University of Cincinnati

Date: 11/5/2013

I, Maria F. Ramirez Bernal, hereby submit this original work as part of the requirements for the degree of Master of Science in Civil Engineering.

It is entitled:
Development of the Bicycle Compatibility Evaluator (BCE) for the city of Cincinnati, OH

Student’s name: Maria F. Ramirez Bernal

This work and its defense approved by:

Committee chair: Heng Wei, Ph.D.
Committee member: Richard Miller, Ph.D.
Committee member: Danilo Palazzo, Ph.D.
Development of the Bicycle Compatibility Evaluator (BCE) for City of Cincinnati, OH

A thesis submitted to the Graduate School of
The University of Cincinnati
in partial fulfillment of the requirement for the degree of

MASTER OF SCIENCE

In the Department of Civil and Architectural Engineering and Construction Management
of the College of Engineering and Applied Science
November 2013

By

Maria Fernanda Ramirez-Bernal

MCP. University of Cincinnati
Road design Specialist. Escuela Colombiana de Ingenieria - Colombia
Civ. Eng. Escuela Colombiana de Ingenieria - Colombia

Committee:
Dr. Heng Wei, Chair
Dr. Richard Miller
Dr. Danilo Palazzo
Abstract

Global warming is a result of most of increasing contaminant components in our atmosphere, many of them generated by motor vehicles. It causes harm to human health, not only to individuals but to communities and towns as a whole. Increasing heat, extreme cold, rising sea levels and flood plains damage each habitat’s equilibrium.

It has been proven that public transit leads to more walking and biking. Surveys report that 63% and 68% of pedestrians and bikers walk at least once a day, respectively. Public transit commuters walk more than car commuters. Walking to those neighborhood destinations accounts for higher frequencies than public transit does. Integration between land use areas, enhancing mixed-use parcels are a complement and support to a public transportation network that includes pedestrians and bikers. Biking infrastructure comparisons have been made among cities and countries; unfortunately, there aren’t many studies that compare and analyze characteristics in metropolitan areas. This lack of analysis explains the gap of infrastructure improvements in cities where using bicycles as active modes of transportation is barely at its beginnings.

Bicycle infrastructure innovation is often related to Master Plans as those quantify the needs of pedestrians and bikers at different levels (city, region, state). As a result of some studies, there are prioritized lists of pedestrian and bicycle corridors according to need and importance to generate an enhanced and improved master plan.

Household travel surveys (HTS) supply multi-level information regarding travel patterns per day. They provide information regarding the purpose of the trips, mode choice, trip length, location of activities and, most importantly for this research, routes taken. A large-scale HTS was been developed for the Cincinnati Metropolitan Area between the months of August 2009 and August 2010. GPS units were
given to each participant household member 12 years or older to carry all day for 3 days. Each household member was requested to provide demographic and location information. Once the survey was concluded, the information was located on the map and compiled as Google Earth maps.

This research will study collected demographic and route data from the household transportation survey and add it to infrastructure attributes from the street network and its corresponding land use zones.

After an adequate statistical analysis, this research will develop a methodology for estimating the coefficients for a linear equation, called the Bicycle Compatibility Evaluator (BCE); that correlates “human and street” characteristics to determine usage of existing infrastructure and evaluate how compatible they are with one another. This process helps establish factors that affect ridership in a particular metropolitan area of study.

This methodology will present a list of potential variables to be included in an evaluation equation and how will they be included in or rejected from the equation. Each one will be either rejected or accepted and to finally estimate their coefficients. Lastly, it will estimate a cut-off match and mismatch compatibility value between streets and individuals.
Para el papá y la mamá, porque son la fuerza que me mantiene viva.

Para Cata, Mona, Panchita y Oswaldo, porque no dejaron de creer en mí. Nunca.

Para Luis Fernando, porque sí.
Acknowledgement

Foremost, I would like to extend my gratitude to my adviser Dr. Heng Wei for his support and guidance, in particular for his great patience in supervising me to conduct research. His work has been indispensable for developing my research topic. It is a great pleasure to work with such a prestigious professional in the transportation field. I am grateful for the opportunity provided by Dr. Heng Wei to participate in a research project for developing the bicycle travel demand forecasting model with the dataset of the Greater Cincinnati area, which was funded by the Ohio Department of Transportation (ODOT). That project gave me the opportunity to interact with transportation engineers and planners.

To Doctors Richard Miller and Danilo Palazzo, thank them for their support and feedback as committee members. I feel great appreciation for Drs. Miller and Palazzo as they had found time to collaborate with my research and provide insightful and valuable comments for its development.

Also, gratitude goes to Doctor Michael Romanos, School of Planning at UC, for his valuable guidelines and support. Dr. Thomas Wuerzer (Department of Community and Regional Planning – Boise State University) for sharing ideas on cartographic data management and applications to bicycle related projects, GIS processes and modeling.

To Mr. Andrew Rohne (OKI) and Mr. James Coppock (Cincinnati City Planning office), who have been a support system, not only by providing the data for this project but the time they’ve taken to explain processes and real hands-on work around the Cincinnati area. To Mr. Tarun Waishal at ESRI and Dr. Peter Kimosop at Bowling Green University for being technical support assistantship team and additional methodology guides.

My appreciation goes to Mr. Nestor Mancipe (PhD candidate) for taking the time to brainstorm and GIS tutoring.
# Table of Contents

Abstract ................................................................................................................................. ii

Acknowledgement .............................................................................................................. vi

List of Tables ....................................................................................................................... ix

List of Figures ...................................................................................................................... x

1. Introduction ...................................................................................................................... 11
   Background ......................................................................................................................... 11
   Problem Identification ....................................................................................................... 12
   Thesis structure .................................................................................................................. 13
   Scope of study ..................................................................................................................... 14
   Hypothesis .......................................................................................................................... 14
   Objectives ......................................................................................................................... 15

2. Related works and literature review ............................................................................. 16
   The impact of urban distribution (land use) in biking and multimodal transportation systems 17
   Bicycle infrastructure studies and evaluation .................................................................... 21
   Transportation demand forecasting methods ...................................................................... 28
   Bicycle travel demand forecasting ................................................................................... 30
   Household travel surveys .................................................................................................. 31
   Cincinnati household travel survey .................................................................................. 32

3. Methodology ................................................................................................................... 34
   Preliminary data ................................................................................................................ 34
   Mode-based household data selection ............................................................................. 41
   Mapping GoogleEarth data .............................................................................................. 43
   Finding HTS segments on the full street network ............................................................. 45
   Demographic data coding and incorporation .................................................................... 47
   Incorporation and recoding of land use data ..................................................................... 49
   “Predisposition to biking” analysis and mapping .............................................................. 49
   Spatial correlation analysis .............................................................................................. 52
   Ordinary least squares ....................................................................................................... 56
   Variable selection and filtering ........................................................................................ 59
   Final results ....................................................................................................................... 59
   Methodology schematics ................................................................................................. 60

4. Data Preparation .............................................................................................................. 61
   Collected data ..................................................................................................................... 61
   Bicycle-only data .............................................................................................................. 65
   Visible bicycle routes ........................................................................................................ 65
   Segments to be evaluated ................................................................................................. 67
   HTS information incorporated into full OKI street network ............................................. 69
OKI network with full HTS information and land use data.

5. Data Analysis .........................................................................................................................74
   Mapped density on “Predisposition to bike”  74
   Spatial correlation analysis results  83

6. Results ........................................................................................................................................84
   Ordinary Least squares  84
   Filtered variables  84
   BCE equation coefficients  87
   BCE validation  93

7. Conclusions ..................................................................................................................................95

8. Directions for future research ....................................................................................................96

References........................................................................................................................................97
List of Tables

Table 1. BEQI Variables ............................................................................................................................ 23
Table 2. Bicycle Mode Index Calculation .................................................................................................. 25
Table 3. Bicycle Compatibility Index Levels ............................................................................................. 26
Table 4. Surveyed Household Statistics .................................................................................................. 33
Table 5. HTS Variable List ....................................................................................................................... 36
Table 6. Organized Household HTS Data (Sample) ................................................................................. 42
Table 7. Organized Trip HTS Data (Sample) ........................................................................................... 43
Table 8. HTS Variables To Be Studied ..................................................................................................... 47
Table 9. Demographic Date Coding ....................................................................................................... 48
Table 10. Land Use Re-coding ................................................................................................................ 49
Table 11. Statistical Significance ............................................................................................................ 55
Table 12. Variable Coding ...................................................................................................................... 59
Table 13. Data Dictionary Characteristics for Household Data from the HTS ......................................... 62
Table 14. Data Dictionary Characteristics For Personal Data From The HTS ........................................ 63
Table 15. Data Dictionary Characteristics For Trip Data From The HTS ................................................ 64
Table 16. Spatial Correlation Results ...................................................................................................... 83
Table 17. OLS Analysis Results .............................................................................................................. 85
Table 18. OLS Analysis Results. First Filter ............................................................................................ 86
Table 19. OLS Analysis Results. Second Filter ....................................................................................... 86
Table 20A. Demographic Variables and Their Coefficients ................................................................. 87
Table 20B. Street Variables and Their Coefficients ................................................................................. 87
List of Figures

