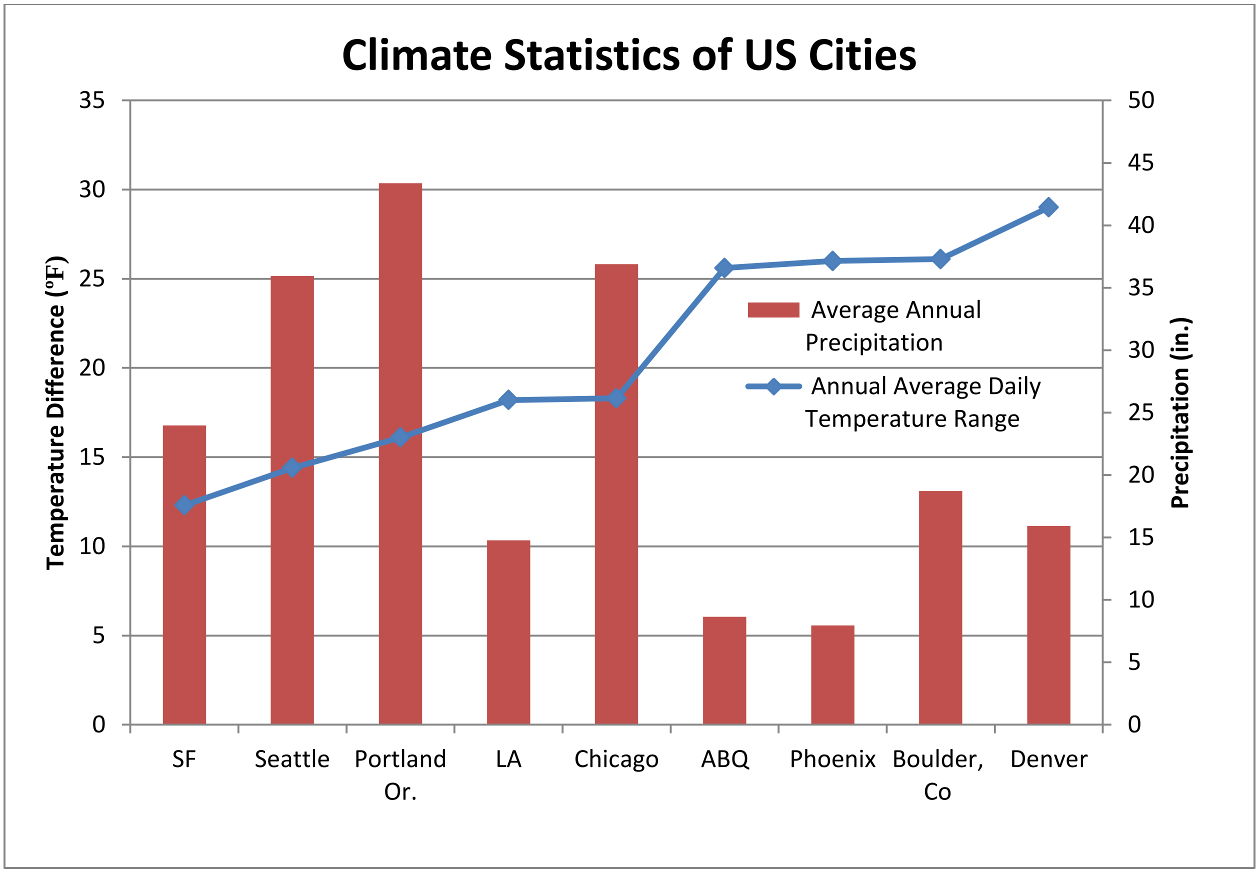

Figure 3-2: Average annual daily temperature ranges and annual precipitations of 9 US cities.

The Daily temperature range is relevant to this study because outdoor activities such as biking and walking become less tolerable in the cold morning hours during winter months, but in the hotter summer months, these activities switch to being more tolerable in the cooler morning and evening hours. The daily temperature ranges in Albuquerque are shown in Figure 3-3 for each month of the year. This graph illustrates a consistently large temperature range for all months of the year. Our hypothesis is that, over the course of the year, bicyclists will shift the time they choose to ride to mirror those times of the day that have more favorable temperatures. This phenomenon is likely to be most prominent in recreational cyclists who have more flexibility in when they make a trip as compared to commuter cyclists. No other studies to the author's knowledge have analyzed the significance of daily temperature variation on bicyclists.

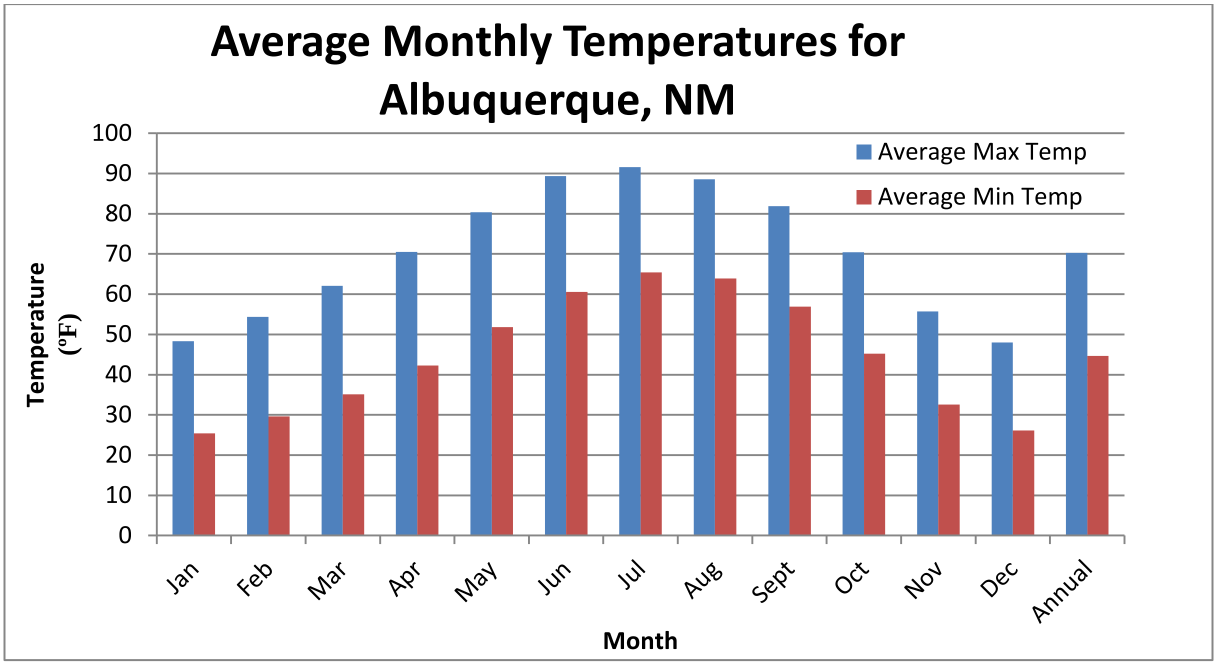

Figure 3-3: Average monthly temperatures for Albuquerque, NM.

Both climate data and bicycle count data were required in order to conduct this study. These sources of data are discussed in detail below.

3.3 Climate Data

The weather dataset for Albuquerque, NM was obtained through the office of the Western Regional Climate Center which is under the National Oceanic and Atmospheric Administration (NOAA). This dataset contains the maximum and minimum temperature, and cumulative rainfall and snowfall for each day of the analysis period.

3.4 Bicycle Count Data

The bicycle count data acquired for this study was collected using loop detectors that are sensitive enough to detect bicyclists as they cross over them. The collection took place at one location on the Bosque bike path, a recreational bicycle trail that parallels the Rio Grande River. This scenic bike trail attracts predominantly recreational cyclists who use the path as a convenient alternative to sharing the road with vehicles. The count data that was received is not perfectly continuous and stretches from 8/23/2012 – 10-08-2012 and 11/9/2012- 6/4/2013. Overall, this amounts to about 8 months of count data.

In the next section, the results of the exploratory data analysis are presented. This step is necessary in order to confirm the significance of weather on bicyclists in the city of Albuquerque. It also validates the results of the bicycle demand model that is developed to quantify the "weather effect".

3.5 Exploratory Analysis

The relationship between temperature and bicycle volume was first investigated by making a graph of these two variables side-by side as shown in (Figure 3-4). This graph shows a general increase in the number of bicycle riders from January to June as the temperature between these months gradually increases.

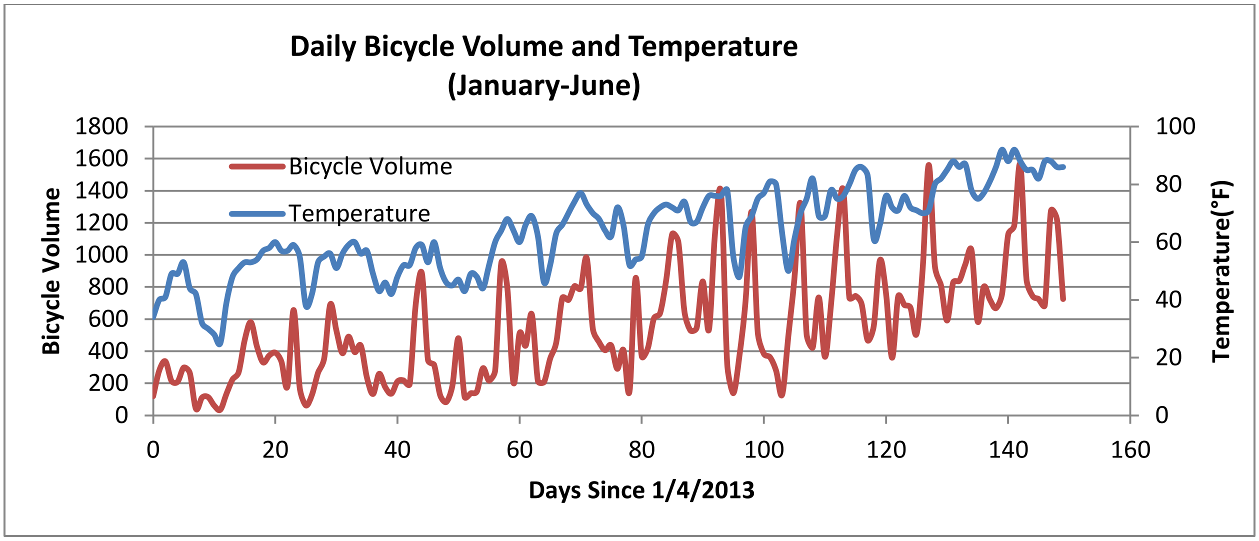

Figure 3-4: Daily Bicycle Volume and Temperature vs Time

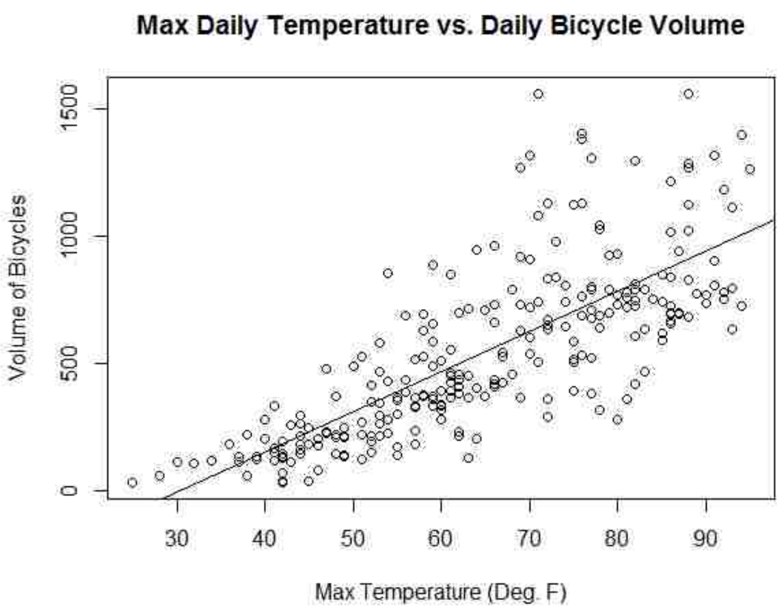

Figure 3-5: Max Daily Temperature vs. Daily Bicycle Volume. (Line of best fit passes through points to illustrate relationship)

Figure 3-5 illustrates the high level of correlation between temperature and the number of bicyclists. The correlation coefficient between the daily bicycle volume and the temperature was calculated to be 0.76 which is consistent with a high degree of correlation.

Weekly trends in bicycle ridership were also investigated. The average daily bicycle volumes the different days of the week are illustrated is shown below (Figure 3-6). As is clear from the graph, the weekends tend to receive much higher bicycle volumes than weekdays. Also, of the weekday volumes, Monday receives the greatest volume and Friday receives the least volume. The Friday

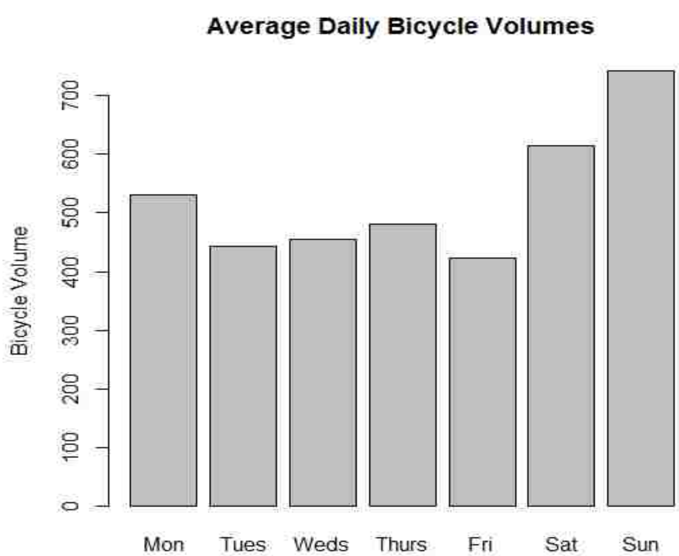

Figure 3-6: Average daily bicycle Volumes for different days of the week.

drop in Bicycle ridership is consistent with the findings of Ahmed, who noted that in both Portland, and Brisbane, Friday received the lowest bicycle volumes of all the weekdays (Ahmed et al., 2012).