FIGURE 1. TRIP DISTRIBUTION – OKI HTS ................................................................. 17
FIGURE 2. TRIP LENGTH – OKI HTS ........................................................................ 19
FIGURE 3. BICYCLE TRIP LENGTH – OKI HTS .......................................................... 20
FIGURE 4. BICYCLE QUALITY ENVIRONMENTAL INDEX ...................................... 24
FIGURE 5. FOUR-STEP TRAVEL FORECASTING MODEL ......................................... 29
FIGURE 6. MAP AND DATA FOLDER FROM CAGIS .................................................. 39
FIGURE 7. CINCINNATI STREET MAP ....................................................................... 40
FIGURE 8. DATA HYPOTHESIS CORRESPONDENCE ............................................ 41
FIGURE 9. TRIP SAMPLE DATA .................................................................................. 42
FIGURE 10. GPS DATA – GOOGLE EARTH DATA ..................................................... 44
FIGURE 11A. CORRECT STREET SEGMENT SELECTION BY ALGORITHM .................. 46
FIGURE 11B. WRONG STREET SEGMENT SELECTION BY ALGORITHM .................... 46
FIGURES 12. LINE DENSITY ....................................................................................... 50
FIGURE 13. LINE DENSITY INPUT DIALOG BOX..................................................... 51
FIGURE 14. DISPERSED FEATURES VS. CLUSTERED FEATURES ............................ 52
FIGURE 15. SPATIAL AUTOCORRELATION INPUT DIALOG BOX ............................... 55
FIGURE 16. ORDINARY LEAST SQUARES FLOWCHART .......................................... 56
FIGURE 17. ORDINARY LEAST SQUARES INPUT DIALOG BOX ................................. 58
FIGURE 18. METHODOLOGY FLOW CHART ............................................................... 60
FIGURE 19A. ROUTE TAKEN BY SUBJECT 101214001 ............................................ 66
FIGURE 19B. ROUTE TAKEN BY SUBJECT 101270004 ............................................ 67
FIGURE 20. RIDDEN ROUTES MAP OVER OKI’S STREET NETWORK ....................... 68
FIGURE 21. RIDDEN ROUTES OVER OKI’S STREET NETWORK (MALE/FEMALE) ...... 70
FIGURE 22. OKI’S STREET NETWORK SELECTED BY LAND USE ............................ 72
FIGURE 23. RIDDEN ROUTES ON THE OKI STREET NETWORK ............................... 73
FIGURE 24A. DENSITY ANALYSIS – LICENSED DRIVER ........................................... 75
FIGURE 24B. DENSITY ANALYSIS – HH BIKES ....................................................... 76
FIGURE 24C. DENSITY ANALYSIS – HH VEHICLES .................................................. 77
FIGURE 24D. DENSITY ANALYSIS – HH INCOME .................................................... 78
FIGURE 24E. DENSITY ANALYSIS – HH STRATA .................................................... 79
FIGURE 24F. DENSITY ANALYSIS – HH TYPE .......................................................... 80
FIGURE 24G. DENSITY ANALYSIS – CHECK ........................................................... 81
FIGURE 25A. ESTIMATED BCE - MAP ................................................................. 89
FIGURE 25B. ESTIMATED BCE – HISTOGRAM ....................................................... 90
FIGURE 26A. BCE RESIDUALS - MAP ................................................................. 91
FIGURE 26B. BCE RESIDUALS – HISTOGRAM ....................................................... 92
FIGURE 27. BCE VALIDATION – INVERSE CUMULATIVE FREQUENCIES .............. 94
1. Introduction

Background

Global warming is a result of most of increasing contaminant components in our atmosphere, many of them generated by motor vehicles. It causes harm to human health, not only to individuals, but also to communities and towns as a whole. Increasing heat, extreme cold, rising sea levels and flood plains damage each habitat’s equilibrium (Koren et al., 2006). On the short term, some individual’s may be affected more than others: minorities, young children and elderly; but in the end, everyone will be at some point touched by the problem.

The City of Cincinnati started a program to include bicycles as part of the city’s multimodal transportation system a couple of years ago and has developed studies to better understand bikers’ behavior and improve the quality of non-motorized infrastructure. A model like this one needs to be based on the user’s personal decisions to use bicycles to move around the city.

In 2010 the Ohio – Kentucky – Indiana Regional Council of Governments (OKI) completed a household travel survey within the Cincinnati Metropolitan area providing GPS devices to a large number of households to gather enough personal, demographic, economic and travel data to be able to predict future traffic from all types of transportation modes and develop studies and programs to provide a better service to the population. According to OKI’s records the number of cyclists has been increasing since 1990s.
Problem Identification

There is little to no research on the relationship between demographic trends and route choosing based on Household Travel Surveys.

This research will present a methodology for developing an equation for the OKI research area to analyze quality of service provided by existing infrastructure considering personal characteristics of the population. This will be called the “Bicycle Compatibility Evaluator”, or BCE. The BCE will be developed as a lineal equation in which all adequate household travel survey variables are included. By using this equation, it will be possible to determine if the street infrastructure is compatible with demographic characteristics from specific groups of the population.

Studies like this will provide local governments’ tools to base their decisions of street improvements based on real biker usage of the network and what those users tend to look for infrastructure-wise. If the BCE indicates non-existing compatibility, then it will be possible to make changes in the existing infrastructure to modify the value of the mismatched value and meet compatibility requirements.

Household surveys are a valuable source of data, but rare due to the high cost and considerable technical requirements; the fact that OKI has completed one for the city of Cincinnati and is willing to share this information for academic reasons is an excellent opportunity that must not go to waste.

GIS data as complete as the one from the city of Cincinnati is not common either. Again, these types of data collections are expensive and time consuming. This project will take the information provided by CAGIS, OKI and other agencies working on the area compile it and analyze it to be able to develop better understanding of bicycle transportation trends.

Changes on human behavior are not within the scope of this research. Such analyses require anthropological studies. Expertise and skills are not considered in this research as they were not originally
considered in the original OKI study and, as part of the agreement confidentiality agreement, the student researcher was not allowed to contact individuals already surveyed in this study.

**Thesis structure**

The structure of this thesis is organized as follows:

Chapter 2. Related works and Literature review: summarizes the state of research in urban physical distribution, multimodal transportation systems, bicycle infrastructure studies, household travel surveys and the Cincinnati household travel survey.

Chapter 3. Methodology of research: presents a step-by-step of the process: selecting, organizing and mapping the data; analysis of cluster patterns; rejecting irrelevant information; estimating coefficients based on selected variables and validating the equation.

Chapter 4. Data preparation: Shows how the real data is cleaned and organized for mapping, and how data is incorporated into a unified database.

Chapter 5: Data Analysis: Presents how data physical organization on the map determines its validity.

Chapter 6: Results: Real data analysis, variable filtering, coefficient estimation and equation validation.

Chapter 7: Conclusions.

Chapter 8: Directions for future research.
**Scope of study**

Data for this research has been provided by OKI, it comes from the eight counties in the Ohio-Kentucky-Indiana Metropolitan Area: Hamilton, Clermont, Butler, Warren (Ohio), Boone, Kenton, Campbell (Kentucky) and Dearborn (Indiana). All variables analyzed and selected in this study come from the OKI household transportation survey database. To keep the concordance of the research, all coding values are kept as they are in the original system.

Data from the Cincinnati Metropolitan Area was collected throughout 12 months, between August 2009 and August 2010. During this period 2,059 households participated in the survey, a total of 3,849 people provided trip data (60,900 trips) from which 689 were bicycle trips.

It is understood that all available database is not large enough to provide highly precise estimations but it will serve the purpose of being a base for putting together a methodology plan and determine how to validate a proposed model.

Also, this study is exclusively dedicated to analyze trips that have already taken place; there is no route or mode choice analysis. The only mode of transportation being analyzed is the bicycle. Also, individual characteristics are based on demographic characteristics found on the survey’s database; following the confidentiality agreement between the University of Cincinnati and OKI there wasn’t any interaction with the riders.

**Hypothesis**

Economic and demographic backgrounds are decisive factors in the bicycle route personal choosing process.

Even though it can be thought that bicyclist tend to choose the shortest path to go from point A to point B, personal route choosing can be based on traffic patterns (slow traffic streets and number and direction of lanes) and land use.
The effect of these variables as it each one of them can be quantified from the household travel survey data.

It will be possible to develop an evaluation tool to pair personal characteristics from riders and street infrastructure characteristics through a lineal equation and determine if they match or not. Such evaluation tool will come in the form of a lineal equation:

$$BCE=\sum_{i=1}^{n} \alpha_i D_i + \sum_{j=1}^{n} \beta_j S_j$$ (1)

Where $\alpha$ and $\beta$ are estimated coefficients for proposed Demographic (D) and Street (S) characteristics. Each one of those characteristics can be determined during the planning stages of a city’s household travel survey.

**Objectives**

To prove the previously stated hypothesis, this project seeks to achieve the following objectives: (1) Develop a methodology to analyze and quantify the impact of demographic background on how the bikers on a metropolitan area make route choice decisions; (2) Determine a link between use of street segments and the characteristics of the bikers that circulate on them; (3) Establish if even though it can be thought that bicyclist tend to choose the shortest path to go from point A to point B, many times the personal route choosing process can be based on other factors like: slow traffic streets, parking lanes and land use; (4) Determine a lineal correlation between variables and the bikers’ individual characteristics and; (5) Develop the Bicycle Compatibility Evaluator (BCE) as an evaluation tool to match riders and street infrastructure.
2. Related works and literature review

“Health interventions that have a beneficial effect on climate change include encouraging patients to reduce the amount of red meat in their diets and to replace some vehicular transportation with walking or bicycling” (Parker, 2011). There is an increasing number of people who ride bicycles to lessen the use of resources (like income and gas), reduce the emission of carbon dioxide and increase their quality of life by making healthier decisions. Nevertheless, there are times when cyclists must use car lanes or sidewalks because there aren’t bike paths in the area putting pedestrians’ lives in danger while trying to protect their own (Sugo et al., 2010).

In some countries, like Hong Kong, cyclists are exposed to high risks of injury and fatality in crashes. It has been proven that some of the significant factors that affect these indexes are demographics, temporal distribution, cyclist behavior, road conditions, helmet use and, alcohol intoxication (Su et al., 2010). Considering how biking as a mode of transportation in the United States, and particularly Cincinnati is at its early stages, these factors, would be a good starting point to understand the reasons why people don’t consider biking a safe way of traveling short distances (Sze et al., 2011).

Countries with higher levels of active transportation (walking or biking) have the lowest obesity rates. Countries in Europe, for example, where residents walk 2.5 times a longer distance when compared to the United States (382 vs. 140km) have a considerably lower obesity rate. It can be stated that “active transportation is inversely related to obesity in countries where is practiced” (Bassett et al., 2008). It’s not definitive, but it suggests one of the explanations for international differences in obesity rates. Adults between the ages 20 and 65, with routines that include active transportation are proven to have better health and enjoyment of physical activities, which leads to healthier states of mind (Lachapelle et al., 2011).
The impact of urban distribution (land use) in biking and multimodal transportation systems

Another topic that will affect the impact of bicycle use is the size of cities, an issue that has been studied for long times. Proposals on the adequate size of a city are debatable. Studies suggest that cities structured as cells are better shaped for bikers than those that are sprawled; areas near the core become denser and trips shorter. Hence, reducing the need to use vehicles and increasing the use of non-motorized vehicles. Gruen’s principle (1973) states: “mass transit modes radiate from dense nuclei outward” (Gruen, 1973).

As an example, Figure 1 shows the Cincinnati household travel survey trip distribution. It can be seen that for the Cincinnati sample, more than four fifths of the trips were done using motor vehicles. 8.8% of the total trips used a bicycle and only. Bicycle is equivalent to one tenth of the motor vehicle trips.

FIGURE 1. TRIP DISTRIBUTION – OKI HTS
Source: OKI, 2010
It has been proven that public transit leads to more walking and biking. Telephone surveys from NJ and CA report that 63% and 68% of pedestrians and bikers walk at least once a day (respectively). Public transit commuters walk more than car commuters (Lachapelle and Noland 2012).

The physical distribution and characteristics of the route can influence people’s attitude towards biking. Associations between geometric characteristics of the route and the surrounding environment have been analyzed to determine how these factors contribute to the topic at hand. In Greater Stockholm (Sweden), the Active Commuting Route Environmental Scale (ACRES) recruited bikers and asked them to ride different routes under different environmental, traffic and greenery surroundings and enquire about their perception of safety (safe/unsafe). In the end, the results shown that those areas perceived as “beautiful, safe and green”, where highly preferred by bikers that those with higher levels of exhaust fume, congestion and “low-directedness” (Wahlgren et al., 2012).

Walk to close neighborhood destinations accounts for higher frequencies than public transit. Integration between land use areas, enhancing mixed-use parcels complement and support to a public transportation network that includes pedestrians and bikers. Figure 2 shows trip length distribution in the Cincinnati metropolitan area HTS. As it can be seen, 50,000 (67% of total) trips are less than 5 miles long; trips that can be easily done by bicycle or public transportation if and when there is adequate infrastructure and connecting networks. (Lachapelle et al., 2012).
Some other studies have shown that large areas like the Washington D.C. Metropolitan Area, which operates the second-busiest heavy transit system in the country, serving 3.5 million people in an over 1.5 square miles area- have studied the potential and implemented ideas of connecting its MetroRail system with other modes of transportation. A 2009 study collected surveys from riders showing that ridership would increase if areas with alternative and complementary activities were connected by the system; sporting events and downtown activities, for example (Jia, 2009). Bikeable communities were considered a “complement to the complement” considering is would be an intermediate and non-expensive mode of transportation in between other modes. Figure 3 presents the OKI bicycle trip length data. It can be seen that all but a dozen of trips are less than 10 miles long, proving the cited findings on length and trip distribution. Bicycle trips are mostly short (less than 2 miles long) which are believed to
be non-commuters. It should be pointed that it will be necessary to create an additional survey in which the bikers can explain the reasons why they chose this form or transportation and the type of trip (commuter, non-commuter).

Interactions among modes of transportation can be studied using vehicle ownership models with trip survey data. Results from Asian studies (Yamamoto, 2009) show that in residential area, ownership distribution is highly impacted by area density based on the type of urban development, whereas in denser areas with green spaces, private vehicle ownership is lower and bicycle ownership is considerably larger when compared to others.
Frantszeskakis analyzed the urban history of Islamabad (Pakistan) as it’s known to becoming one of the fastest growing cities in Asia with the inclusion of the city of Rawalpindi to its urban core. He analyzes the innovations of the city’s urban plans and the impact on “categories of movements” or modes of transportation on trip lengths. According to his analysis, these lengths will be larger as the urban area expands and the population and car ownership increases, not only because lengths between destination and origin points become longer but because the road network will “not be subject to the capacity and related serious environmental problems” these changes are due to bring (Frantzeskakis, 2009).

Studies within the United States have provided insight on whether active transportation, most specifically biking, can help Americans fit in the recommended healthy levels of physical activity and its biased by infrastructure quality and characteristics. In the Portland Metropolitan Area (Oregon), 60% of a group of surveyed cyclists rode for more than 150 minutes per week for utilitarian (not exercise) purposes. Most of these trips took place on bicycle boulevards, bicycle lanes and separate paths, which encourages adult bicycling (Dill, 2009).

**Bicycle infrastructure studies and evaluation**

Biking infrastructure comparisons have been made among cities and countries, unfortunately, there aren’t many studies that compare and analyze characteristics in metropolitan areas. This lack of analysis explains the gap of infrastructure improvements in cities where using bicycles as active modes of transportation is barely at its beginnings. Bonham (2008) has compared urban areas in Australia and the United States to develop an attribute analysis method and their interaction. The study concludes that there are five variables that affect those who cycle and to where they cycle: urban context, cycling context, local government policies and programs, culture of travel and, socio-economic and demographic characteristics (Bonham et al., 2008). Bonham’s study focuses on the latter and it incorporates it along with biker levels and expertise. The study’s results show that downtown areas with multimodal land use are places where bikers commute. Other areas, like suburbs and uptown areas haven’t been studied but they should be, based on disaggregation by gender and income level.
In the United States, the Safe Routes to School (SRTS) program was designed to support safe transportation and children’s active commuting to school. Areas in where these programs are implemented improve the bicycling environment, not only for children but adults as well. Surveys in SRTS areas estimate that in large urban areas, 65.5 million people benefit from the projects, making this a great contributor to physical activity among the population (Watson et al., 2008).

Regarding quality of service and infrastructure, it is known that the Florida Department of Transportation (FDOT) created a model that predicts the way bicyclists perceive the street network environment. It is based on the reasoning of the Level of Service (LOS) models (Petritsch et al. 2007). The data for this service index was obtained from the “Ride for Science field data collection event and video simulations”, in which participants provided feedback on how well or bad the road network fulfilled their needs and safety perception. This method is based on a Pearson correlation analysis and it interlocks 700 combined real-time cyclists’ observations. Data is then analyzed then as subgroups and cross sections (age, gender and experience, and residence location). It is believed that this methodology is reliable based on its correlation coefficient (R² = 0.74) and it can be shifted to a large percentage of US metropolitan areas.

Bicycle infrastructure innovation is often related to master plans as those quantify the needs of pedestrians and bikers at different levels (city, region, state). As an example, New Jersey’s authorities created a statewide inventory as part of their bicycle and pedestrian master plan which included existing infrastructure and related trip attractors like transit stations, commercial areas and schools; all on a GIS environment. Interactive maps ease the input process for users and help deliver infrastructure solutions for the population. As a result of the studies, there’s a prioritized list of pedestrian and bicycle corridors based on need and importance to generate an enhanced and improved master plan (Swords et al., 2004).

The Bicycle Environmental Quality Index (BEQI) is the result of a quantitative, observational survey to assess bicycle environment on roadways and evaluate streetscape improvements to be made to
promote bicycling in San Francisco (Realmuto, 2013). The San Francisco Department of Public Health (SFDPH) identified five main categories that embody physical environmental factors for bicyclists: intersection safety, vehicle traffic, street design, safety, and land use. From these categories, 22 indicators seemed to promote or discourage bicycle riding and connectivity to other modes of transport, as shown in Table 1.

<table>
<thead>
<tr>
<th>INTERSECTION SAFETY</th>
<th>VEHICLE TRAFFIC</th>
<th>STREET DESIGN</th>
<th>SAFETY</th>
<th>LAND USE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dashed intersection bicycle lane.</td>
<td>Number of vehicle lanes.</td>
<td>Presence or a marked area for bicycle traffic.</td>
<td>Bicycle/pedestrian scale lighting.</td>
<td>Line of site.</td>
</tr>
<tr>
<td></td>
<td>Parallel parking adjacent to bicycle lane route.</td>
<td>Connectivity of bike lanes.</td>
<td>Pavement type/condition.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Traffic volume.</td>
<td>Driveway cuts.</td>
<td>Street slopes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Percentage of heavy vehicles.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


Results from the BEQI reveal the relative quality of the biking environment at a street-level scale in select San Francisco neighborhoods. Use of the BEQI can translate environmental variables into a set of provisions for a healthy bicycle environment and a BEQI assessment can inform neighborhood planning and prioritize improvements through the land use plans and environmental assessments.
The North Central Texas Bicycle Need Index’s (from 1990 census) goal is to identify zones with high bicycle use to justify the need for bicycle facilities. It is used to find and identify zones that could increase the bicycle usage rate by improving the number and quality of non-motorized infrastructure (Turner et al., 1997).

The BCI preliminary process calculation goes as follows:

\[
BMS = 0.0299 \times AGE + 0.05459 \times LE + 0.00053 \times ED + 0.00335 \times WPR + 0.00026 \times PRD + 0.05 \times HR + 0.00398
\]  

by definition of the authors:

- **BMS**: Bicycle mode share
- **AGE**: percentage of residents under 16 years of age;
- **HR**: Number of hours worked per week;
- **LE**: Percentage of land devoted to employment uses;
- **PD**: Population density;
- **ED**: Employment density;
- **PRD**: Population density of residential land uses;
- **WPR**: Ratio of workers to population.

Considering how biking is an activity that is directly related to the human body (as it requires physical activity) as part of the study, the research has proven to include additional variables such as:
commute length, gender, age, annual income level, presence of bicycle facilities, climate, topography and personal perception of the rider. Unfortunately, the preliminary version didn’t include children and additionally, it was found that some of the variables were correlated among each other, so a simpler version was put together.

Each one of the zones, five variables are calculated and normalized by a region-wide factor to obtain an INDEX SCORE:

\[
\text{index score} = \frac{\text{traffic survey zone factor value}}{\text{regionwide factor value}}
\]  

Table 2 shows the Index Score amplification process by a factor given by:

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>DEFINITION</th>
<th>FACTOR</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance</td>
<td>% trips shorter than 5 miles</td>
<td>3</td>
<td>D x 3</td>
</tr>
<tr>
<td>Land Use</td>
<td>T</td>
<td>2</td>
<td>LU x 2</td>
</tr>
<tr>
<td>Median HHI</td>
<td></td>
<td>2</td>
<td>HH x 2</td>
</tr>
<tr>
<td>Population Density</td>
<td></td>
<td>1</td>
<td>PD x 1</td>
</tr>
<tr>
<td>Employment Density</td>
<td></td>
<td>1</td>
<td>ED x 1</td>
</tr>
<tr>
<td>Total</td>
<td>Sum</td>
<td></td>
<td>SCORE</td>
</tr>
</tbody>
</table>


The final SCORE is the region’s Bicycle Mode Index.