Previous research has classified bicyclist into these two categories based on the ratio of weekday traffic/ weekend traffic (Ahmed et al., 2012). In our data set, the Weekday/Weekend bicycle traffic ratio was calculated to be 0.69, which is less than 1 and is indicative of a predominantly recreational bicyclist population (Ahmed et al., 2012). This makes good sense, for the Bosque bike trail is located in a remote area, far from many of the major business and employment centers where residents are likely to work.

3.6 Analysis and Results

To quantify the significance of weather on bicyclists, a bicycle demand model was developed using ordinary least squares regression. The explanatory variable of daily bicycle counts was modeled as a function of two different weather parameters, temperature and rainfall. To account for the effect of varying demand across the different weekdays, six "day of week" variables were included as explanatory variables. The seventh day of the week, Sunday, was used as the reference day, and was not included. The general form of the fitting function is shown below:

$$y = \alpha + \beta_1 R + \beta_2 T + \beta_{3,\mathrm{n}} \qquad (3\text{-}1)$$

Where

$$\begin{aligned} y &= \text{Number of Bicyclists} \\ R &= \text{Precipitation (inches)} \\ T &= \text{Temperature } (^\circ\text{F}) \\ n &= \text{Day of Week} \end{aligned}$$

Several different combinations of rain and temperature variables were experimented with in 4 different regression models. The model results are shown in (Table 3-1). The coefficients of each variable are shown in the table and their associated t-values are shown in parenthesis. (All the significant variables are highlighted.) Model 1, equivalent to the fitting function above, contains linear terms for both rain and temperature and will henceforth, be referred to as the linear-base model.

A rain squared and temperature squared term were examined guided by the findings of Ahmed who noted that a rain squared term performed the best out of the many terms he tried and also noted that the relationship between ridership and temperature was parabolic in shape (Ahmed et al., 2012). However, from our results, Ahmed’s findings could not be confirmed. The goodness of fit (measured by the r-squared value) for the model containing the rain squared term (Model 4) and the goodness of fit for the model with the temperature squared term were nearly identical to the linear-base model.

Postulating that bicyclists in Albuquerque may be more affected by the presence of rain rather than the actual quantity of precipitation, a binary rain term was also experimented with in Model 2. However, this variable was found to be statistically insignificant.

Explanatory Variable	Units	Model1 (linear-base)	Model2	Model3	Model4
Rain	Riders/inch	-1325.5		-1324.0	
		(-2.776)		(-2.704)	
Rain Squared	Riders/inch^2				-5783.9
					(-1.929)
Rain(Binary)	------------------		-50.1		
			(-1.166)		
Temp	Riders/°F	16.0	15.9		16.0
		(-22.43)	(-22.009)		(-22.23)
Temp Squared	Riders/(°F)^2			0.1	
				-21.6	
Day of Week					
Mon	------------------	-219.4	-219.0	-217.7	-218.2
		(-5.155)	(-5.077)	(-4.989)	(-5.084)
Tues	------------------	-278.6	-281.4	-281.1	-277.9
		(-6.454)	(-6.427)	(-6.352)	(-6.386)
Wed	------------------	-319.9	-326.6	-315.8	-319.0
		(-7.295)	(-7.335)	(-7.024)	(-7.211)
Thurs	------------------	-310.4	-326.2	-304.9	-312.7
		(-7.147)	(-7.433)	(-6.85)	(-7.132)
Fri	------------------	-339.1	-350.9	-335.4	-345.0
		(-7.926)	(-8.138)	(-7.648)	(-8.02)

Sat	------------------	-145.6	-144.7	-144.8	-146.1
		(-3.42)	(-3.351)	(-3.318)	(-3.405)
R-Squared		0.70	0.70	0.69	0.70

Table 3-1: Model results of the 4 trail models tested in the analysis. Coefficient values of each variable are shown (upper value) and the t-values for each variable are also shown (lower value). All highlighted variables are statistically significant based on a t-distribution to the 95% level of confidence.

Each of the R-Squared values was determined to be nearly equivalent in each of the four models tested (roughly 0.70). This R-squared value denotes that nearly 70% of the model variance can be explained by the explanatory variables.

According to the results of Model 1, a 10 degree increase in temperature will increase the number of bicyclists by about 160. Model 1 also indicates that the presence of 0.1 inches of rain will decrease the volume of daily bicyclists by about 130. It should be noted that, although rain was determined to be statistically significant in several of the trial models, the number of days in the analysis period that received rain was very small, so these results should be taken with a grain of salt.

Above, we discussed that a natural property of desert climates is the large daily temperature range between the daily maximum and minimum temperatures. Due to this, we previously postulated that recreational cyclists, unbridled by time constraints, will choose to ride during the time of day with the most comfortable temperature. In the hotter summer months, the most comfortable time of day would be in the cooler morning and evening hours, and in the colder winter months, the warmer afternoon hours would be most comfortable for riding.

The hourly distribution of bicyclists for different months of the year was examined to test this hypothesis that bicyclists might shift their riding times throughout the year to times of the day having the most favorable temperatures. A graph of the daily distribution of bicyclists for several different months is shown in Figure 3-7. In colder months like January and December, the distribution has a single peak occurring in the middle of the day, but in the warmer months of May and September, the distribution has 2 peaks (bimodal). For the month of September, there is a morning peak which occurs at about 10:00AM and an afternoon peak that occurs at about 5:00PM. In contrast, in January, the coldest month of the year, a single peak occurs at 3:00PM. There is a gradual transition between the single peak distribution (colder months) and the bimodal distribution (warmer months). Unfortunately, there is no data presently available for June and July, which are the hottest months of the year. Nonetheless, this trend appears to confirm our hypothesis. The most comfortable riding temperatures for bicycling are in the cooler morning and evening hours during the summer, and shift to warmer afternoon hours during the winter.

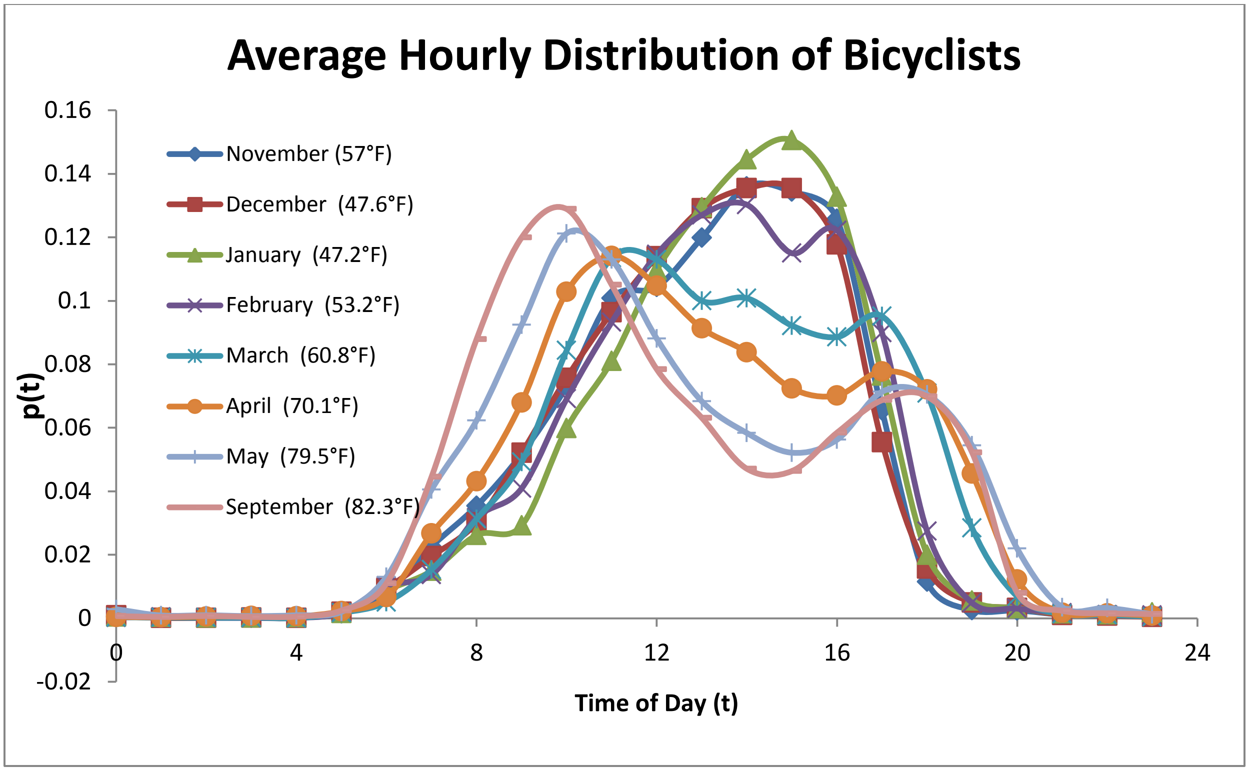

Figure 3-7: Distribution of bicyclists for different months of the year. (Average daily high temperatures in parenthesis.) Lines through data points calculated with spline to guide the eye.

From this distribution of bicyclists, it seems that two groups of bikers are being analyzed. One group of bikers ride early in the day (morning riders), and another group of bikers enjoy riding later in the day (afternoon riders). To further analyze this result quantitatively, each monthly distribution was deconvoluted into two separate distributions to represent the two rider groups. The deconvolution was done by fitting the sum of two Gaussian functions to the data using the method of least squares. The Gaussian sum function that was used is shown below:

$$f(t) = \frac{A_1}{\sigma_1\sqrt{2\pi}} e^{\frac{-(t-\tau_1)^2}{2{\sigma_1}^2}} + \frac{A_2}{\sigma_2\sqrt{2\pi}} e^{\frac{-(t-\tau_2)^2}{2{\sigma_2}^2}} \tag{1-2}$$

where σ_1, σ_2; τ_1, τ_2; and A_1, A_2 are the variances, means and relative weights of the two distributions representing the morning and afternoon rider groups respectively and t represents the time of day a count was recorded.

Figure 3-8illustrates the method of least squares fit for the distribution of bicyclists in the month of September.

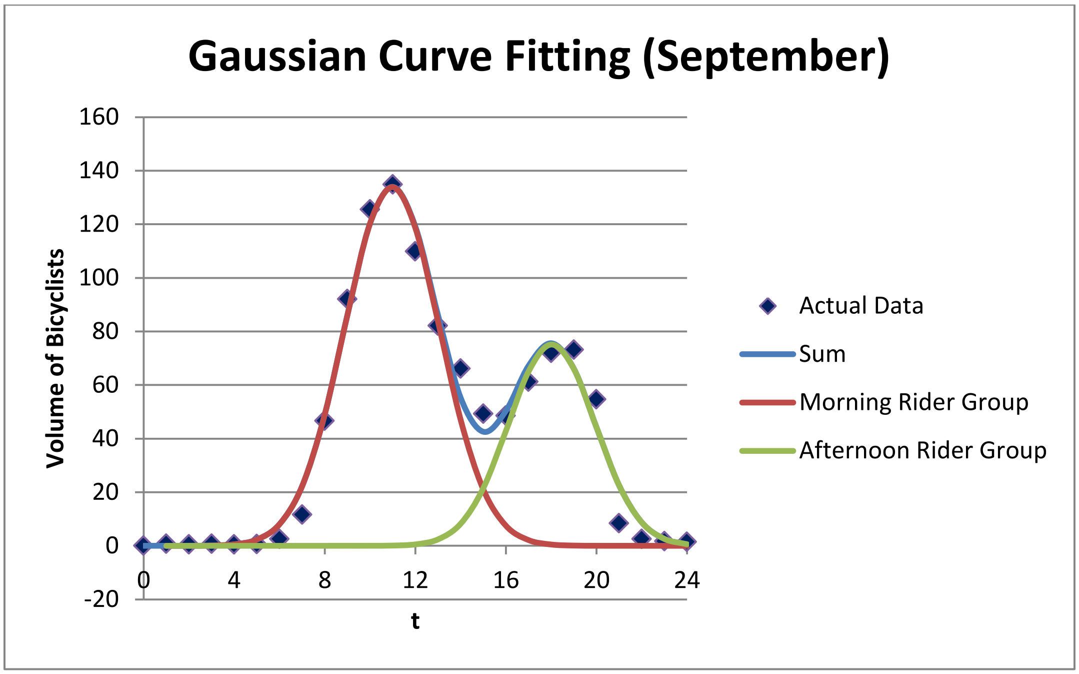

Figure 3-8: Gaussian Curve Fitting of Bicyclist Distribution

Of the two statistical parameters extracted from the fit (variance and mean), only the two means were further analyzed. The assumption was made that the later mean time represents the afternoon rider group and the earlier mean time represents the morning rider group.

The means of each of the two fitted Gaussian curves are plotted below, as a function the average monthly temperature (Figure 3-9). The left graph shows that the mean riding times of the morning rider group undergo a three hour shift with season. The negative slope shows that warmer temperatures are associated with earlier riding times. For the afternoon riders (right graph) the slope is positive, showing that warmer temperatures are associated with later riding times.

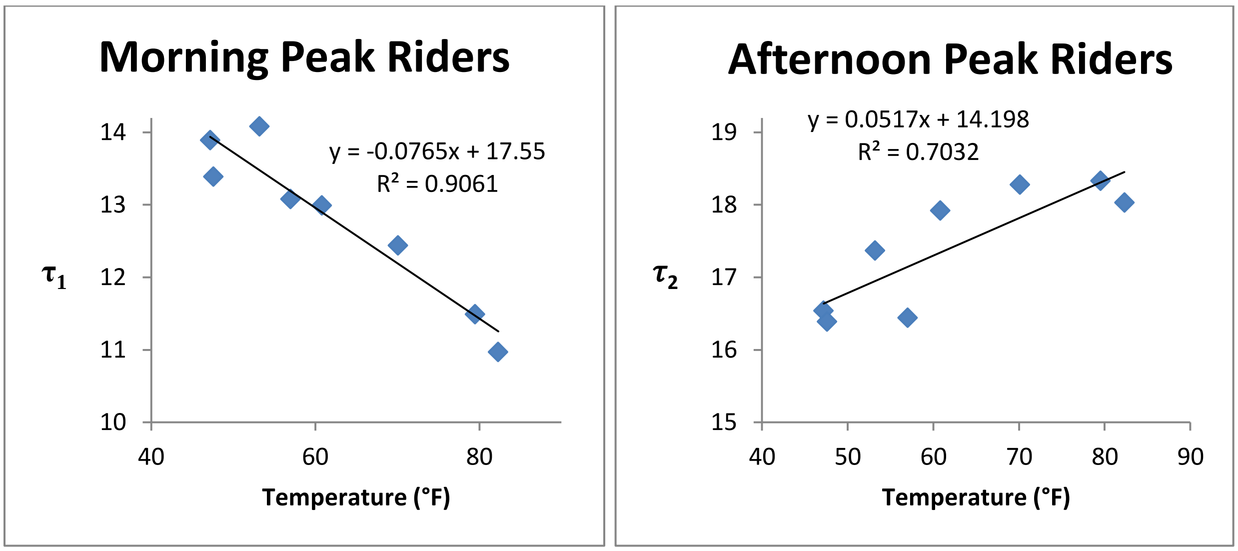

Figure 3-9: Left graph shows the mean riding times for the morning peak rider group as found from the Gaussian curve fitting. The negative slope shows that they prefer to ride earlier in the day during warmer months. Right graph shows the mean riding times for the afternoon rider group, and here the positive slope shows that they prefer to ride later as the days get warmer.

3.7 Summary of Findings

A bicycle demand model was created in order to examine the significance of weather on bicycle ridership using ordinary least squares regression. This model was able to explain about 70 percent of the variation in daily bicycle volumes and is comparable in “goodness of fit” to the other literatures. It should be noted that the model results are limited by the lack of a complete year’s data. In addition, having bicycle data for the months of July and August during

Albuquerque's "monsoon season" might have yielded interesting results to explain how bicyclists react to hard rains and thunderstorms.

The developed bicycle demand model gives relevant insight into how daily bicyclist volume changes with weather. It should be noted that this model is location specific and pertains to a predominantly recreational bicycle population. Further research was also done to analyze the effect of temperature changes throughout the year on the hourly distribution of bicyclists. The hourly distribution of bicyclists was analyzed for each month of available data. These distributions were deconvoluted into two bicycle groups, a morning rider group and an afternoon rider group to determine the average riding times of each group. A shift in average riding times was observed for warmer and colder months and this confirms the hypothesis that bicyclists do indeed shift their riding times throughout the course of the year as a result of changing temperatures. If we had data for June and July, it is expected that the hot temperatures of these months would have caused even greater shifts in riding times for the morning and evening rider groups and would have further verified our result.

This phenomenon of shifting bicycle riding times is important from a transportation safety perspective. The rising and setting of the sun can blind both bicyclists and vehicles making the evening and morning hours more dangerous times for bike riding. This study shows that recreational bicyclists are more likely to ride during morning and evening hours to avoid high temperatures during the summer months, which suggests that bicyclists have a higher risk of an accident with a vehicle during this time of the year. Further research might investigate the correlates of bicycle vehicle accidents to see if a shift in riding times due to temperature is a significant factor of bicycle-vehicle collisions.

Chapter 4: Urban Non-Motorized Travel Mode Choice

4.1 Introduction

As was previously noted, other countries like the Netherlands and Germany utilize non-motorized transportation on a much greater level than people in the U.S. Actually, in spite of Canada's colder climate, Canadians utilize NMT for work trips significantly more than Americans, and bike to work nearly three times more often than Americans and walk to work twice as often (Pucher, Transportation Quarterly). A comparison study between the U.S. and Canada concluded that the stricter land use and transportation policies present in Canada are the primary reason behind their higher cycling levels (Pucher and Buehler, 2006). Evidence from a study in the Netherlands found that transportation policies affecting the safety of cyclists and the average number of stops during a route helped explain varied bicycle demand between different municipalities (Niemeier, 1996). These transportation and land use policies are integral in shaping the built environment around us, and factors of the built environment such as denser, and more diverse land uses are thought encourage more transit and non-motorized trips.

If future public policies are to successfully promote non-motorized transportation, and better utilize the limited assets earmarked to do so, a more fundamental understanding of the relationship between NMT and the built environment in the U.S. is vital. Adding to the limited body of work in this area, a second research study is presented here which attempts to quantify the built environment vs. NMT relationship for the population of Seattle, WA. Guided by the previous literature in this field, disaggregate demand models, developed from individual trip information collected from the 2006 regional household survey and GIS spatial data describing several prominent built environment factors, for the purpose of quantifying the effect of different built environment factors on non-motorized mode choices. This study adds to the existing literature and helps to paint a more complete picture of non-motorized transportation and the built environment around us.

This specific case study is structured as follows: a description of the datasets used and the data analysis process is explained, followed by an illustration and description of the analysis results. Lastly, a discussion of the implications of these results is given, as well as suggestions for further research.

4.2 Dataset Description and Compilation

The 2006, Puget Sound Region Council's (PSRC's) house hold travel survey was instrumental for conducting this study. (The more current 2014 PSRC travel survey data only recently became available after the time this research was completed). This vast databank contains 48 hour trip diaries for individuals living in 4,600 randomly selected households throughout the region. In addition to individual trip attributes, this dataset also includes basic demographic information for each trip maker, as well as to latitude and longitude coordinates of all trip origins and destinations. Only a subset of this data, consisting of approximately 16,000 individual trips, which were under a five mile threshold distance and that were located within the Seattle city limits, was selected for further analysis. The five mile threshold distance encompassed the vast majority of all biking and walking trips and was chosen in order to exclude excessively long trips requiring an unreasonably high level of physical effort.

Using the trip location information obtained from the trip survey data, trip origins and destinations were plotted to an ArcGIS map that included an ArcGIS feature data set containing the following layers:

-Prominent Bicycle and pedestrian paths and trail locations within the Puget Sound region

-Roadway network data

-Topography data

-Crosswalk location data

(ArcGIS feature data set layers were obtained from the Washington Geodatabase Archive (WAGDA) online data archive)

Figure 4-1 illustrates the bicycle/ shared path network in Seattle. The four path types shown in this map are: “StripedBikeLane2Lanes”, which indicates the presence of a striped bike lane on both sides of a road also shared with automobile traffic, “BicycleTrack”, an exclusive bicycle path reserved only for bicyclists, “StripedBikeLane”, a striped bicycle lane on only one side of a road, and “SharedPath”, an exclusive path which serves all forms of non-motorized transportation.

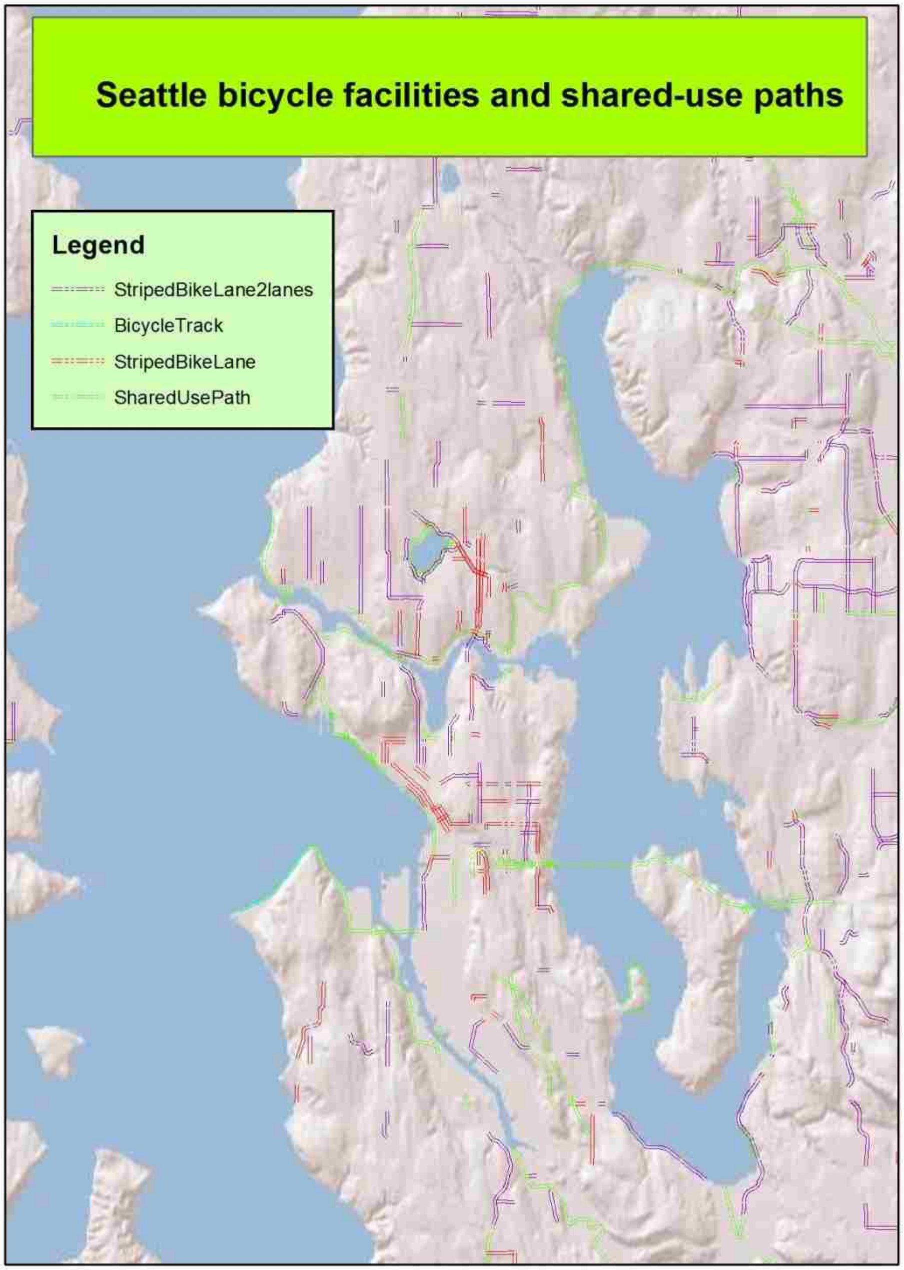

Figure 4-1: Map of Bicycle Facilities and Shared-Use paths

Guided by the prior studies reviewed, the following built environment attributes were extrapolated for the origin and destination of each trip using built-in spatial analysis tools included in the ArcGIS software:

Block density, distance to nearest path, and cross-walk density.

A brief description of each of these variables, including why they are relevant, and details on how they were calculated follows. A detailed step-by-step description of the spatial analysis and data compilation can be found in Appendix I.

Block Density:

Conventional urban landscape and planning policies emphasize the "three D structure: *density*, *diversity*, and *design*. The school of thought is that through densifying urban areas, as well as building diversely, so that multiple land uses are close together, while being mindful of the different needs of the people in an urban region, a better community can be established. In this study, block density is used as an indicator of overall land density described in the three d land use policy structure.

Individual blocks were identified by adjoining intersecting roadway segments; a block was defined as a space enclosed on all sides by road segments. The map below (**Figure 4-2)**, illustrates the block size geography in the Seattle area. These areas are broken into three categories: *small blocks* (<1.6 acres), *medium blocks* (>1.6 acres and <4.2 acres), *large blocks* (<25.3 acres and >4.2 acres). (For size reference, 1 acre is approximately 2/3rds of a football field)

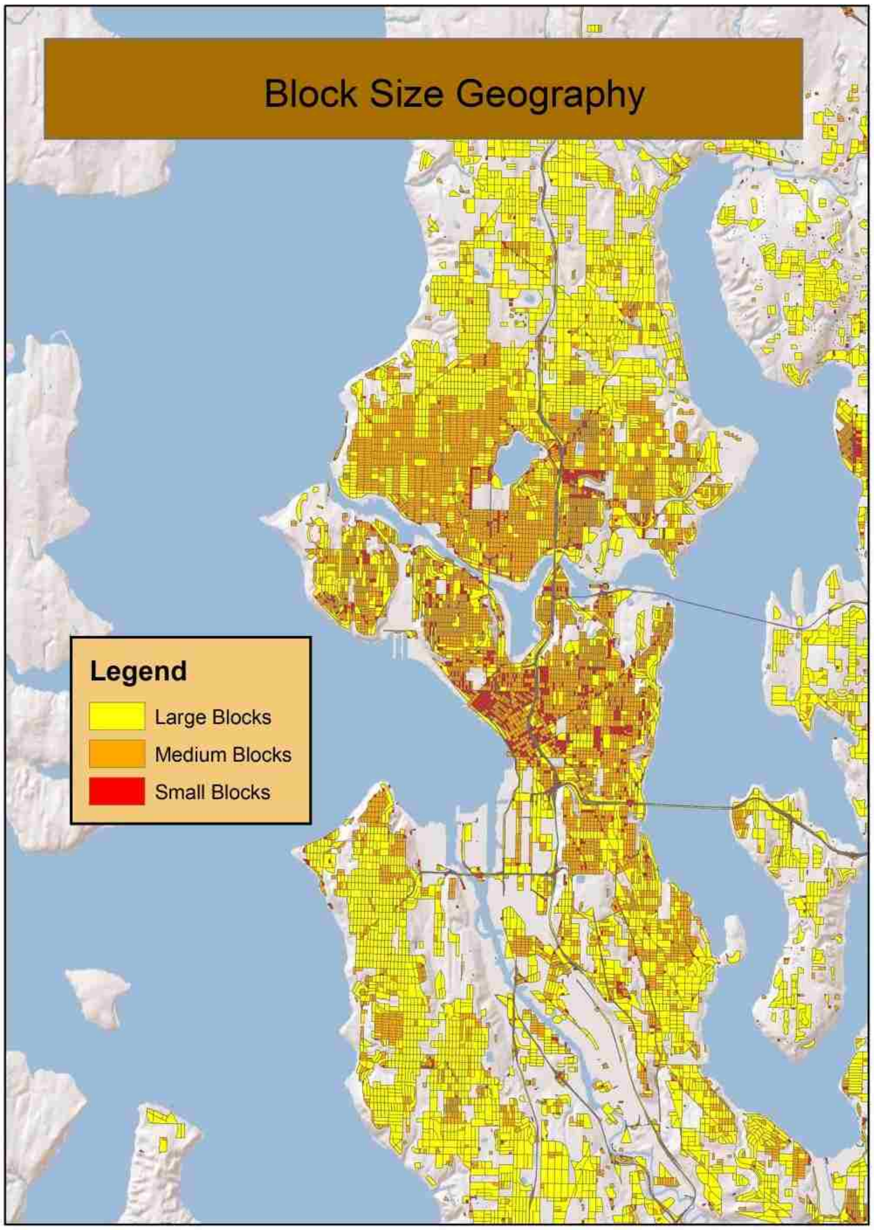

Figure 4-2: *Block Size Geography Map of Seattle*. **Note:** The majority of small blocks are located in the central business district (CBD) in downtown Seattle. Block sizes gradually increase as one moves away from the CBD.