The advantages of this methodology are that it relates demographics with land use, which is something that has not been done in recent studies and even though is a multiple regression model, its calibration does not require complex calculations.

The problem with this model is that its $R^2$ value has been estimated for several case studies and the resulting average is 0.40, which is not highly representative for the sample.

The model does not consider the “bike to school” population, which in some areas of the country accounts for a high percentage of non-motorized transportation.
The Bicycle Compatibility Index (BCI), measures what the Highway Capacity Manual calls Level of Service (LOS) for cars, but for bicycles (Landis et al., 1997). The concept is not yet defined in design manuals and is mostly limited to “the impact of the bicycles on motor vehicles”. What the BCI does is evaluate the number of lanes and directions of travel, the geometry of curve and parking lanes, traffic volumes, speed, driveway density, presence of sidewalks and medians and type of roadside development (land use). A lower BCI value in an area represents extremely high bicycle compatibility; of course, a higher value, represents low bicycle compatibility. Having this in mind, the Highway Safety Research Center established BCI ranges associated with LOS and classify them according to “Compatibility levels”, shown in Table 3.

<table>
<thead>
<tr>
<th>LOS</th>
<th>BCI range</th>
<th>Compatibility Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>≤ 1.50</td>
<td>Extremely High</td>
</tr>
<tr>
<td>B</td>
<td>1.51 –</td>
<td>Very High</td>
</tr>
<tr>
<td>C</td>
<td>2.31 –</td>
<td>Moderately High</td>
</tr>
<tr>
<td>D</td>
<td>3.41 –</td>
<td>Moderately Low</td>
</tr>
<tr>
<td>E</td>
<td>4.41 –</td>
<td>Very Low</td>
</tr>
<tr>
<td>F</td>
<td>&gt; 5.30</td>
<td>Extremely Low</td>
</tr>
</tbody>
</table>

Source: Landis et al, 1997

The BCI is calculated according to the equation:

\[
BCI = 3.67 - 0.996BL - 0.41BLW - 0.498CLW + 0.002CLV + 0.0004OLV \\
+ 0.022SPD + 0.506PKG + 0.26AREA + AF
\]

(4)

Where:
BL = presence of a bicycle lanes or paved shoulder > 0.9m (no = 0, yes = 1)
BLW = Bicycle lane (or paved shoulder) width, meters (to the nearest tenth).
CLW = curb lane width, meters (to the nearest tenth).
CLV = curb lane volume, (vehicles per hour in one direction).
OLV = Other lanes volume – same direction (vehicles per hour).
SPD = 85TH percentile speed of traffic, kph.
PKG = presence of a parking lane with more than 30% occupancy (no = 0, yes = 1).
AREA = type of roadside development (residential = 1, other type = 0).
AF = adjustment factor for truck volumes, parking turnover and right turn volumes
The *Copenhagenize Index* (2013) is applied to cities and it assigns them a score for their efforts towards making bicycling a form of transportation that is accepted, practical, and most importantly feasible. The Copenhagenize’s motto is: “Every city used to be bicycle friendly before planners and engineers started to change the paradigm and plan for cars and relegate bicycle users, pedestrians and public transport users to third class citizens”. Its goal is to challenge cities to modernize by implementing bicycle infrastructure and policies.

Cities are given 0 – 4 points in 13 different categories with the possibility to additional points (12) for unique efforts or results. A maximum of 64 points could be awarded, which is later normalized to 100.

Categories:

- Advocacy: level of influence of advocacy groups.
- Bicycle Culture: establishment of bicycling as regular mode of transportation or “sub-culture”.
- Bicycle Facilities.
- Bicycle Infrastructure: presence of bike lanes, bike paths, sharrows.
- Bike Share Programs.
- Gender Split.
- Modal Share For Bicycles.
- Modal Share Increase Since 2006 (year when urban of cycling kicked off world-wide).
- Perception of Safety.
- Politics.
- Social Acceptance.
- Urban Planning.
- Traffic Calming.
Transportation demand forecasting methods

Transportation forecasting has traditionally followed a four-step model. A model first implemented on mainframe computers in the 1950s at the Detroit Area Transportation Study and the Chicago Area Transportation Study (CATS). The process starts with land use forecasting; generally, this is done for an area as a whole, providing total amounts for the local land use analysis (Zhou et al., 2009). Then population and employment forecasting amounts are estimated for zones (Traffic Analysis Zones, TAZ) within the region.

The classic four steps in the model (Tian et al., 1998) are shown in Figure 5 ("Metropolitan washington council of governments." 2013), and are as follows:

1. Trip generation: to determine trip origin and destination frequency in each zone (by trip purpose) as a function of demographics of households and land use, as well as socio-economic factors.

2. Trip Distribution: to match origins with destinations.

3. Mode Choice: to estimate the number of trips between origins and destination using a particular mode.

4. Route Assignment: for trips between an origin and destination by a particular mode.
Activity-based models predict traffic based on individuals’ location and the very specific activities they are conducting (school, work, shop, for example). These methods use behavioral theories while focusing on individuals and households (Cervero, 1996). Even though many U.S. metropolitan planning organizations rely on the traditional four-step model, some experts believe that activity-based provide better predictive capability due to their sensitivity to behavioral changes.

According to Mr. Andrew Rohne at OKI, many large MPOs in the country are moving (or have moved) towards an Activity-based model traffic estimation including Chicago, New York, Atlanta,
Miami, Houston, Minneapolis, Cleveland, Columbus, Cincinnati, Denver, Dayton, San Francisco, Sacramento and Jacksonville. Additionally, Knoxville, TN and Evansville, IN have a hybrid model.

The major premise of the activity-based models is that travel occurs from activities performed by people (the human component) and the travel decisions those activities include (personal schedule and mode, for example) (Khademi et al., 2011). Travel becomes one of many attributes of a system; a component of an activity scheduling decision.

The main advantage of the activity-based models is the possibility to include additional variables that cannot be considered on four-step models. Such possibilities present an additional use for these models, and that the capability to predict emissions and study air quality issues like risks due to exposure and the derived studies related to health impact. Demographic characteristics can be analyzed as well by creating user categories (age, gender, economic activity, for example) (Kim et al., 2008).

**Bicycle travel demand forecasting**

Commuting in the US can be done using private vehicles, public transportation, active transportation (bicycling or walking) or “working from home”. It has been estimated that the latter, even though it doesn’t involve real commuting or long distance displacement, creates a different type of traffic in which workers travel shorter distances and require less and less motorized forms of transportation since they generally move around a smaller action radii. In general, non-motorized mode of transportation analysis is empirical if not theoretical. There is a belief that households with higher income tend to work from home more than those with lower incomes; ironically, those households don’t bike or walk as much as others. Higher education levels are closely related to non-motorized modes. Car ownership, gender, race and neighborhood features affect modal choices as well (Dill, 2009).
Existing travel demand and traffic assignment models are calibrated for their particular needs. Demand models are usually calibrated to include non-motorized modes. There isn’t much evidence of the proportion and the reasons people travel routes deviating from the shortest-path or least-cost routes generated by transportation models. Despite of mode, it has been found that people do not detour far off the shortest route. It has been accounted that differences in the infrastructure suggest that bike commuters detour at will; some commuters prefer routes with more bicycle facilities (traffic-calming features, bike racks, and signage) than shortest-path routes (Winters et al., 2010).

Household travel surveys

Household travel surveys (HTS) supply multi-level information regarding travel patterns per day. They provide information regarding the purpose of the trips, mode choice, trip length and location of activities and, most importantly for this research, routes taken. “Diary information is retrieved through computer-assisted telephone interview (CATI) that compared with GPS measured travel suggests trip under-reporting ranges from 20 to 30 percent” (Forrest et al., 2005). It has been proven that human errors contribute because they create inaccuracies in origin-destination location errors and respondent fatigue, among others.

Household travel surveys have been integrated in active life and health studies by identifying trip purposes, lengths and related activities (Merom et al., 2010). The most common method consists of a repetitive sampling during a short time to the same household to evaluate trends and customs; others study behavior (or, in this case, travel tendencies) of the same household over an extended period of time to elaborate an extended-over-time trend study. Prior to surveying each household, a request for participation was sent.

Some household surveys are complemented with face-to-face interviews or, as stated before, self-reported questionnaires. Not only to obtain personal information regarding each one of the people living
on the studied area, but also to gather information about specific activities taking place on the destination and origin points. All this, to avoid making assumptions regarding the type of activity a person does at a particular location based on his/her age, gender or income.

Results of these surveys are normally used to determine trends and develop models by studying very detailed variables from actual travel information such as trip duration and length and mode of transportation. Also, household surveys help identify subgroups to make active travel interventions (Merom et al., 2010).

**Cincinnati household travel survey**

As an example of procedure and methodology, Howard et al. developed a suggested survey for expert cyclist to help determine the links between the routes they took and their existing conditions, and created profiles for non-motorized commuters around the Phoenix metropolitan area. Information was collected on personal diaries, regarding individual commuting (home-to-work). Additional data included gender and age characteristics, along with trip frequency based on street segments commonly used. Analysis included comparisons between the real routes and shortest distance, shortest time and safest route alternatives. As a result, it was possible to prove that bicyclist adjust their routes to fill the gaps infrastructure doesn’t meet (Howard et al., 2001).

A large-scale HTS was completed for the Cincinnati Metropolitan Area between the months of August 2009 and August 2010. GPS units were given to each household member 12 years or older to carry everywhere for 3 days. Each unit was set to collect data (second-by-second basis) and once the data collection period was finished, the devices were retrieved and data downloaded. In addition, each household and each inhabitant was requested to provide information forms regarding work and school places and the two most frequented locations. Also, they were asked about reasons for lack of personal data if any or lack of travel. Once the survey was concluded the information was located and used in the
GPS processing (OKI, 2011). Table 4 provides information and statistics regarding number of surveyed households.