From this map, one can see that the smallest blocks, corresponding to the most densely developed areas, are heavily concentrated in the South Lake Union/ downtown region of Seattle. As you move away from downtown, the block sizes tend to get larger, as is noted by a higher concentration of large blocks.

The established size thresholds for each of the block size categories were determined from the 33rd, 66th, and 98th percentile block areas. A distribution of block sizes is shown in Figure 4-3. The three blue shaded regions correspond to the three block size categories.

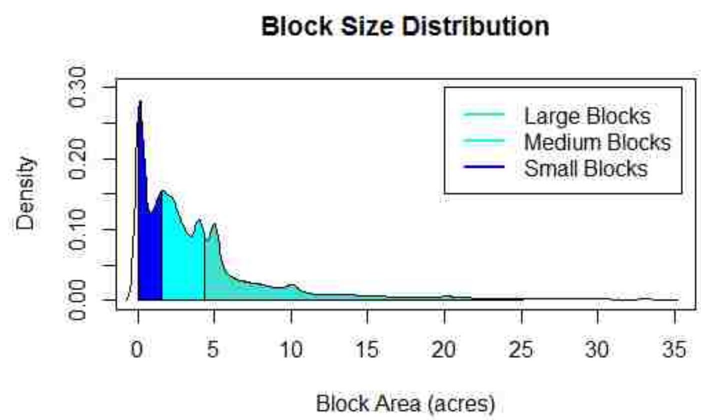

Figure 4-3: Block Size Distribution. Three shaded regions indicate the three block size categories

However, it is the block densities, not block sizes that are of interest. In this study, a surrogate indicator for the block density is defined as the number of blocks within a one mile radius of a trip origin/destination. The distribution of the block densities for the trip origins of trips less than five miles and within the Seattle city limits is shown Figure 4-4.

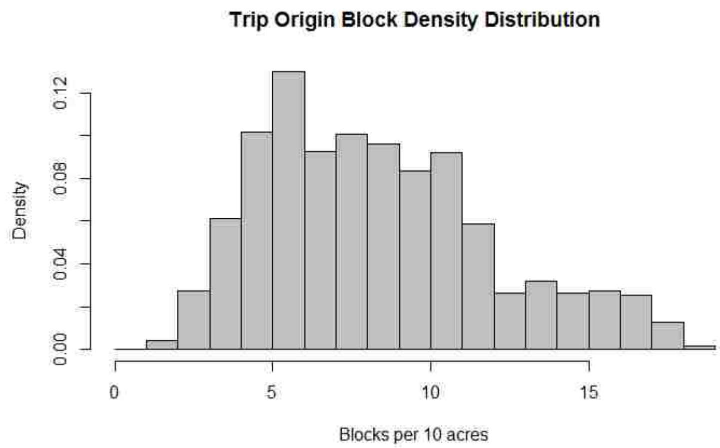

Figure 4-4: Distribution of block densities as computed from trip origins

This distribution is bimodal and indicates that densities tend to fit into two categories, those less than 1 block per acre and those greater than 1 block per acre.

Distance to Nearest Path

Distance to nearest path was used as a proximity measure of the distance between a given trip origin/destination, to the nearest bicycle trail/ path. The trails and path facilities included in this analysis are shown in map 1 above. This nearest distance metric provides a reasonable estimate of the bicycle/ pedestrian network topology and is illustrated as a 3-D contour surface shown in the map below (Figure 4-5).

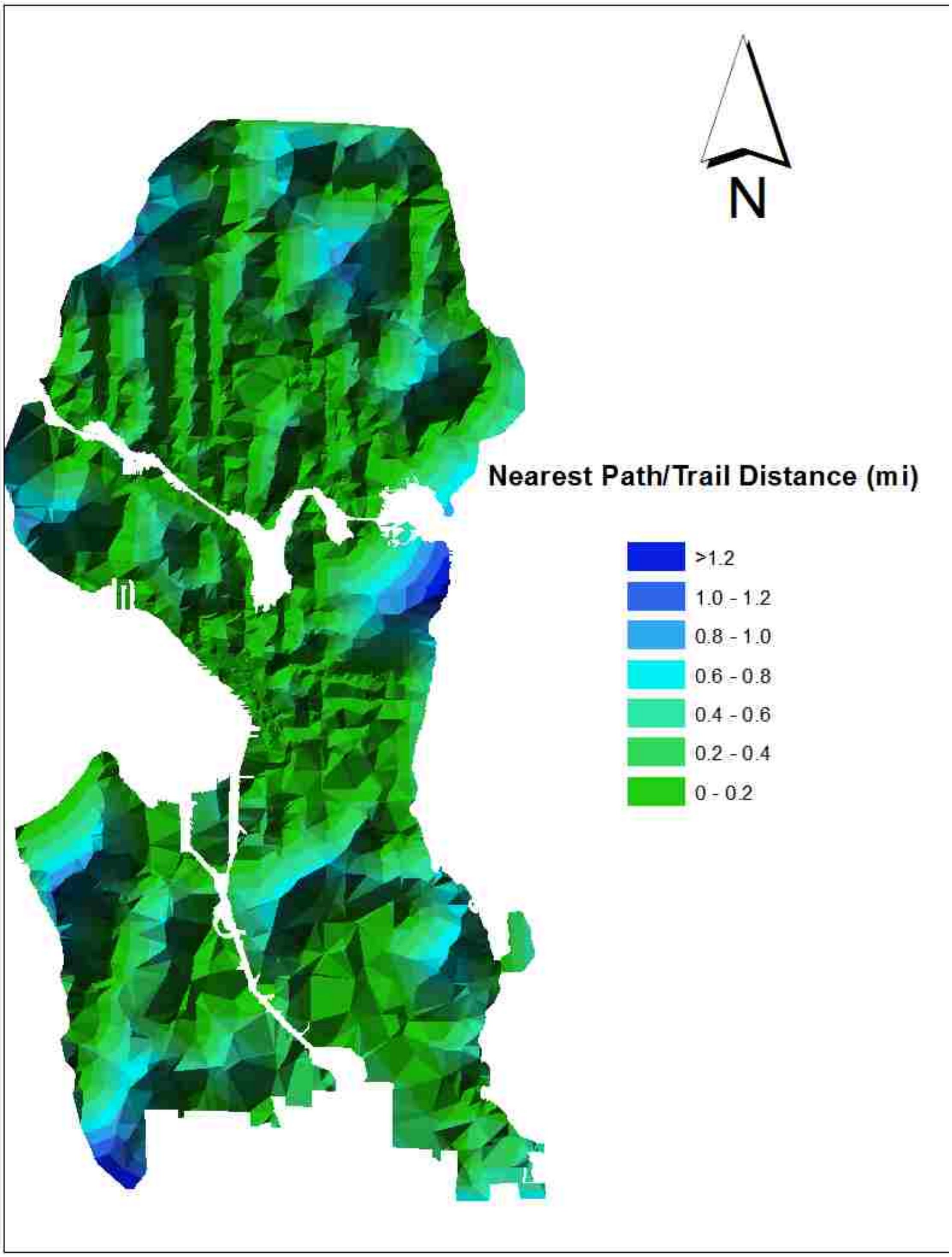

Figure 4-5: Non-Motorized Network Topology: This map illustrates the nearest distance from given trip origins/destinations to non-motorized paths/trails.

Regarding this map (shown above), regions with higher bicycle/pedestrian connectivity and accessibility are indicated by the darker green shaded areas. In contrast, the dark blue regions indicate areas with relatively low bicycle/ pedestrian connectivity and accessibility.

Cross Walk Density

Cross walk densities were calculated in a nearly equivalent fashion to block densities, except, instead of counting the number of blocks within a 1 mile radius of a trip origin/destination, the number of cross walks encompassed by this circle was quantified. It was thought that areas with higher levels of cross-walks would be more conducive to non-motorized road users. Cross walks provide safer places for pedestrians to navigate across streets, and used in tandem with pedestrian warning and yield signs, can improve the safety of pedestrians, not to mention, make them feel more comfortable while walking.

A map showing the densities of cross walks in Seattle is shown in Figure 4-6. Highest density areas are shaded red, while the lowest density regions are shaded purple.

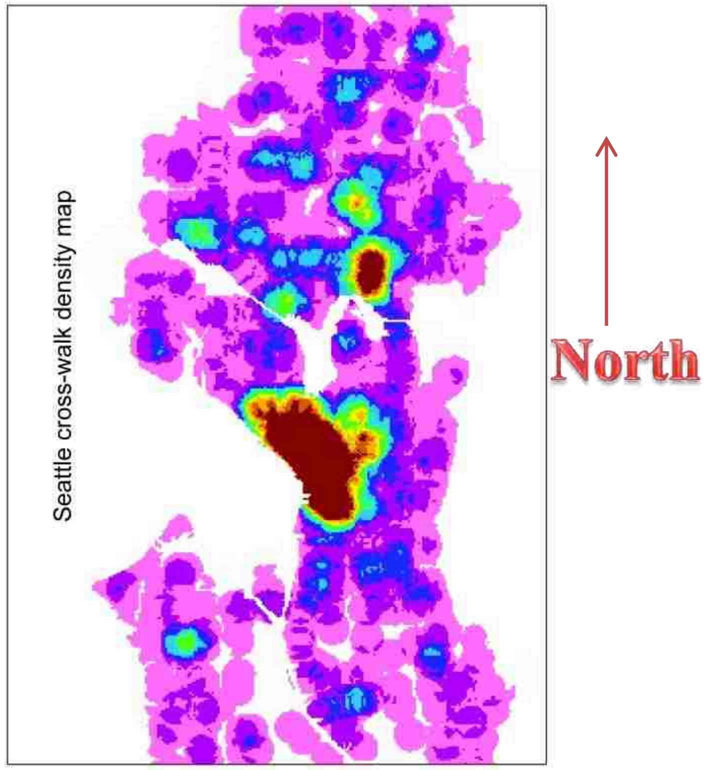

Figure 4-6: Spectral Map of Cross-Walk Densities in Seattle

An additional variable, *slope,* was generated by incorporating USGS elevation data in ArcGIS, and obtaining the average slope between origin and destination locations of each trip. The purpose of this variable was to account for the impedance of landscape topography. Darkness and weather, well known hindrances of non-motorized transportation were also accounted for with *Darkness*, a binary indicator of whether a trip was made during non-daylight hours, and Rain, a three category ordinal variable describing the quantity of daily precipitation coinciding with the date of a given trip. These indicators were based upon daily weather and sunrise/sunset data sources for the Puget Sound region that was obtained from the National Oceanic Atmospheric Administration’s databank.

Table 4-1 lists the built environment, demographic, and impedance variables that were considered in this study. Although analyzing the relationship between the built environment and non-motorized transportation was the primary objective of this paper, in modelling this relationship, it was also necessary to account for both demographic and natural impedance factors. It has been noted that failing to account for these other factors may lead to a biased conception of the built environment and non-motorized travel relationship.

Variable	Units	Type	Description
Impedance Factors			
Distance	*miles*	*continuous*	distance between O-D as recorded in survey
Slope	*ft/mile*	*continuous*	absolute value of difference in elevations between O and D divided by the O-D distance
Rain	----	*ordinal*	1= light rain , 2 = moderate rain, 3= heavy rain (based on quartile values from annual precipitation distribution)
Darkness	----	*binary*	equals 1 if after dusk / before dawn, else 0
Built Environment Factors			
PathDistO	*feet*	*continuous*	Euclidean distance between trip O/D and nearest shared pedestrian trail / bicycle path
PathDistD	*feet*	*continuous*	
XWalksCountO	*cross-walks*	*continuous*	Number of cross-walks within 1 mile radius of trip O/D
XWalksCountD	*cross-walks*	*continuous*	
BlocksD	*blocks*	*continuous*	Number of blocks witin 1 contained in 1 mi radius buffer of trip O/D
BlocksO	*blocks*	*continuous*	
Demographic Factors			
Sex	----	*binary*	Gender of trip taker (1 =male, 0=female)
Age	----	*ordinal*	Age category of trip taker (as indicated from survey question)
Ldrv	----	*binary*	Has valid driver's license (1 if yes, 0 if no)
Edu	----	*ordinal*	Highest level of education attained (as indicated from household survey)
hinccat2	----	*ordinal*	Household income level (from survey)
Hhnumveh	*vehicles*	*discrete*	number of vehicles owned per household

Table 4-1: Description and list of the impedance, built environment, and demographic factors used for analysis

4.3 Research Design

In quantifying the effect of the built environment on non-motorized forms of transportation, previous researchers developed mode choice models to predict the probability of an individual choosing mode i, given a set of demographic, impedance and built environment factors. Previous studies looked at trips holistically, irrespective of trip purpose. The trip purpose however may carry significant weight in a persons' mode choice. In addition, various explanatory factors such as block density, found significant to one trip purpose, maybe superfluous to another. Furthermore, various inherent characteristics specific to a given type of trip (shopping, work, leisure, etc.), warrant different requirements of the trip taker. For instance, shopping trips require carrying of goods; for work and social occasions, arriving well-groomed is often necessary; work trips require the need to arrive on-time differ compared to leisure/shopping trips which have more flexible timelines. Not accounting for trip purposes in this research area may lead to misinterpretations of non-motorized travel behavior.

In this study, trip data was divided by purpose in order to better understand this elusive relationship between the trip taker, and the built environment. The purpose of each trip was determined from the trip survey and allowed users to choose the most relevant trip purpose from a list of nine total trip purposes including*: work*, *school*, *home*, *social, shopping, eating, recreational, personal business,* and *other*. The trip purposes of *recreational* and *other* were excluded from further analysis due to their ambiguous nature, and to exclude exercise trips. Although a large portion of non-motorized trips are generated in the name of exercise, we were primarily interested in studying the interactions between motorized and non-motorized transportation modes and therefore chose to exclude these trips. The remaining 7 trip purposes were broken down into four main trip categories

-*work/school*

-*eating/social*

-*shopping*

-*personal business*

The trips included in a given trip category can be thought of as a distinct data subset. A breakdown of trips showing the portion of biking and walking trips within each trip category is shown in Table 4-2. It should be noted that the total number of trips, net trips, is 7977 and although the original number of trips selected from the original travel survey data was approximately double this 16,000. The reason for this is that half of the trips were randomly selected in order to train our model and the other half were reserved for model testing and validation. (Randomly splitting a dataset into two equal halves is a common statistical approach often used in machine learning applications and it allows for an unbiased validation of a given prediction model (Dill and Gliebe, 2008).) Overall biking and walking trips constitute to just over one quarter of all the trips analyzed. Percentagewise, biking was most often used for commuting purposes, constituting nearly 5% of the total work/school trips, while walking was most commonly used for eating/social trips, constituting nearly 31% of all eating/social trips.

Trip Purpose	Work/School	Shopping	Personal Business	Eating/Social	Net Trips
Number/ % of Bike Trips	127 / 4.7%	22 /1.3%	17 / 1.0%	20 /1.1%	186 /2.33%
Number/ % of Walking Trips	625 / 23.3%	374 /21.4%	330 /19.7%	586 /31.4%	1915 /24.0%
Total Number of Trips	2687	1746	1677	1867	7977

Table 4-2: Proportion of Biking and Walking Trips by Trip Category

In order to quantify this relationship, a binary mode choice model was used. This model predicts the probability of an individual selecting one particular mode choice among all other possible choices and is of the following form:

$$P(j)_i = \frac{e^{u_i}}{1 + e^{u_i}}$$

and

$$u_i = \beta_o + \boldsymbol{\beta} \cdot \boldsymbol{X_i}$$

where

$P(j)_i$= probability of individual *i* choosing mode *j*.

$\boldsymbol{X_i}$ **=** covariate matrix for individual *i,* including all relevant built environment, demographic, and impedance factors listed in table 1

$\boldsymbol{\beta} =$ matrix of covariate coefficients as estimated using method of maximum likelihood

In this case, we were interested in both the propensities of individuals selecting either walking or biking travel modes and therefore, two different forms of binary mode choice models were developed: one for bicyclists, and another for walkers. These two types of models were fit for each of the trip purpose data subsets to predict, among a set of all possible mode choices, the probability of biking or walking respectively. The result of this model fitting process was a total of 8 trip purpose specific choice models.

However, before each of these 8 models were fit, lasso regression was used to determine with of the 16 variables listed in Table 1 were relevant, and which should be excluded from the model fitting. A popular class of penalized regression, lasso regression has been noted to provide both improved model prediction power, while also acting as an effective variable selection method (Dill and Gliebe, 2008). Lasso regression was implemented using the two step procedure described below.

1) Perform lasso regression using 10-fold cross validation to determine the optimum value of the Lasso penalty term, *Lambda*.
2) Using the previously determined optimum *lambda* penalty term, refit the model on the entire dataset to yield the best fit model containing the optimum set of selected set of variables

Although Lasso regression is a powerful technique used for feature selection with the goal of minimizing the overall test error of a model, it has some notable drawbacks, one of which being, that it is not able to provide the conventional variable significance measures of un-penalized regression methods, and also, because the variables must be standardized prior to the analysis, the model can be difficult to interpret, especially when ordinal and binary variables are present in the model. Because of this, lasso regression was only used for feature selection purposes, after which, the dataset was refit using the best set of selected variables using normal, non-penalized regression.

4.4 Estimation Results and Discussion

The results of each of the mode specific binary bike and walk models are illustrated in Table 4-3 and 4-4. The coefficients of each covariate, as well as their statistical significance, are illustrated in these tables. The presence of a shaded row for a given variable indicates that this feature was not selected by the lasso regression and is therefore deemed to be irrelevant for explaining non-motorized trip propensity. For brevity's sake, the discussion of these results has been summarized to include only the most important key findings.

	Work/School Trips				Shopping Trips			
Variable	Coefficient	Standard error	Probability		Coefficient	Standard error	Probability	
(Intercept)	5.97E-01	1.05E-01	1.64E-08	***	4.70E-01	5.66E-02	< 2e-16	***
distance	-1.51E-01	7.46E-03	< 2e-16	***	-1.19E-01	6.99E-03	< 2e-16	***
Slope	1.83E-05	1.38E-04	8.94E-01		-8.37E-04	1.47E-04	1.60E-08	***
Rain	6.51E-03	2.68E-02	8.08E-01					
Darkness					7.94E-02	5.19E-02	1.27E-01	
PathDistD	-2.28E-05	9.93E-06	2.20E-02	*				
PathDistO	-2.13E-05	8.83E-06	1.59E-02	*	-1.18E-06	7.57E-06	8.76E-01	
XWalkscountD	6.03E-05	4.75E-05	2.05E-01		2.81E-04	4.32E-05	1.11E-10	***
XWalkscountO	1.45E-04	3.23E-05	7.83E-06	***				
BlocksD	3.20E-05	7.93E-05	6.87E-01		-2.99E-05	7.41E-05	6.87E-01	
BlocksO					2.16E-04	5.27E-05	4.37E-05	***
agerng	-2.34E-03	7.19E-03	7.45E-01		-1.30E-02	4.30E-03	2.50E-03	**
ldrv	1.05E-02	3.94E-02	7.89E-01					
edu	-3.50E-03	3.18E-03	2.70E-01		-4.36E-04	1.95E-03	8.23E-01	
sex	1.92E-02	1.84E-02	2.97E-01		2.79E-02	1.70E-02	9.96E-02	.
hinccat2	-3.34E-03	3.41E-03	3.27E-01		-1.37E-03	2.91E-03	6.39E-01	
hhnumveh	-2.65E-02	1.03E-02	1.04E-02	*	-5.54E-02	9.81E-03	1.94E-08	***
	Personal Business Trips				Social/Eating Trips			
Variable	Coefficient	Standard error	Probability		Coefficient	Standard error	Probability	
(Intercept)	4.88E-01	1.19E-01	4.59E-05	***	7.94E-01	8.14E-02	< 2e-16	***
distance	-1.21E-01	9.82E-03	< 2e-16	***	-1.93E-01	7.44E-03	< 2e-16	***
Slope	-4.04E-04	1.82E-04	2.72E-02	*	-1.03E-03	1.25E-04	3.15E-16	***
Rain	2.97E-02	3.72E-02	4.25E-01		2.78E-02	2.35E-02	2.36E-01	
Darkness	7.45E-02	1.41E-01	5.97E-01		-1.34E-02	5.08E-02	7.92E-01	
PathDistD					-3.56E-06	7.95E-06	6.55E-01	
PathDistO	-1.18E-05	1.08E-05	2.77E-01		-2.19E-05	8.02E-06	6.39E-03	**
XWalkscountD	2.13E-04	6.23E-05	6.65E-04	***	3.14E-04	4.00E-05	7.18E-15	***
XWalkscountO								
BlocksD	-1.66E-04	1.08E-04	1.26E-01		-1.30E-04	7.39E-05	7.96E-02	.
BlocksO	1.99E-04	7.25E-05	6.05E-03	**	6.59E-05	5.26E-05	2.10E-01	
agerng	-1.62E-02	6.65E-03	1.53E-02	*	-1.76E-02	4.89E-03	3.34E-04	***
ldrv	1.29E-02	4.28E-02	7.63E-01		-2.67E-02	2.96E-02	3.67E-01	
edu	-1.10E-03	2.45E-03	6.54E-01		1.02E-03	1.40E-03	4.66E-01	
sex	-8.32E-03	2.44E-02	7.34E-01		3.27E-02	1.68E-02	5.16E-02	.
hinccat2					1.43E-03	1.20E-03	2.35E-01	
hhnumveh	-3.28E-02	1.34E-02	1.47E-02	*	-3.84E-02	9.04E-03	2.28E-05	***

Table 4-3: Trip purpose specific binary logit model results for walking.

To begin this discussion, let us first focus on purpose specific walking models, the results of which are illustrated above (Table 3). Among the impedance factors, distance was found to be highly statistically significant and negative in sign. Rain and darkness were found to be insignificant for all four trip purposes and slope was found to be highly significant for shopping and social/ eating trips. Demographic factors did not play a significant role in the binary walking models; however, age was statistically significant for shopping and social/eating trips. The negative sign of the variable age indicates that older individuals have a lower propensity to walk for these types of trips.

Several built environment factors were found to be significant within several of the purpose specific walking models. *BlockO,* a variable indicating the density of blocks within a 1 mile buffer region of a trip origin, was found to be highly significant for all trip purposes save work/school trips, where it was determined to be irrelevant by the lasso feature selection procedure. *XwalkO,* indicating the cross walk density near a trip origin, was found to be highly significant for work/school trips, but was otherwise found to be irrelevant for the other trip purpose models. The positive sign suggests that higher cross walk densities are more conducive to walking trips. *XWalkD,* was found to be statistically significant for shopping, personal business, and social/eating trips and like *XWalkO,* was also found to be positive in sign. The marginal statistical significance and negative sign of the *PathDistO* coefficient for work/school and social/eating trips, indicates that closer proximity of trip origins to pedestrian /bicycle pathways tend support more walking trips and increase the likelihood of an individual choosing to walk.

Regarding bicycle mode choice models, the purpose specific binary bicycle choice models are illustrated in Table 4. It should be noted that the models for social/eating and personal business trips has been excluded from this paper because all of the variables were deemed irrelevant by the lasso regression variable selection. This means that of the candidate variables, none pose any predictive power in bicycle mode selection. This effect is definitely due to a lack of bicycle trips for these trip purposes. (Personal business and eating/social trips by bicycle constituted only 17 and 22 of total trips respectively).

Comparing the work/school trips models between biking and walking modes, demographic factors in the biking model were found to be overall much more statistically significant. Both income and gender were found to be highly significant in explaining the propensity to bicycle. *PathDistD, XWalkCountO and BlocksD/O* were all found to be very significant. The negative sign of the *PathDistD* coefficient is disconcerting and suggests that closer proximities to pedestrian/bike trails are associated with a decrease in bicycle activity. This may indicate there's disconnect between the locations of paths and the work/schools of individuals who commute by bicycle, however more research is necessary to confirm this theory. The negative sign on the coefficient for *XWalkcountO* suggests that bicyclists have an affinity towards areas with lower densities of cross-walks. Crosswalks often are located at busy intersections that often hinder bicyclists by forcing them to stop. Previous research has determined that cities with higher stop frequencies for bicyclists are likely to decrease bicycle demand (James et al., 2013), thus this finding is supported.

Work/School Trips					Shopping Trips			
Variable	Coefficient	Standard error	Probability		Coefficient	Standard error	Probability	
(Intercept)	1.45E-01	4.66E-02	1.83E-03	**	4.27E-02	4.26E-02	3.17E-01	
distance	3.32E-03	3.21E-03	3.02E-01		-3.39E-03	3.63E-03	3.50E-01	
Slope					4.65E-05	7.46E-05	5.33E-01	
Rain	2.97E-03	1.21E-02	8.06E-01		-5.55E-03	1.00E-02	5.79E-01	
Darkness	-3.70E-02	7.91E-02	6.40E-01		-2.34E-02	2.92E-02	4.23E-01	
PathDistD	-1.25E-05	4.46E-06	5.29E-03	**	-8.61E-06	3.96E-06	2.98E-02	*
PathDistO	5.34E-06	3.98E-06	1.79E-01					
XWalkscountD	4.22E-05	2.21E-05	5.61E-02	.	-5.55E-05	2.20E-05	1.17E-02	*
XWalkscountO	-1.13E-04	2.20E-05	3.31E-07	***				
BlocksD	-1.05E-04	3.91E-05	7.41E-03	**	3.93E-05	4.01E-05	3.27E-01	
BlocksO	1.32E-04	3.71E-05	3.93E-04	***	5.61E-06	2.75E-05	8.38E-01	
agerng	-5.94E-03	3.07E-03	5.29E-02	.	-5.46E-04	2.48E-03	8.26E-01	
ldrv	-7.25E-02	1.70E-02	2.05E-05	***	-1.76E-02	1.71E-02	3.02E-01	
edu	1.40E-03	1.39E-03	3.13E-01		7.12E-04	9.03E-04	4.30E-01	
sex	3.83E-02	8.13E-03	2.67E-06	***	5.19E-03	8.73E-03	5.52E-01	
hinccat2	3.92E-03	1.47E-03	7.74E-03	**	1.12E-03	1.50E-03	4.55E-01	
hhnumveh	-1.42E-02	4.48E-03	1.50E-03	**	-2.32E-03	5.15E-03	6.52E-01	

Table 4-4: Results of purpose specific binary logit models for biking. (Note: Eating/social trips excluded for a complete lack of selected features.)