**Table 4. Surveyed Household Statistics**

<table>
<thead>
<tr>
<th>Statistic</th>
<th>GPS Complete</th>
<th></th>
<th>GPS Incomplete</th>
<th></th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
<td>Percent</td>
<td>Number</td>
<td>Percent</td>
<td></td>
</tr>
<tr>
<td>Households</td>
<td>2,059</td>
<td>78.9%</td>
<td>549</td>
<td>21.1%</td>
<td>2,608</td>
</tr>
<tr>
<td>Persons</td>
<td>3,849</td>
<td>82.7%</td>
<td>807</td>
<td>17.3%</td>
<td>4,656</td>
</tr>
<tr>
<td>Travel Days</td>
<td>13,210</td>
<td>83.2%</td>
<td>2,670</td>
<td>16.8%</td>
<td>15,880</td>
</tr>
<tr>
<td>Trips</td>
<td>60,900</td>
<td>84.2%</td>
<td>11,336</td>
<td>15.8%</td>
<td>72,236</td>
</tr>
<tr>
<td>Average Daily Household Trip Rate</td>
<td>8.62</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>8.62</td>
</tr>
<tr>
<td>Average Daily Person Trip Rate</td>
<td>4.61</td>
<td>-</td>
<td>4.25</td>
<td>-</td>
<td>4.55</td>
</tr>
<tr>
<td>Average Weekday Household Trip Rate</td>
<td>9.46</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>9.46</td>
</tr>
<tr>
<td>Average Weekday Person Trip Rate</td>
<td>5.06</td>
<td>-</td>
<td>4.64</td>
<td>-</td>
<td>9.99</td>
</tr>
<tr>
<td>Average Trip Distance (all days)</td>
<td>6.11 miles</td>
<td>-</td>
<td>6.29 miles</td>
<td>-</td>
<td>6.14 miles</td>
</tr>
<tr>
<td>Average Trip Distance (weekdays)</td>
<td>6.21 miles</td>
<td>-</td>
<td>6.48 miles</td>
<td>-</td>
<td>6.25 miles</td>
</tr>
<tr>
<td>Average Trip Time (all days)</td>
<td>0:13:07</td>
<td>-</td>
<td>0:13:17</td>
<td>-</td>
<td>0:13:09</td>
</tr>
<tr>
<td>Average Trip Time (weekdays)</td>
<td>0:13:05</td>
<td>-</td>
<td>0:13:21</td>
<td>-</td>
<td>0:13:07</td>
</tr>
<tr>
<td>Average Daily Travel Time (all days)</td>
<td>01:22:11.1</td>
<td>-</td>
<td>01:19:27.1</td>
<td>-</td>
<td>01:21:44.4</td>
</tr>
<tr>
<td>Average Daily Travel Time (weekdays)</td>
<td>01:21:10.5</td>
<td>-</td>
<td>01:19:26.6</td>
<td>-</td>
<td>01:20:53.7</td>
</tr>
</tbody>
</table>

Source: OKI, 2012
3. **Methodology**

As previously stated, the data required by this research will be extracted from the database of the Ohio-Kentucky-Indiana regional council of Governments’ Household transportation survey. This data will be filtered and located on the maps according to mode of transportation and location. Then, it’ll be possible to develop a simple model to study the effect of particular variables based on infrastructure characteristics in the city of Cincinnati’s metropolitan area.

**Preliminary data**

Data for this study comes from the Greater Cincinnati Household Travel Survey developed by OKI, who joined forces with the Ohio Department of Transportation and the Kentucky Transportation Cabinet to elaborate a complete household travel survey mostly regarding trip length, trip purpose, transit usage and auto occupancy.

The goal of this survey is to collect data “regarding the movement of people throughout the region” (http://www.oki.org – accessed January 10, 2013) and update the region’s travel demand models. The last survey of its kind was conducted in 1995.

OKI has reported generalized data and travel trends on their website and has shown comparisons between current and previous results.

The process for this survey was as follows:

1. Households are surveyed through short phone and internet interviews to gather general information on households and their inhabitants.
2. Participants (or their parents if 15 years old or younger) are provided a GPS device to carry for three days.
3. Follow-up questions are addressed to some of the participants.

As part of the user’s agreement, OKI compromised to use this survey’s data in the strictest confidence. Personal information will not be released to anyone that’s not closely related to the project or its
procedures. Meaning, the public, several government agencies, employers and specially marketers will not have access to this information. The results of this survey are presented so it’s not possible to trace or identify any participants.

Table 5 shows the full list of variables compiled by the HTS:
Table 5. HTS Variable List

<table>
<thead>
<tr>
<th>HOUSEHOLD DATA</th>
<th>PERSONAL DATA</th>
<th>LOCATION</th>
<th>TRIP INFO</th>
</tr>
</thead>
<tbody>
<tr>
<td>• MPO Designation</td>
<td>• HH/Person ID of Trip Maker</td>
<td>• Location Description</td>
<td>• Departure Day of Week</td>
</tr>
<tr>
<td>• Household Status as to Transit Access Area, University, Other</td>
<td>• Relationship of PERSONID to Head of Household</td>
<td>• Location or Building Name</td>
<td>• Departure Travel Date</td>
</tr>
<tr>
<td>• Area Type of HH</td>
<td>• Gender of Trip Maker</td>
<td>• Location address</td>
<td>• Departure Time</td>
</tr>
<tr>
<td>• County (Self-Reported)</td>
<td>• Age of Tripmaker</td>
<td>• Location Cross-streets</td>
<td>• Arrival Travel Day</td>
</tr>
<tr>
<td>• Household Travel Period</td>
<td>• Age of Tripmaker by Category</td>
<td>• Location City</td>
<td>• Arrival Day of Week</td>
</tr>
<tr>
<td>• Number of Persons in Household</td>
<td></td>
<td>• Location State</td>
<td>• Arrival Travel Date</td>
</tr>
<tr>
<td>• Number of Workers in Household</td>
<td></td>
<td>• Location Zip code</td>
<td></td>
</tr>
<tr>
<td>• Number of Students in Household</td>
<td></td>
<td>• Location Latitude</td>
<td></td>
</tr>
<tr>
<td>• Number of Drivers in Household</td>
<td></td>
<td>• Location Longitude</td>
<td></td>
</tr>
<tr>
<td>• Number of Household Vehicles</td>
<td></td>
<td>• Location Longitude</td>
<td>• Trip Distance</td>
</tr>
<tr>
<td>• Household Type</td>
<td></td>
<td>• GIS Address</td>
<td>• Trip Speed</td>
</tr>
<tr>
<td>• Number of Bicycles in Household</td>
<td>• Number of Household Vehicles</td>
<td>• GIS State</td>
<td>• Trip Duration</td>
</tr>
<tr>
<td>• Household Income</td>
<td>• Number of Workers in Household</td>
<td>• GIS Zip code</td>
<td>• Place (Origin)</td>
</tr>
<tr>
<td>• Number of Household Vehicles</td>
<td>• Number of Persons in Household</td>
<td>• GIS Traffic Analysis Zone</td>
<td>• Activity (Origin)</td>
</tr>
<tr>
<td>• Number of Students in Household</td>
<td>• Number of Students in Household</td>
<td>• GIS Match Status</td>
<td>• Longitude (Origin)</td>
</tr>
<tr>
<td>• K thru 12 Student Present in Household</td>
<td></td>
<td></td>
<td>• Latitude (Origin)</td>
</tr>
<tr>
<td>• Post Secondary Student Present in Household</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Total GPS Trips Made by Household</td>
<td>• Location ID from the Location Table</td>
<td></td>
<td>• Mode of Trip</td>
</tr>
<tr>
<td>• Total Trips Made by Children in Household</td>
<td>• Confidence in GPS Trip Recorded Record</td>
<td></td>
<td>• Driver or Passenger</td>
</tr>
<tr>
<td>• Total Trips Made by Household</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Child Diary Collected for Household Child Members</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Home Address</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Latitude</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Home Address Longitude</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As it’s shown, the HTS provides very detailed information, not only about each household but also about each resident and each one of the trips each one of them made during the surveyed days.

This research is also going to use the Cincinnati Area Geographic Information Systems office (CAGIS) street infrastructure information.

CAGIS is a consortium that “provides Hamilton County’s (Cincinnati-based) agencies with up-to-date and real-time cartographic information to improve decision making processes regarding public services” (http://cagismaps.hamilton-co.org/ - accessed January 4, 2013). All this is possible, in real time, through integration of GIS technology and workflow software. Through CAGIS is possible for local government business operators develop permits and licenses to enforce codes and work orders and enhance customer service and billing systems. Up to date, CAGIS has reported to reach approximately 80% of local government and utility information through addresses and precise locations.

CAGIS manages more than 200 map information layers. All of them built in cooperation with adjunct agencies. All this information is the basis to create the agency’s framework of aid on decision making, support analysis and community planning. Topographical data is updated every five years through flyovers over Hamilton, Butler, Warren and Clermont counties.

CAGIS data provides planimetric, administrative, transactional, utility and raster data in which aerial photographs are integrated. The accuracy of this data meets the National Map Standards (NMAS) for 1”=100” base map (horizontal and vertical).

Figure 6 is a map created using county and census track information data from CAGIS. Each one of the counties of the OKI region has been identified with a different color and each one of the census tracks has been traced as a polygon. Internally, each one of them has been characterized through a database (attribute table) containing the corresponding demographic information.
Figures 7 shows an example of a street network map created using the CAGIS data. Along with the type of street attribute (highway, arterial, secondary) the corresponding data base contains geometry and design characteristics per segment.
FIGURE 6. MAP AND DATA FOLDER FROM CAGIS
FIGURE 7. CINCINNATI STREET MAP
Considering this research hypothesis and CAGIS’ available data, it has been determined that the infrastructure characteristics to be analyzed: maximum speed, number of lanes (one or two day streets) and land use.

The maximum allowed speed will help differentiate the streets where traffic moves slower. As it was presented before, maneuvering around cars is a hassle for some bikers and higher speeds and they believe that it poses a safety threat.

The number of lanes and being a one/two way street can help determine if there is enough space for bikers to maneuver around possible parked vehicles and avoid “dooring”.

Land use has been chosen because, as stated in the hypothesis, denser land uses encourage non-motorized transportation as it’s possible to complete several tasks on the same trip.

Figure 8 shows the correspondence between infrastructure characteristics from this research’s hypothesis and GIS data to be attached to the route line layer.

<table>
<thead>
<tr>
<th>Slow traffic</th>
<th>Max allowed speed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parking Lanes</td>
<td>Number of lanes + direction</td>
</tr>
<tr>
<td>Land Use</td>
<td>Land Use</td>
</tr>
</tbody>
</table>

**FIGURE 8. DATA HYPOTHESIS CORRESPONDENCE**

**Mode-based household data selection**

As is presented in table 2, information on each individual data comes with excessive information, more than this research requires. Which is an aggregate value to the process, but it can be a distracter from the main objective of this research.