Contrasting between biking and walking modes, cross-walk densities were found to be inversely related to bicycle trips as compared to the strong positive relationship that was discovered between walking cross walk densities. This difference highlights upon differences between bicycle promoting policies and policies aimed at providing more walkable communities. It emphasizes the need to provide different considerations when planning for these different NMT modes.

Overall proximity measures to pedestrian/ bicycle paths were not found to of significance for walking trips, and were only marginally significant for bicycle work/school trips. Higher density areas, represented by block density were mainly found to contribute to non-motorized activity; however, among work/school biking trips, a strong negative relationship between destination block density and the trip propensity suggests the opposite may be true.

Differences in the relationships between NMT and different trip purposes support that trip takers weigh their decisions to choose non-motorized transportation at least partly upon the purpose of the trip. Also, demographic variables like sex, and income were found to be both relevant and significant for bicycle work/school trips, however were much less significant for walking trips.

From the results presented, a combination of both built environment, impendence, and demographic factors were shown to be significant in explaining non-motorized transportation mode choices. Previous literature has noted that this type disaggregate study suffers if self-selection is not accounted for. Self-selection is essentially the idea that individuals select their living locations and the places they travel based on their individual taste preferences for particular modes of transportation. Unfortunately, it is not possible to say that the results of our study and the strong associations that were established between several built environment factors and NMT are not biased on account of this phenomenon. Without confirmation of this association, it is not possible to confirm that pro-bicycle and walking transportation policies will be effective for promoting a more diverse array of transportation options. Future studies could address this issue through the analysis of before and after treatment studies, or by administering surveys to individuals assessing the effects of hypothetical transportation policies aimed at improving the built environment.

3.5 Conclusion:

The large disparity between percentages of non-motorized trips taken in the US compared to other prominent first world nations is evidence that improvements within our country's transportation network are necessary. Transportation policies often aim at improving areas to be safer and more convenient for non-motorized forms of transportation through alterations to the built environment. A few conventional examples of such policies include installation new non-motorized friendly infrastructure such as new bicycle/ pedestrian pathways, suburban land densification, as well as transit oriented development. However, the evidence that that changes in built environment factors implemented by new transportation policies will help decrease this NMT disparity are limited. In analyzing the effect of the built environment on NMT for individuals living in Seattle, while controlling for demographic factors and the natural hindrances of weather and darkness, this study helps to add to the limited knowledge of this illusive relationship, although our understanding of this subject is still very opaque. More studies are necessary if the true nature of this relationship is to be understood. The results of this study suggest that there are significant differences between the factors influencing bicycle propensity vs. walking propensity, and also, that significance of this relationship changes based upon the purpose of the trip.

Chapter 5: Conclusions and Future Work

Congestion, pollution, and an obesity epidemic are major issues in the U.S. right now. Our lust for the convenience of the personal automobile is only worsening these issues. Galvanized by the problems facing our nation's current non-sustainable transportation system, government stakeholders are now beginning to consider non-motorized forms of transportation as viable and healthy alternatives to the automobile. However, in order for recent budget increases in non-motorized transportation to help get more Americans out of the driver's seat, a more transparent understanding of pedestrian and bicycle travel behavior is necessary. The work presented here seeks to contribute to our limited and opaque knowledge of this under-studies mode group.

The body of work that is the backbone of this thesis is routed in two studies. In the first study, the relationship between weather and bicycle ridership was investigated, for a predominantly recreational population of bicyclists in Albuquerque, NM. This study, like similar studies on this topic, confirmed that weather is a significant factor in quantifying and predicting bicycle volumes. In spite of the relatively warm climate of the Northern-Chihuahuan dessert, in which the city of Albuquerque, NM is located, bicyclists are still very sensitive to even slight changes in ambient temperatures. A fact that can be confirmed statistically in the least squares regression model results shown in Table 3-1, as well as visually in the "Temperature vs. Bicycle Volume" graph illustrated in Figure 3-4. Although previous research on the relationship between bicycle demand and weather is far from scarce, confirming the significance of weather for an arid climate, like that of Albuquerque is valuable, and may also transfer to other arid cities throughout the South Western United States. With the exception of Lewin (Lewin, 2011), no other known research has studied the weather's effect on bicyclists in an arid climate. This research provides an additional study of weather's effect on bicyclists in arid climates. Future research might include a variable for wind speed in the linear regression model as this was previously found to be significant (Mirada-Moreno and Nosal, 2011). Also, this study confirmed that daily temperature variations influence the times which bicyclists choose to ride. Additional research should be done to see how weather influences different types of bicyclists, in particular, to determine if recreational cyclists and commuter cyclists are affected by weather differently.

The main contribution of this work however lies in the shifting hourly bicycle peak vs. temperature relationship which was discovered by applying a novel deconvolution method to hourly distributions of bicycle demand for separate months of the year. This analysis confirms the hypothesis that bicyclists do indeed shift the time of day that they ride throughout the course of the year, to those times of day with the most favorable temperatures. This seemingly-intuitive discovery provides not only insight into the relationship between mother nature and bicycle riders, but is useful for transportation planners and engineers everywhere as they seek to design more livable communities that include well-designed facilities for bicyclists and motorists alike.

The second study presented, documents the investigation of the relationship between the built environment and non-motorized travel choices. This study utilized disaggregate-level trip survey data from the 2006 Puget Sound Regional Council Household Travel Survey, as well as rich ArcGIS spatial database. Four purpose specific binary mode choice models were developed for both biking and walking transportation modes. The results show that the built environment is indeed closely tied to travel mode choice and that the magnitude, direction, and significance of this relationship are dependent upon the purpose of each trip. A major limitation of this study was the lack of overall bicycle trips. This made the modelling process difficult and resulted in a lack of statistically significant factors. Having access to a larger dataset would have likely yielded different and more accurate results.

Appendix I.

Detailed steps for data compilation, processing and analysis:

This section describes in fine detail, the steps and processes necessary to duplicate the research method and analysis completed to quantify the relationship between non-motorized travel mode choice and the built environment. It is structured as follows. 1) A detailed description of each dataset is given. 2) Data compilation procedures conducted in Microsoft SQL server and ARCGIS are described.

Part 1: The Data sets

Before delving into the details about each data set used, as well as the processes and vivid details of this analysis, it would first be helpful to refresh the reader as to the purpose and what is to be the end result of this analysis.

As has been done in a modest quantity of published studies, the purpose of this research is to quantify and better understand the significance between the built environment and non-motorized travel behavior. In undertaking this endeavor, we hope to better understand some of the hindrances and non-motorized transportation, and the reasons behind its unsuccessful popularization in the U.S.

In quantifying this relationship developed discrete mode choice models using a combination of both trip and travel data information, and also data on the built environment. However, it has been shown that a person's mode choice may also be heavily influenced by climate, weather, individual demographics of the trip taker, not to mention, personal attitudes, and perceptions of safety and convenience. Therefore, these factors were also included in the model framework, shown below (**Figure A-1**).

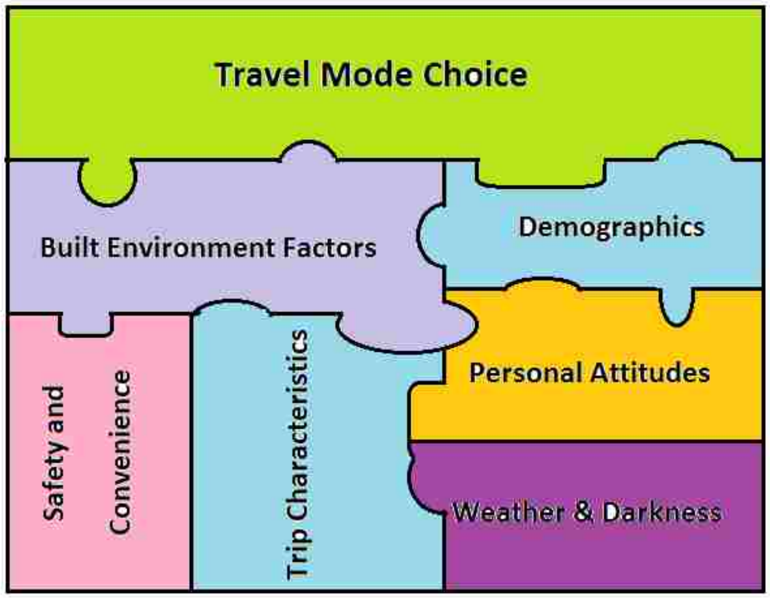

Figure A-1: Model Framework of Travel Mode Choice

Like any other research study, or project, we were subject to constraints and in our analysis, we are limited to the data which is available to us, therefore several attributes and factors listed in the model frame work were overlooked.

Household Travel Surveys

Household travel surveys are commonly administered in metropolitan areas for the purpose of gaging the travel behaviors and general demographic information of individuals in an area. They are also integral in the formation of the long range and short range transportation plans. In Washington's Puget Sound Region, The Puget Sound Regional Travel Survey is a survey administered every 8 years. For this study the 2006 PSRC travel survey was used. (At the time of this study, the 2014 PSRC travel survey data had not yet become available.) The survey data consisted of 3 main parts, household information, and personal information, and personal trip diary information for each household member.

Relevant data collected in the household information part of the survey included:

-*Household location*

-*Household size*

-*Household income*

-*Number of Household vehicles*

Important data collected in the personal information part of the survey included:

-*Age*

-*Sex*

-*Education level*

-*Employment status*

-*Driving license status*

Useful trip related data collected in each 48-hour travel survey, and administered to each household member included:

-*Trip destination location*

-*Trip origin location*

-*Trip purpose*

-*Mode of transportation*

-*Length of trip (distance)*

-*Length of trip (time)*

-*Time of day for each trip*

Acquiring the household data was step 1. The next step would be to acquire the built environment data, which will be described in detail in the next section. The last step would be to combine these two data sets together into one comprehensive disaggregate data set from which travel behavior on personal mode choice could be analyzed, and the link between the built environment and travel mode choice could be better understood.

Built Environment GIS Data Set

As mentioned previously, acquiring detailed geospatial information pertaining to the land use and urban landscape characteristics of the Puget Sound Region was a crucial part of this study. Fortunately, the acquisition of this data was not incredibly difficult. A detailed geospatial database was compiled using data accessed online through the Washington State Geospatial Data Archive (WAGDA) website.

After browsing this immense data archive, the following geospatial data sets were acquired:

-Roadway network

-Bicycle route network

-Sidewalks

-Pedestrian Crosswalks

A description of each of these geospatial data sets follows:

Roadway Network:

This comprehensive shape file contains all of the roadway centerlines in King County.

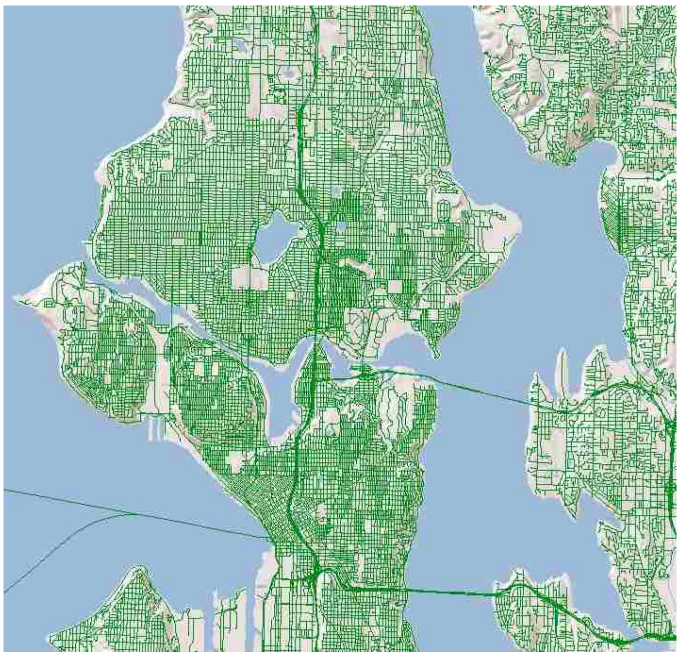

Figure A-2: Map Illustration of Roadway Network Data (*.shp*) in Seattle.

Bicycle Route Network:

A shape file containing all the trails and roads with bike facilities extracted from PSRC's transportation geodatabase. Each separate polyline segment represents a path or trail and contains following attributes: name of bicycle facility, type of bicycle facility (shared path, bicycle lane, etc.) for each side of the road.

.


Figure A-3: Map of Bicycle Network in Seattle

Sidewalks

A complete shape file data set containing the entire sidewalk network in Seattle

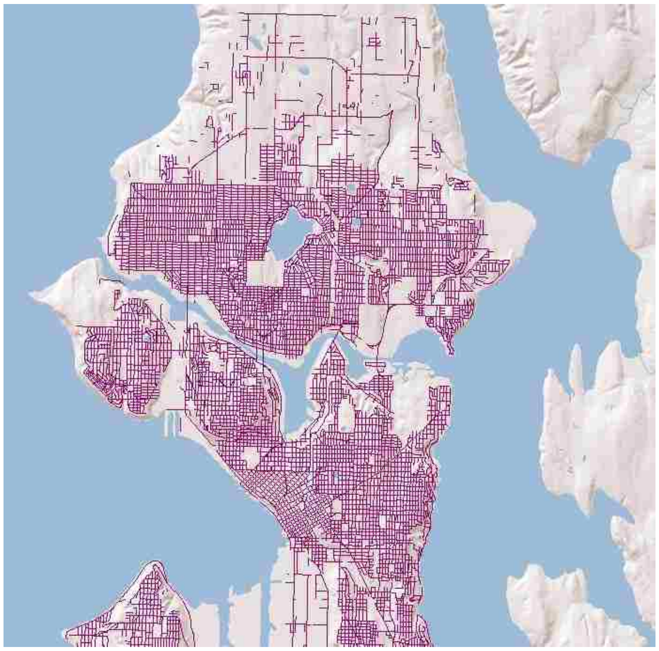

Figure A-4: Map of the Side-Walk Network in Seattle

Pedestrian Cross-walks

A complete shape file data set of all the cross-walks in Seattle. **Note:** This data, although relevant was not included in the modelling process due to a lack of time. Future research may include sidewalk connectivity/topology metrics as potentially influence factors in non-motorized mode choice behavior.

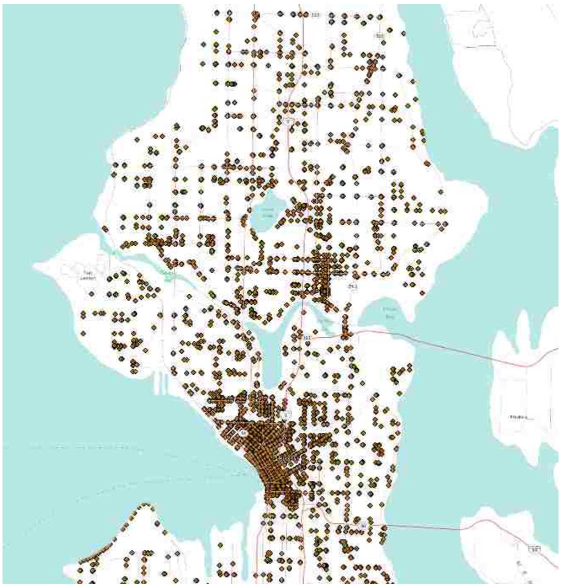

Figure A-5: Map of Pedestrian Cross-Walks in Seattle

Other Datasets

Three other data sources that were used were, weather data, sunset/sunrise data, and elevation data. Both the weather and sunrise/sunset data sources were obtained through the National Oceanic and Atmospheric Administration. These data represent a relevant piece of the modelling framework identified earlier.

The Weather data was queried from a weather station located in the heart of Seattle. The weather station made daily observation of several parameters; however for this study we only were interested in the presence of rainfall. Therefore, a simplified data dataset with the *Date* and a binary Rain indicator was created.

Sunrise/sunset was obtained for the Seattle area, and contained the sunrise time and the sunset time for every day of the year.

The Elevation dataset was obtained from the USGS data archive, in the form of a raster data file.

Part 2: Data Compilation Procedures

After obtaining all the necessary data sets, all that was left to do was combine and link these sources of data into one composite disaggregate trip data set. Together, our multiple datasets formed what will henceforth be referred to as our "study database". Essentially, the "study database consisted of the 4 key data sources:

(1) Weather data, (2) Sunrise/set data, (3) PSRC travel survey data, and (4) Built Environment data.

However, before these data sources could be amalgamated, several spatial attributes were first computed. These spatial attributes were computed using several built-in ArcGIS functions. A detailed description of the spatial computation follows.

From the built environment dataset, which was obtained through WAGDA, several metrics were created which reflect important land use and planning factors. As a surrogate metric for land use density, block densities were computed for each trip origin and destination. Block density was calculated as the number of blocks contained within a circle with a radius of 1 mile from the location of a trip origin/destination. A block in this instance was defined as the space contained in an area completely surrounded by streets segments. For this computation, a "block" shape file, containing the spatial information about each block, was created using the roadway network shape file. The Analysis Tool Box function "Feature to Polygon" was used in order to accomplish this task.

A filter was put on the block feature data set to eliminate excessively large block sizes, which were likely to be public spaces such as parks, as well as industrial areas. The threshold cut off for this value was based on the 98th percentile block size of the distribution of block sizes. After cleaning the block dataset to exclude excessively large blocks, circular buffer regions were created around each trip O/D of 1 mile in radius using the built-in "Buffer" analysis tool located in the ArcGIS toolbox. Then, using a spatial join, between the buffer region and block features was created. The number of blocks contained in each buffer zone was computed. This count value represents the block density at each O/D.

As a surrogate measure for the network topology and the accessibility of bicycle and pedestrian facilities, the near distance metric was computed. This metric was calculated as the shortest distance from a given trip O/D to the nearest bicycle/ pedestrian facility. In order to calculate this distance, "near dist" a spatial attribute representing the closest distance from a given O/D to a bicycle/ pedestrian path was added to the "trip location" feature data set. This "near dist" attribute was computed using the "Near" analysis function of the ArcGIS network toolbox. It was thought that trips occurring in areas with shorter "near distances" would be more conducive to non-motorized activities.

The third built environment metric that was computed was the density of pedestrian cross-walks in the vicinity of a given origin/destination. This density metric was computed in the same manner as the block densities. Using a spatial join on the cross-walks feature data set and the buffer region data set, the quantity of cross-walks in each buffer zone was computed. The spectral map, shown below, represents the density of x-walks in Seattle. Red colors indicate very dense cross walk areas, whereas blue and purple areas indicate low cross-walk density areas.

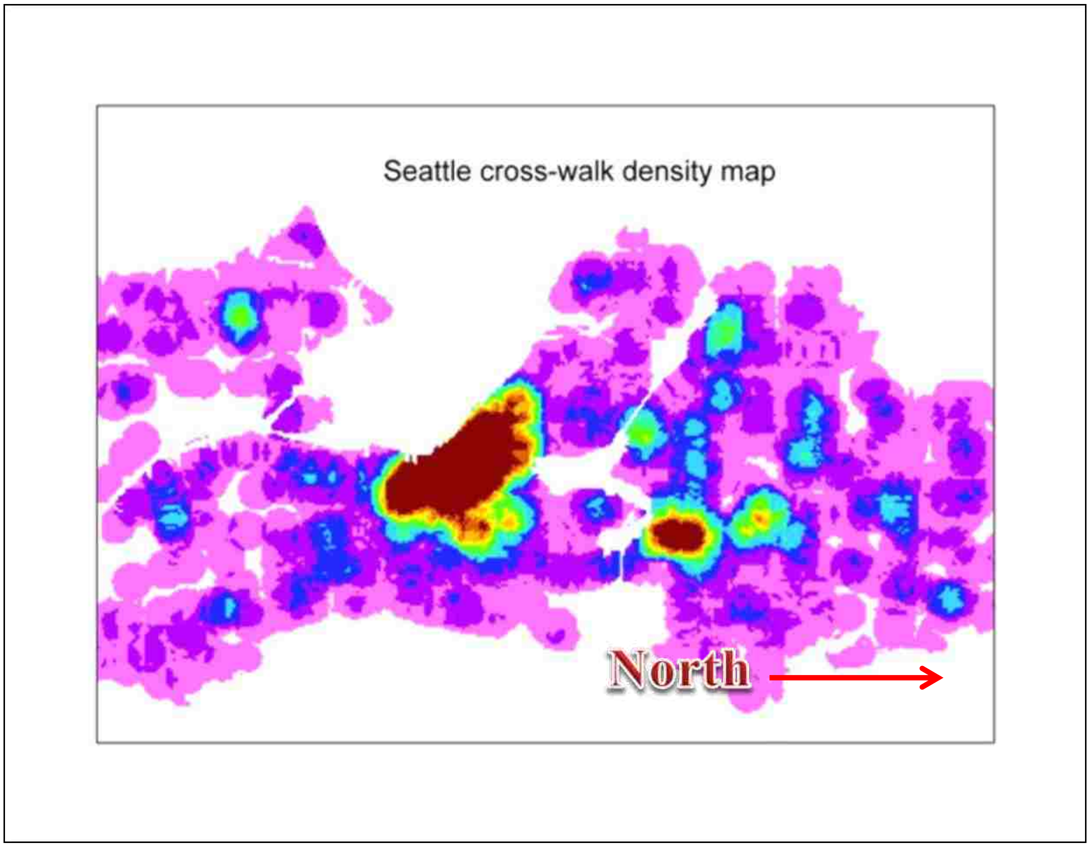

Figure A-6: Spectral Density Map of Seattle Cross-Walks

Incorporating the elevation raster data file, the "hilliness" of each trip could be computed. In this study, the "hilliness" of a given trip was computed as the average slope between each trip origin and destination. This was computed as the elevation difference between origin, i, and destination j, divided by trip distance as reported in the survey. The average slope formula is shown below (equation (1)).

(1) $$Slope\ Average_{i-j} = \frac{|elevation_i - elevation_j|}{d_{i-j}}$$

where,

$*\ elevation_i\ and\ elevation_j\ represent\ the\ elevations\ in\ feet\ of\ trip\ origin\ i, and\ destination\ j$

$*\ d_{i-j} =\ distance\ in\ miles\ between\ i\ and\ j\ as\ reported\ in\ survey$

This "Slope average" was computed for each trip origin and destination in the trip dataset.

This process

The ArcGIS data computations have now been described in full. Altogether, 4 attributes were computed for each trip origin and destination by spatially joining the trip location information to the built environment dataset. The net result of this is one large "spatial data" table. The first 5 rows of this data table are shown below (**Table A-1**).

recid	qno	pernum	tripnum	Distance (mi)	BlocksD	BlocksO	PathDistD (ft)	PathDistO (ft)	XWalksCountD	XWalksCountO	Slope (ft/mi)
23366	14613	1	4	2.81	322	230	2306.1	3220.6	171	100	5.1
23390	14613	2	5	3.2	201	259	3123.3	1820.2	36	17	5.1
23402	14613	3	7	1.1	159	156	549.0	441.8	67	43	3.4
23413	14613	4	9	1.1	159	156	549.0	441.8	67	43	3.4
23414	14613	4	10	1.1	156	159	441.8	549.0	43	67	3.4

Table A-1: *Compiled Spatial Data Table:* Trip specific attributes as reported in survey are shown in blue. Computed built environment metrics and "slope" shown in red.

The last portion of the data compilation involved linking the weather data table and the personal and household information collected in the trip survey, to the previously-generated spatial trip data table illustrated above. This linking and data amalgamation process is described in the following section.

Data Amalgamation

In order to arrive at our end result, a comprehensive disaggregate trip data file, the following 5 different data tables were joined together: *spatial data table*, *personal characteristics table*, *household characteristics table*, *weather data table*, and the *trip characteristics table.* The schema diagram illustrating the relationships between these five tables is illustrated below

(**Figure A-7**). The attributes are also listed for each table with attribute keys underlined and in bold font. The *personal characteristics table* has 2 keys, *qno*, to denote the unique household an individual is a member of, and *pernum*, to distinguish a unique individual from a given household. All of the other tables have only 1 key which is unique. *Recid,* the key of both the *spatial data table* and the *trip characteristics* table is a unique number that was assigned to each individual trip.

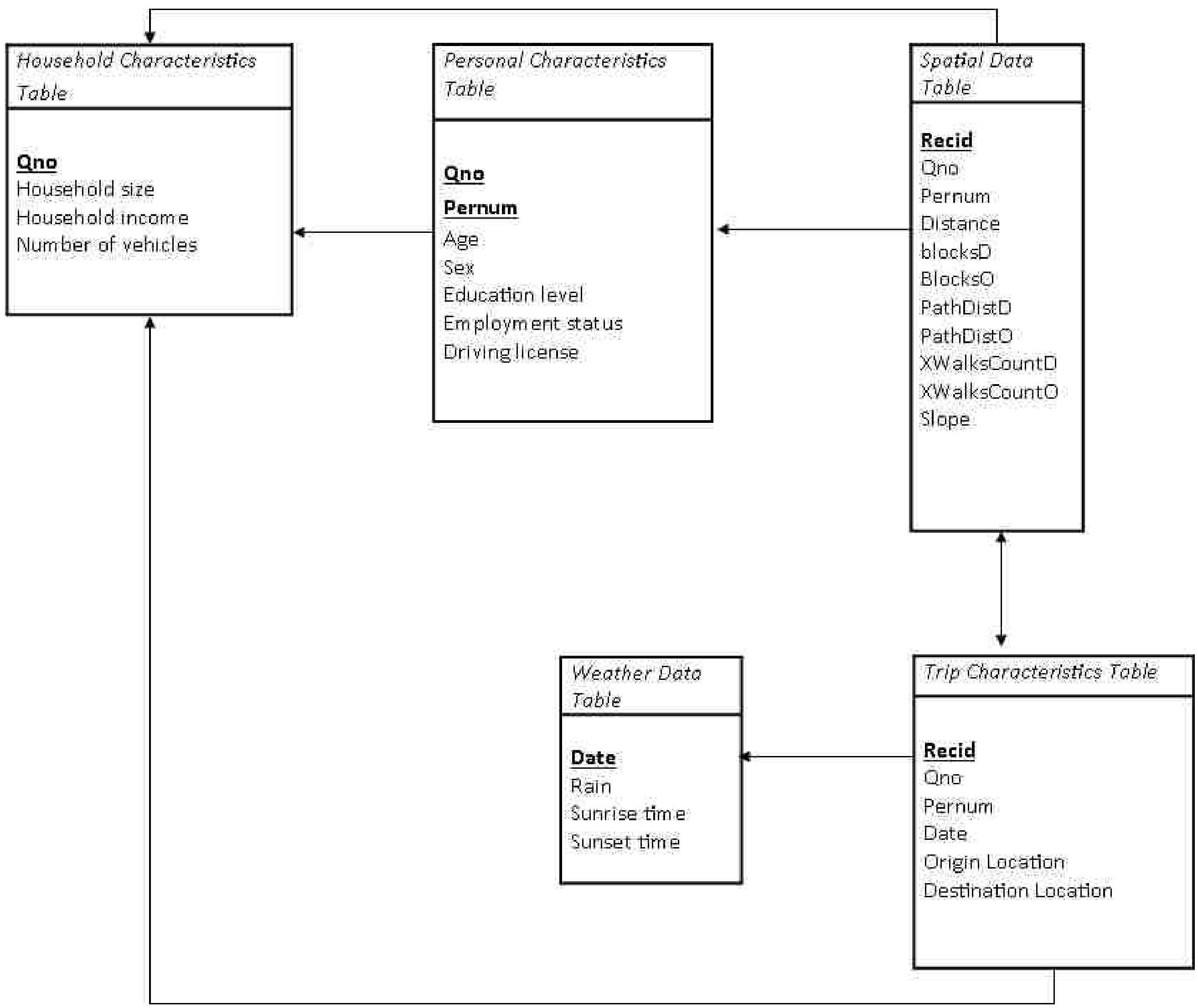

Figure A-7: Data Relationship Diagram

A total of 6 relationships exist between the five tables and the nature of each relationship is denoted by the configuration of the arrow connection between them. A one-to-many relationship exists from the *characteristics table* to the *personal characteristics table,* indicating that a given household can have multiple occupants. In addition, two more one-to-many relationships exist from the *household characteristics table* to the *spatial data table and* the *trip characteristics table,* which indicate that each trip is taken by a person who can only live at one household. A one-to-many relationship also exists from the *personal characteristics table* to the *spatial data table*, indicating that each individual may take multiple trips. The one-to-one relationship exists from the *trip characteristics table* to the *spatial data table.* Lastly, a one-to-many relationship exists from the *weather data table* to the *trip characteristics table,* indicating that each trip may have only one weather event and a single sunrise and set time.