Figure 9 presents a clipping, as an example of how trip data is presented in the original dataset.

Table 6 presents a clean and organized sample of the household data showing each of their ID numbers, strata, type, number of bicycles and income range. Table 7 presents the corresponding travel information per trip coded as which member of the household took the trip, date, time, type of origin.
point and activities that take place there along with the longitude and latitude coordinates of the point and other information presented on table 6. The most important attribute is the MODE of transportation used on the trip is included, for this research (Bicycle is mode = 4)

Table 6. Organized Household HTS Data (Sample)

<table>
<thead>
<tr>
<th>HHID</th>
<th>HHSTRATA</th>
<th>TYPE</th>
<th>BICY</th>
<th>INCOME</th>
</tr>
</thead>
<tbody>
<tr>
<td>100005.0</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>100019.0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>100020.0</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>100026.0</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>100029.0</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>100039.0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>100040.0</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>100049.0</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>100050.0</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>100052.0</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>100079.0</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>100081.0</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>100082.0</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>
Trip data like the one on table 7 will be filtered and only those trips that used bicycles as mode will be considered. Also, data from table 6 will be included so each trip user is described using his/her household’s characteristics.

### Mapping GoogleEarth data

The GPS data from the survey comes as *.kmz packages, a type of file which is possible to work on Google Earth. Each file is named: householdid_personid_dayofsurvey_dayoftheweek.

Figure 10 shows the original data visualized on GoogleEarth. Each color represents a trip which, by Google Earth definitions, is the one where the distance and space traveled use the same mode of transportation without “considerably long” stops. A considerably long stop is one that takes more than an average red-light signal. Google Earth defines as a TOUR as the group of trips taken by ONE person in ONE day.
FIGURE 10. GPS DATA – GOOGLE EARTH DATA
Source: HTS, OKI, 2010
Each one of the trips from the household travel individuals from the database has a corresponding Google Earth map generated by GPS data. It is an overload of data. Which it makes it necessary to use the previously filter “Bicycle-only data” to select the map files that contain bicycle trips and proceed to locate them to be able to visualize them as GIS files.

Finding HTS segments on the full street network

Route data comes in the form of “point layers” which does not match the line layers in which the street data is provided by CAGIS. In this case, it’s necessary to transfer all route information and its corresponding demographic attributes to the street layers.

Each one of the routes came as a single layer, so transferring demographic data to each one of them was done as a database management process using Microsoft Access.

Then, a short algorithm was developed using ArcMap’s model builder in which lines from the original street segment were selected based on their proximity to each one of the routes. In this case, the proximity was determined by “trial and error” to be 5 feet.

As it is shown in figures 11a and 11b, for some layers, this process selected more (or less) street segments that the corresponding user rode on. Unfortunately, it was necessary to correct these errors by hand. Fortunately, errors like the ones shown occurred in 10% of the layers.

In figure 11A, it’s shown street segments matching the route. In figure 11B, the two crossing streets selected by the query are not part of the original route.
FIGURE 11A. CORRECT STREET SEGMENT SELECTION BY ALGORITHM

FIGURE 11B. WRONG STREET SEGMENT SELECTION BY ALGORITHM
This query also left out rides that took place on parks which, again, are not part of this research because bike trails are not part of the shared street network. This was not a mistake, but an intentional process.

**Demographic data coding and incorporation**

As the Google Map route data was converted to Arc Map compatible files, each one was assigned a unique value that represented the biker that rode it. This value then used as the common attribute to joined the demographic table to the route attributes.

As a preliminary approach, 10 of these characteristics will be used as variables:

Activity based (2): type of activity taken place at a) origin and b) destination.

Rider’s personal data (3): a) gender, b) age and c) having or not a driver’s license.

Rider’s household data (5): a) household strata (closeness to a transit station or a university or none of them), b) type of area where the household is located, c) number of bicycles in household, d) household income and, e) number of vehicles in household.

Table 8 presents a compilation of the selected variables and the corresponding data set they belong to in relationship to the full data presented on the first stage of this research methodology “Preliminary data”.

<table>
<thead>
<tr>
<th>HOUSEHOLD DATA</th>
<th>PERSONAL DATA</th>
<th>TRIP INFO</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Household Status as to Transit Access Area, University, Other</td>
<td>• Gender of Trip Maker</td>
<td>• Activity (Origin)</td>
</tr>
<tr>
<td>• Area Type of HH</td>
<td>• Age of Trip maker by Category</td>
<td></td>
</tr>
<tr>
<td>• Number of Bicycles in Household</td>
<td>• Licensed Driver</td>
<td></td>
</tr>
<tr>
<td>• Household Income</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Number of Household Vehicles</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 9 shows all demographic data coding, following OKI’s system.
<table>
<thead>
<tr>
<th>TYPE OF VARIABLE</th>
<th>VARIABLE</th>
<th>VALUES</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HOUSEHOLD DATA</strong></td>
<td>Household Accessibility</td>
<td>1=Transit, 2=University, 3=Other</td>
</tr>
<tr>
<td></td>
<td>Area Type of HH</td>
<td>1=CBD, 2=Urban, 3=Suburban, 4=Rural, 97=Outside study area, 99=HH Unassigned</td>
</tr>
<tr>
<td></td>
<td>Number of Bicycles in Household</td>
<td>15=15 or more bicycles, 98/99 = DK/RF</td>
</tr>
<tr>
<td></td>
<td>Household Income</td>
<td>1=Less than $25,000, 2=$25,000 to $49,999, 3=$50,000 to $74,999, 4=$75,000 or Above, 5/6 = DK/RF</td>
</tr>
<tr>
<td></td>
<td>Vehicles in Household</td>
<td>3 = 3 or more</td>
</tr>
<tr>
<td><strong>PERSONAL DATA</strong></td>
<td>Gender of Trip Maker</td>
<td>1=Male, 2=Female, 8/9 = DK/RF</td>
</tr>
<tr>
<td></td>
<td>Age of Trip maker</td>
<td>1=Under 5, 2=5 to 12, 3=13 to 15, 4=16 to 17, 5=18 to 24, 6=25 to 34, 7=35 to 44, 8=45 to 54, 9=55 to 64, 10=65 to 74, 11=75 to 84, 12=85 and over, 998/999 = DK/RF</td>
</tr>
<tr>
<td></td>
<td>Licensed Driver</td>
<td>1=Yes, 2=No, 8/9 = DK/RF, 997=Ineligible to drive</td>
</tr>
<tr>
<td><strong>TRIP INFO</strong></td>
<td>Activity at Origin/Destination</td>
<td>1=At home, 2=Paid work, 3=School, 4=Volunteer Work, 5=Pick-Up / Drop Off Person, 6=Social, Recreational, Church, 7=Catch a Bus, Train or Airplane, 8=Transfer From One Bus, Train or Airplane to Another, 9=Shop, 10=Personal Business, 11=Eat Meal, 12=Go for a Drive, 13=Work Related, 14=School Related, 15=Other, 99=DK/RF</td>
</tr>
</tbody>
</table>
Incorporation and recoding of land use data

With the routes and their demographic data allocated on the street network (infrastructure data) it is now possible to incorporate land use data to the network.

To do this, land uses were organized by hierarchy (Lachapelle et al. 2011) and each one was assigned a numerical code as shown on table 10.

<table>
<thead>
<tr>
<th>Land use name</th>
<th>Symbol</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residential</td>
<td>RES</td>
<td>10</td>
</tr>
<tr>
<td>Commercial</td>
<td>COM</td>
<td>9</td>
</tr>
<tr>
<td>Institutional</td>
<td>INS</td>
<td>8</td>
</tr>
<tr>
<td>Agriculture</td>
<td>AGR</td>
<td>7</td>
</tr>
<tr>
<td>Industrial</td>
<td>IND</td>
<td>6</td>
</tr>
<tr>
<td>Undeveloped</td>
<td>UND</td>
<td>5</td>
</tr>
<tr>
<td>Unclassified</td>
<td>UNC</td>
<td>4</td>
</tr>
<tr>
<td>Train</td>
<td>TRN</td>
<td>3</td>
</tr>
<tr>
<td>Water body</td>
<td>OHR</td>
<td>0</td>
</tr>
<tr>
<td>Other</td>
<td>---</td>
<td>0</td>
</tr>
</tbody>
</table>

“Predisposition to biking” analysis and mapping

Predisposition to biking is measured as element that can predispose an individual to bike or not based on the preliminary demographic variables from the HTS.

This predisposition to biking will be represented as how the bikers are located according to their demographic characteristics and how much did they used the street network, represented as “density analysis”

The Density analysis tool from Arc Map takes known attribute values and extends them across the area estimating a value for each location to show where the line features are concentrated. In other words, it distributes a measured quantity throughout the landscape to produce a continuous surface (Silverman 1986).
Only the portion of a line contained within the circular area, like shown in Figure 12, centered at each grid cell is considered when calculating the density. When there aren’t any lines inside this circle, the assigned value is cero.

Larger values of the radius parameter produce a more generalized density, and smaller values result in a more detailed surface.

Figure 12 illustrates how the Density Analysis is calculated. The estimate value for each grid cell is a sum result: using a constant radius value, a circle centered in each cell is drawn to determine the study area per grid cell. Then, the length of each one of the lines contained by the circle (in this case: Lines 1 and 2) is measured (L1 and L2). After that, the attributes of each line are extracted from the data base (A1 and A2, in case) and multiplied for the corresponding length. Finally, the density at each point corresponds to the sum of all the products of lengths and attributes are divided by the area of the circle.

\[
Density = \frac{\sum_{i=1}^{n} L_i A_i}{Area\ of\ circle}
\]  

(5)
Figure 13 presents the toolbox dialog box in which:

Input polyline features: the layer to be analyzed.

Population field: the attribute to be studied. In this case, 6 of the 10 attributes were analyzed: Licensed driver, number of bicycles at households, number of vehicles at household, household income, household accessibility to transit and household location.

Output raster: the name and the location of the file with the results.

**FIGURE 13. LINE DENSITY INPUT DIALOG BOX.**
**Spatial correlation analysis**

This type of analysis estimates how correlated are the spatial features based on their location and attribute values using the Global Moran's I statistic. 