Knowing these relationships between the different tables was crucial to assembling the final dataset. In order to comprise the final dataset, a database was created using Microsoft SQL Server that included each of the 5 tables shown in the data relationship diagram (**Figure A-7**). Then 4 separate queries were performed in order to create one large data table which included all the necessary attributes pertinent to the original model framework described previously. Variables to be contained in the final dataset can be grouped into three factor categories: general impedance factors, built environment factors, and demographic factors, and are depicted in the table below (**Table A-2**).

Variable	Type	Description
Impedance Factors		
Distance	continuous	distance between O-D as recorded in survey
Slope	continuous	absolute value of difference in elevations between O and D divided by the O-D distance
Sunrise time	continuous	time of day of sunrise
Sunset time	continuous	time of day of sunset
Rain	continuous	(based on quartile values from annual precipitation distribution)
Built Environment Factors		
PathDistO	continuous	Euclidean distance between trip O/D and nearest shared pedestrian trail / bicycle path
PathDistD	continuous	
XWalksCountO	discrete	Number of cross-walks within 1 mile radius of trip
XWalksCountD	discrete	
BlocksD	discrete	Number of blocks witin 1 contained in 1 mi radius buffer of trip O/D
BlocksO	discrete	
Demographic Factors		
Sex	binary	Gender of trip taker (1 =male, 0=female)
Age	ordinal	Age category of trip taker (as indicated from survey question)
Ldrv	binary	Has valid driver's license (1 if yes, 0 if no)
Edu	ordinal	Highest level of education attained (as indicated from household survey)
hinccat2	ordinal	Household income level (from survey)
Hhnumveh	discrete	number of vehicles owned per household

Table A-2: Description of Final Dataset Variables

Further analysis was done do convert the continuous rain variable, "Rain", into an ordinal rain variable, to indicate the quantity of rain, ranging from light rain (=1) to heavy rain (=3), and the Sunrise and Sunset variables into one binary variable *darkness*, to indicate whether a trip took place before sunrise or after sunset. A description of these variables is illustrated in the table below.

Variable	Units	Type	Description
Rain	----	*ordinal*	1= light rain , 2 = moderate rain, 3= heavy rain (based on quartile values from annual precipitation distribution)
Darkness	----	*binary*	equals 1 if after dusk / before dawn, else 0

Table A-3: *Rain and Darkness Variables:* The continuous rain variable was converted into a three level ordinal variable with value thresholds based upon the quantile values from the annual precipitation distribution. The *Darkness* variable was created from the sunrise and sunset values and is a binary indication of whether or not a trip took place during non-daylight hours.

An updated table of the final model data set variables including the new Rain and Darkness variables is shown below (**Table A-4**).

Variable	Units	Type	Description
Impedance Factors			
Distance	*miles*	*continuous*	distance between O-D as recorded in survey
Slope	*ft/mile*	*continuous*	absolute value of difference in elevations between O and D divided by the O-D distance
Rain	----	*ordinal*	1= light rain , 2 = moderate rain, 3= heavy rain (based on quartile values from annual precipitation distribution)
Darkness	----	*binary*	equals 1 if after dusk / before dawn, else 0
Built Environment Factors			
PathDistO	*feet*	*continuous*	Euclidean distance between trip O/D and nearest shared pedestrian trail / bicycle path
PathDistD	*feet*	*continuous*	
XWalksCountO	*cross-walks*	*continuous*	Number of cross-walks within 1 mile radius of trip O/D
XWalksCountD	*cross-walks*	*continuous*	
BlocksD	*blocks*	*continuous*	Number of blocks witin 1 contained in 1 mi radius buffer of trip O/D
BlocksO	*blocks*	*continuous*	
Demographic Factors			
Sex	----	*binary*	Gender of trip taker (1 =male, 0=female)
Age	----	*ordinal*	Age category of trip taker (as indicated from survey question)
Ldrv	----	*binary*	Has valid driver's license (1 if yes, 0 if no)
Edu	----	*ordinal*	Highest level of education attained (as indicated from household survey)
hinccat2	----	*ordinal*	Household income level (from survey)
Hhnumveh	*vehicles*	*discrete*	number of vehicles owned per household

Table A-4: Updated Variable Table

Resources

1. Ahmed, F., Rose, G., Figliozzi, M., & Jakob, C. (2012). Commuter Cyclist's Sensitivity to Changes in Weather: Insight from Two Cities with Different Climatic Conditions. In *Transportation Research Board Annual Meeting, 91st, 2012, Washington, DC, USA* (No. 12-2233).
2. Akar, G., & Clifton, K. J. (2009). Influence of individual perceptions and bicycle infrastructure on decision to bike. *Transportation Research Record: Journal of the Transportation Research Board*, *2140*(1), 165-172.
3. Brezina, E., & Kramer, M. (1970). *AN INVESTIGATION OF RIDER BICYCLE AND ENVIRONMENTAL VARIABLES IN URBAN BICYCLE COLLISIONS* (No. Study Rept).
4. Center for Disease Control and Prevention. Heart Disease Facts. (2015) Source: http://www.cdc.gov/heartdisease/facts.htm. Accessed Jan. 12, 2015.
5. Cervero, R., & Duncan, M. (2003). Walking, bicycling, and urban landscapes: evidence from the San Francisco Bay Area. *American journal of public health*,*93*(9), 1478-1483.
6. Cervero, R., Sarmiento, O., Jacoby, E., Gomez, L.F., Neiman, A. (2009) "Influences of Built Environments on Walking and Cycling: Lessons from Bogota." *International Journal of*
7. De Hartog, J. J., Boogaard, H., Nijland, H., & Hoek, G. (2010). Do the health benefits of cycling outweigh the risks?. *Environmental health perspectives*,*118*(8), 1109.
8. Dill, J., Gliebe, J. (2008) "*Understanding and Measuring Bicycling Behavior: A Focus on Travel Time and Route Choice.*" Final Report Oregon Transportation Research and Education Consortium-RR-08-03.
9. Dunn, A.L., Marcus, B.H., Kampert, J.B., Garcia, M.E., Kohl, H.W., Blair, S.N., (1999). " Comparison of lifestyle and structured interventions to increase physical activity and cardiorespiratory fitness: A randomized trail." *Journal of American Medical Association* 281(4), 327-334.
10. El Esawey, M., Lim, C., Sayed, T., & Mosas, A. I. (2013). Development of Daily Adjustment Factors for Bicycle Traffic. *Journal of Transportation Engineering*.
11. Ewing, R., Cervero, R. (2010) "Travel and the Built Environment: A Meta-Analysis." J*ournal of the American Planning Association*. 76 (3): 265-294.
12. Gallop, C., Tse, C., & Zhao, J. (2012). A Seasonal Autoregressive Model of Vancouver Bicycle Traffic Using Weather Variables. In *Transportation Research Board 91st Annual Meeting* (No. 12-2119).
13. Handy, S. L., Boarnet, M. G., Ewing, R., & Killingsworth, R. E. (2002). How the built environment affects physical activity: views from urban planning.*American journal of preventive medicine*, *23*(2), 64-73.
14. Hanson, S., & Hanson, P. (1977). Evaluating the impact of weather on bicycle use. *Transportation Research Record*, (629).
15. Haskell WL, Lee IM, Pate RR, Powell KE, Blair SN, Franklin BA, et al. 2007. Physical activity and public health: updated recommendation for adults from the American College of Sports Medicine and the American Heart Association. Med Sci Sports Exerc 39:1423–1434.
16. James, G., Witten, D., Hastie, T., Tibshirani, R. (2013). An Introduction to Statistical Learning with Applications in R. NewYork, New York: Springer.

17. Khan, M., Xiong, X., Kockelman, K. (2014) "Models for anticipating non-motorized travel choices, and the role of the built environment." *Transport Policy*, 35: 117-246.
18. Koplan, Jeffrey, Dietz, William. (1999). Caloric Imbalance and Public Health Policy‖, JAMA.
19. Kuzmyak, J., R., Kockelman, K., Bowman J., Bradle M., Lawton, K., Pratt, R. H. (2011).*Estimating Bicycling and Walking for Planning and Project Development*; Task 3 Interim Report. NCHRP Project No. 08-78.
20. Lee, I. M., & Skerrett, P. J. (2001). Physical activity and all-cause mortality: what is the dose-response relation?. *Medicine and science in sports and exercise*, *33*(6; SUPP), S459-S471.
21. Lewin, A. (2011). Temporal and weather impacts on bicycle volumes. In *Transportation Research Board 90th Annual Meeting* (No. 11-2536).
22. Lindsey, G., Wilson, J., Rubchinskaya, E., Yang, J., & Han, Y. (2007). Estimating urban trail traffic: Methods for existing and proposed trails.*Landscape and urban planning*, *81*(4), 299-315.
23. Miranda-Moreno, L. F., & Nosal, T. (2011). Weather or not to cycle.*Transportation Research Record: Journal of the Transportation Research Board*, *2247*(1), 42-52.
24. Moudon, A. V., Lee, C., Cheadle, A. D., Collier, C. W., Johnson, D., Schmid, T. L., & Weather, R. D. (2005). Cycling and the built environment, a US perspective. *Transportation Research Part D: Transport and Environment*,*10*(3), 245-261.
25. Nankervis, M. (1999). The effect of weather and climate on bicycle commuting.*Transportation Research Part A: Policy and Practice*, *33*(6), 417-431.
26. Niemeier, D. A. (1996). Longitudinal analysis of bicycle count variability: results and modeling implications. *Journal of Transportation Engineering*, *122*(3), 200-206.
27. Noland, R. B., & Kunreuther, H. (1995). Short-run and long-run policies for increasing bicycle transportation for daily commuter trips. *Transport Policy*,*2*(1), 67-79
28. Pratt, R. (2012). *Traveler Response to Transportation System Changes: Pedestrian and Bicycle Facilities*. Transit Cooperative Highway Research Program (TCRP) Report 95, Chapter16.
29. Pucher, J., & Buehler, R. (2006). Why Canadians cycle more than Americans: a comparative analysis of bicycling trends and policies. *Transport Policy*, *13*(3), 265-279.
30. Pucher, J., Transportation Quarterly, 98-1: Source: http://www.ibike.org/library/statistics-data.htm. Accessed 2/12/2015
31. Rietveld, P., & Daniel, V. (2004). Determinants of bicycle use: do municipal policies matter?. *Transportation Research Part A: Policy and Practice*, *38*(7), 531-550.
32. Rodríguez, D. A., & Joo, J. (2004). The relationship between non-motorized mode choice and the local physical environment. *Transportation Research Part D: Transport and Environment*, *9*(2), 151-173.
33. Saelens, B. E., & Handy, S. L. (2008). Built environment correlates of walking: a review. *Medicine and science in sports and exercise*, *40*(7 Suppl), S550.
34. Saelens, B. E., Sallis, J. F., & Frank, L. D. (2003). Environmental correlates of walking and cycling: findings from the transportation, urban design, and planning literatures. *Annals of behavioral medicine*, *25*(2), 80-91.

35. Santos, A., McGuckin, A.N., Nakamoto, H.Y., Gray, D., Lis, S., (2011). "*Summary of Travel Trends: 2009 National Household Travel*." Federal Highway Administration, Washington, DC.
36. Schroeder, P., Wilbur, M. (2013). "*2012 National survey of bicyclist and pedestrian attitudes and behavior, volume 1: Summary report*." (Report No. DOT HS 811 841 A). Washington, DC: National Highway Traffic Safety Administration.
37. Sener, I.N., Eluru, N., Bhat, C.R., (2009). "An analysis of bicycle route choice preferences in Texas." *Transportation* 36(5), 511-539.
38. U.S. Department of Transportation, Federal Highway Administration, 2009 National Household Travel Survey. URL: http://nhts.ornl.gov.
39. Warburton, D. E., Nicol, C. W., & Bredin, S. S. (2006). Health benefits of physical activity: the evidence. *Canadian medical association journal*, *174*(6), 801-809.