The spatial correlation tool determines how dispersed or how clustered the features in a map by calculating five values: Moran's I Index, Expected Index, Variance, z-score, and p-value. A graphic definition of clustered and dispersed data is presented in Figure 14 (ESRI 2013).

![Dispersed Features vs. Clustered Features](image)

**FIGURE 14. DISPERSED FEATURES VS. CLUSTERED FEATURES**  
Source ArcGIS resources, 2013

The Global Moran’s spatial correlation tool estimates geographical distribution based on feature location and value at the same time. Given a set of features and an associated attribute, it evaluates whether the pattern expressed is clustered, dispersed, or randomly distributed. The tool calculates the Moran's I Index value and both z- and p-scores to evaluate the significance of that Index.

This tool is an inferential statistic, meaning, the results are interpreted considering a null hypothesis. This null hypothesis states that the attribute being analyzed is randomly distributed among the features in the study area (Fotheringham et al. 2000).
This tool is used to estimate the strongest autocorrelated neighborhoods in a study area over time; for example, to map trends of cultural (racial, religious, ideological) segregation or isolation.

The Moran’s I statistic is calculated using the equation:

$$I = \frac{n \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} z_i z_j}{S_0 \sum_{i=1}^{n} z_i^2}$$  \hspace{1cm} (6)$$

Where:

- $z_i = $ deviation of an attribute for feature $i$ from its mean.
- $w_{ij} = $ spatial weight between feature $I$ and $j$,
- $n = $ total number of features,
- $S_0 = $ aggregated spatial weights.

$S_0$ is calculated using the equation:

$$S_0 = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}$$  \hspace{1cm} (7)$$

And the $z_i$-score for the statistic is calculated as:

$$z_i = \frac{I - E[I]}{\sqrt{V[I]}}$$  \hspace{1cm} (8)$$

Where:

$$E[I] = \frac{-1}{n-1}$$  \hspace{1cm} (9)$$

and

$$V[I] = E[I^2] - E[I]^2$$  \hspace{1cm} (10)$$
This particular tool estimates the mean and the variance for each attribute and then it subtracts them from the mean calculating then, a deviation from the mean.

When two attribute values of neighboring features are either larger or smaller than the mean, the result will be positive. In the mixed case: one value is smaller and the other is larger than the mean, the result will be negative. In any case, having a larger deviation from the mean will result in a larger statistic value. The index will be positive for clustered values (high values close to high values, low values close to low values); negative for features that seem to “repel” one another; and will tend to zero for those where high and low value features tend to balance negative and positive products (Gelfand 2010).

Figure 15 presents the toolbox dialog box in which:

Input feature features: the layer to be analyzed.

Input field: the attribute to be studied. In this case, the 14 attributes (10 personal + 3 street + 1 check) were analyzed at a time.

Generate report (yes/no): it will create an html life with the corresponding probability analysis results (probability bell, Moran’s index, z-score and p-value)

Conceptualization of spatial relationships: Specifies how spatial relationships among features are conceptualized. It can be determined by inverse distance (nearby features have a larger influence on the calculations), inverse distance squared, e-distance squared, fixed distance (neighboring features inside a pre-specified distance are weighted 1, all others are weighted 0), zone of indifference (objects within a specific critical distance are weighted as 1, the rest diminish with distance), contiguity edges only, contiguity edges by corners or, spatial weights can be personalized by the user. In this case, the selected method was INVERSE DISTANCE.

Distance Method: Specifies how distances are calculated from each feature to feature. It can be estimated as Euclidean or Manhattan (meaning, straight line or along x and y axes).
As it can be seen, the index is normalized by the variance so the values fall in a -1.0 and +1.0 range.

The Expected and Observed Indexes are compared to estimate a z- and a p-scores to be able to indicate if the difference is statistically significant or not (GETIS and ORD 1992).

This statistical significance is explained on Table 11:

<table>
<thead>
<tr>
<th>The p-value is not statistically significant.</th>
<th>The null hypothesis cannot be rejected. Quite possibly, the spatial distribution of feature values is the result of random spatial processes.</th>
</tr>
</thead>
<tbody>
<tr>
<td>The p-value is statistically significant, and the z-score is positive.</td>
<td>The null hypothesis is rejected. The distribution of high and low values in the dataset is more spatially clustered than would be expected if related spatial processes were random.</td>
</tr>
<tr>
<td>The p-value is statistically significant, and the z-score is negative.</td>
<td>The null hypothesis is rejected. The distribution of high and low values in the dataset is more spatially dispersed than would be expected if related spatial processes were random.</td>
</tr>
</tbody>
</table>
Ordinary least squares

A regression analysis evaluates relationships between attributes of a feature making it possible to predict where and when an event is likely to occur; or, as an inverse process, examine causes of events to better understand why they occur the way they do (Mitchell and Environmental Systems Research Institute (Redlands, Calif.) 1999).

Ordinary Least Squares (OLS) the most frequently used regression technique. For GIS, it is the starting point for all spatial regression analyses because it develops a model of the variable or process under study creating a single regression equation to represent it. The full process is explained in Figure 16.

**FIGURE 16. ORDINARY LEAST SQUARES FLOWCHART**
Source: ArcGIS resources, 2013
A model’s performance is evaluated by a multiple R-squared and adjusted R-squared values which range from 0.0 to 1.0. Adding additional variables to a model will result in a higher Multiple R-squared value but a lower Adjusted R-Squared value because it reflects the complexity of the model.

The OLS analysis will also calculate items like the equation coefficient for each featured (which will now be a variable), Robust Probability, and Variance Inflation Factor (VIF).

A T-test value is estimated to determine if an explanatory variable is statistically significant, considering a null hypothesis in which the coefficient is equal to zero. When the p-value is considerably small the chances of the coefficient being zero are also small.

The Variance Inflation Factor (VIF) for every variable quantifies such variable’s multicollinearity by estimating an index that measures how much does its variance increases due to it.

The VIF for each i-variable is calculated as:

\[
\text{VIF} = \frac{1}{1-R_i^2} \tag{11}
\]

Where \(R_i^2\) is the coefficient of determination of the regression equation in which the i-variable becomes the independent term and the rest of the variables define it.

Some authors define high levels of multicollinearity with the VIF is higher than 5 (Allison 1999). From this point, others have defined that 7.5 is a multicollinearity level in which the variable should be removed from the equation and estimate the effect of this removal; the most probable explanation is that this variable’s contribution to it is represented by another variable. It has been proposed as well, that a VIF value of 10 is a cut off value (Hair 2006).

Figure 17 presents the toolbox dialog box in which:
Input feature features: the layer to be analyzed.

Unique OD field: a particular identificator for each one of the elements in the layer.

Output feature class: the name and destination of the file to be created with the results.

Dependent variable: the numeric field with the values to be modeled.

Explanatory variables: a list of the fields from the attribute table that will explain the dependent variable and will be accompanied by coefficients.

FIGURE 17. ORDINARY LEAST SQUARES INPUT DIALOG BOX
Variable selection and filtering

Considering all multicolinearity standards (R², p-value, z-value, robusticity index, Moran’s statistic), it is now possible to reject or accept each one of the suggested variables and iterate the spatial correlation analysis to obtain the optimum combination of variables and their corresponding coefficients towards defining compatibility between bikers and streets.

Final results

The result of this process will be a Bicycle Compatibility Evaluator (BCE) applicable to every street segment of the OKI area: Hamilton, Clermont, Butler, Warren (Ohio), Boone, Kenton, Campbell (Kentucky) and Dearborn (Indiana):

\[ BCE = \sum_{i=1}^{n} \alpha_i D_i + \sum_{j=1}^{n} \beta_j S_j \]  

(12)

Where \( \alpha \) and \( \beta \) are the corresponding estimated coefficients for each one of the proposed Demographic (D) and Street (S) characteristics (table 12):

<table>
<thead>
<tr>
<th>Table 12. Variable Coding</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Demographic (D) i-values</strong></td>
</tr>
<tr>
<td>1 = Rider’s gender</td>
</tr>
<tr>
<td>2 = Rider’s age</td>
</tr>
<tr>
<td>3 = Rider is a licensed driver</td>
</tr>
<tr>
<td>4 = Rider’s activity at origin</td>
</tr>
<tr>
<td>5 = Rider’s activity at destination</td>
</tr>
<tr>
<td>6 = Number of bicycles at rider’s HH</td>
</tr>
<tr>
<td>7 = Number of motor vehicles at rider’s HH</td>
</tr>
<tr>
<td>8 = Rider’s HH income</td>
</tr>
<tr>
<td>9 = Rider’s HH accessibility</td>
</tr>
<tr>
<td>10 = Rider’s HH location</td>
</tr>
</tbody>
</table>

As it has been noted, it is possible that not all variables are included in the equation due to their significance in describing the relationship between bicyclists and streets.
Methodology schematics

Figure 18 shows this research methodology:
4. Data Preparation

Collected data

As stated before, data from the Cincinnati Metropolitan area collected by OKI has been collected and it’s now ready to be analyzed. Collected data has been presented on table 7.

The HTS provides very detailed information, not only about each household but also about each resident and each one of the trips each one of them made during the surveyed days.

Every variable value collected on the HTS was been coded accordingly to a “Data Dictionary” that has been provided along with all datasets. The following tables explain where the pertaining variables are found and their coding.

Table 13 shows this research’s selected variables from the HTS survey related to household characteristics. The HHID is the ID number assigned to each one of the participating households, it is a numeric variable. HHSTRATA represents the “strata: or relative location of the household respecting to a Transit station or a University, any other location will be coded as “other”. AREATYPE is the type of area in which the household is located: as within the Central Business Distric (CBD, or Downtown area in this case), urban, suburban or rural areas. Household located outside the study area were also accounted for. TOTVEH is the number of vehicles own by the household members, it’s a discreet variable in which the upper range is 10, any other value higher than this will be coded as 10 as well. BICYC is the number of bicycles owned by the members of the household and, just like the number of vehicles is a discreet variable, with a maximum value of 15. INCOME is measured as value ranges from 1 to 4 and two additional 5 and 6 for those households that did know or didn’t wanted to provide such information.
### Table 13. Data Dictionary Characteristics for Household Data from the HTS

<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
<th>Values / Notes</th>
<th>Record Type</th>
<th>Type</th>
<th>Data Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>HHID</td>
<td>Household ID Number</td>
<td>Numeric Variable</td>
<td>Household Record</td>
<td>Assigned</td>
<td>Assigned</td>
</tr>
<tr>
<td>HHSTRATA</td>
<td>Household Status as to Transit Access Area, University, Other</td>
<td>1=Transit 2=University 3=Other</td>
<td>Household Record</td>
<td>Sample</td>
<td>Sampe</td>
</tr>
<tr>
<td>AREATYPE</td>
<td>Area Type of HH</td>
<td>1=CBD 2=Urban 3=Suburban 4=Rural 97=Outside study area 99=HH Unassigned</td>
<td>Household Record</td>
<td>Sample</td>
<td>GIS Based</td>
</tr>
<tr>
<td>TOTVEH</td>
<td>Number of Household Vehicles</td>
<td>10=10 or more vehicles</td>
<td>Household Record</td>
<td>Survey Response</td>
<td>Recruitment</td>
</tr>
<tr>
<td>BICYC</td>
<td>Number of Bicycles in Household</td>
<td>15-15 or more 98=Don't Know 99=Refused</td>
<td>Household Record</td>
<td>Survey Response</td>
<td>Recruitment</td>
</tr>
<tr>
<td>INCOME</td>
<td>Household Income</td>
<td>1=Less than $25,000 2=$25,000 to $49,999 3=$50,000 to $74,999 4=$75,000 or Above 5=Don't know 6=Refused</td>
<td>Household Record</td>
<td>Survey Response</td>
<td>Recruitment</td>
</tr>
</tbody>
</table>

Source: OKI, 2010

Table 14 contains the selected variables from the HTS survey related to user personal characteristics. Just like the previous one, HHID is the ID number assigned to each one of the participating households, it is a numeric variable. PERSONID is the unique identifier for each one of the members of the household as a numeric variable. HHPERSONID is a combined code to identify each one of the users in the entire survey, containing the HHID where they belong and their personal identifier. GENDER represents the gender of the user, including two classes: 8 = don’t know and 9 = refused to provide the information. As it has been stated before, one of the conditions of usage of this data is to preserve and respect participants’ anonymity. AGE2 is the age of the user fitted into 12 categories; again, the user was presented with the possibility of not replying (998) or refusing to do it (999).
Table 14. Data Dictionary Characteristics For Personal Data From The HTS

<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
<th>Values / Notes</th>
<th>Record Type</th>
<th>Type</th>
<th>Data Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>HHID</td>
<td>Household ID Number</td>
<td>Numeric Variable</td>
<td>Household Record</td>
<td>Assigned</td>
<td>Assigned</td>
</tr>
<tr>
<td>PERSONID</td>
<td>Person ID</td>
<td>Numeric Variable</td>
<td>Person Record</td>
<td>Assigned</td>
<td>Assigned</td>
</tr>
<tr>
<td>HHPERSONID</td>
<td>HH/Person ID of Trip Maker</td>
<td>HHID * 1000 + PERSONID</td>
<td>Person Record</td>
<td>Assigned</td>
<td>Assigned</td>
</tr>
<tr>
<td>GENDER</td>
<td>Gender of Trip Maker (PERSONID)</td>
<td>1=Male 2=Female 8=Don't know (VOL)</td>
<td>Person Record</td>
<td>Survey Response</td>
<td>Recruitment</td>
</tr>
<tr>
<td>AGE2</td>
<td>Age of Tripmaker by Category</td>
<td>1=Under 5 2=5 to 12 3=13 to 15 4=16 to 17 5=18 to 24 6=25 to 34 7=35 to 44 8=45 to 54 9=55 to 64 10=65 to 74 11=75 to 84 12=85 and over 998=Don't know 999=Refused</td>
<td>Person Record</td>
<td>Calculated</td>
<td>Recruitment</td>
</tr>
<tr>
<td>LICENSED</td>
<td>Licensed Driver</td>
<td>997= Cannot drive</td>
<td>Person Record</td>
<td>Response</td>
<td>Recruitment</td>
</tr>
</tbody>
</table>

Source: OKI, 2010

Finally, table 15 presents the variables that have been selected to study the trips, HHID and PERSONID will help link each trip to the corresponding household and participant (hence, economic and demographic data). The travel day helps make each “tour” unique, considering how each person was asked to collect data for three days. O_ACT and D_ACT are the identifiers for the activities that take place at the origin and destination points respectively. The most important piece of data in this table is the MODE of trip, in which this research’s focus will be code number 4. Based on this table, only households that rode bicycles on any of the days of the study will be considered, and only users that rode a bicycle on any of those three days will be accounted for in this study.
<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
<th>Values / Notes</th>
<th>Record Type</th>
<th>Agency</th>
<th>Type</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>HHID</td>
<td>Household ID Number</td>
<td>Numeric Variable</td>
<td>Household Record</td>
<td>Abt SRBI</td>
<td>Assigned</td>
<td>Assigned</td>
</tr>
<tr>
<td>PERSONID</td>
<td>Number of Household Member</td>
<td>Numeric Variable (Range 1-15)</td>
<td>Person Record</td>
<td>Abt SRBI</td>
<td>Assigned</td>
<td>Assigned</td>
</tr>
<tr>
<td>Travel Day</td>
<td>1=Day 2=Day 3=Day</td>
<td>HHID * 1000 + PERSONID from HH * 100 + TRIPSEQ + Day</td>
<td>Trip Record</td>
<td>PlanTrans</td>
<td>Assigned</td>
<td>Assigned</td>
</tr>
<tr>
<td>O_ACT</td>
<td>Activity (Origin)</td>
<td>1=At home 2=Paid work 3=School 4=Volunteer Work 5=Pick-Up / Drop Off Person 6=Social, Recreational, Church 7=Catch a Bus, Train or Airplane 8=Transfer 9=Shop 10=Personal Business 11=Eat Meal 12=Go for a Drive 13=Work Related 14=School Related 15=Other 99=DK/RF</td>
<td>Trip Record</td>
<td>PlanTrans</td>
<td>From Prompted Recall Only</td>
<td>From Prompted Recall Only</td>
</tr>
<tr>
<td>MODE</td>
<td>Mode of Trip</td>
<td>1=Motor Vehicle 2=Bus 3=Walk 4=Bicycle 5=Driver of Auto/van/truck 6=Passenger 7=Driver of Carpool 8=Passenger of Carpool 9=Driver of Vanpool 10=Passenger of Vanpool 11=Bus 12=Demand Response Bus 13=School Bus 14= Taxi / paid limo 15=Motorcycle/Moped 96=Other 98=Unknown</td>
<td>Trip Record</td>
<td>PlanTrans</td>
<td>Imputed from GPS and From Prompted Recall</td>
<td>Imputed from GPS and Prompted Recall</td>
</tr>
<tr>
<td>D_ACT</td>
<td>Activity (Destination)</td>
<td>1=At home 2=Paid work 3=School 4=Volunteer Work 5=Pick-Up / Drop Off Person 6=Social, Recreational, Church 7=Catch a Bus, Train or Airplane 8=Transfer 9=Shop 10=Personal Business 11=Eat Meal 12=Go for a Drive 13=Work Related 14=School Related 15=Other 99=DK/RF</td>
<td>Trip Record</td>
<td>PlanTrans</td>
<td>From Prompted Recall Only</td>
<td>From Prompted Recall Only</td>
</tr>
</tbody>
</table>

Source: OKI, 2010
GIS data comes from the shared data information agreement between the University of Cincinnati and CAGIS. Data is available for educational purposes.

**Bicycle-only data**

Using the codes from table 15, all data from the trips is filtered so only trips that used a bicycle will be taken in consideration.

Once this data is prepared, the corresponding HHID and PERSONID from each one of the trips is used to select the economic and demographic information from tables 13 and 14 correspondingly.

With data from tables 13, 14 and 15 filtered; isolated bicycle data, it’s possible to create a mega data base as a single one where each trip data contains demographic information from the user that took the trip and economic from the household to where this user belongs to.

**Visible bicycle routes**

The Google Earth maps collection from the Household Travel Survey represents each one of the trips, which is why it’s necessary to select just the ones that used a bicycle on a particular day. Just like the data was selected, those maps must be selected/filtered as well and presented as the example on figure 7 on the previous chapter.

As it was explained before, each route is currently composed of a series of points (dots) that represent the geographical point where the rider was at every second of his/her ride (given by the GPS’ precision). To be able to analyze them as part of the street network, it’s necessary to convert them into segments.

This process was done by selecting the street segment within a radius of 3-feet from each of the points, as shown in figure 10 (previous chapter).
The result was a series of GIS layers, in which every one of them corresponds to a trip. As an example, a zoomed section of the city network showing two of these routes is presented on figures 19A and 19B.

FIGURE 19A. ROUTE TAKEN BY SUBJECT 101214001
Segments to be evaluated

Once all the routes were converted into segments, it was possible to create a single layer, containing all the routes from the HTS.

It is important to emphasize that each one of them has a unique identifying attribute related to the rider that will later be used to incorporate its corresponding demographic information.

Figure 20 shows the entire HTS bicycle data overlapped on the full OKI street network for better orientation.
FIGURE 20. RIDDEN ROUTES MAP OVER OKI'S STREET NETWORK
**HTS information incorporated into full OKI street network**

Once all the routes were converted into segments, it was possible to create a single layer, containing all the routes from the HTS. The goal to doing this is to compare those segments that were actually ridden and those that were not used.

In addition to the infrastructure characteristics already present in the layer (maximum speed, one/two ways) each one of the demographic attributes from the bicycle routes will become a new attribute for the street layer. Segments used by a biker will contain all the “personal information” from that person and those that were not used during the time of the survey will have CERO as the attribute values on those new fields.

At this point, it’s possible to “ask” the OKI map to show streets that were ridden by a specific demographic group based on their location.
FIGURE 21. RIDDEN ROUTES OVER OKI'S STREET NETWORK (MALE/FEMALE)
As an example, Figure 21 presents all the street segments and shows which ones were ridden by women and which were ridden by men.

**OKI network with full HTS information and land use data.**

The last step in the data incorporation process is adding land use data to the all the entire street network. The reason for doing this to the full network is because this research considers land use to be a permanent characteristic of each segment; unlike the presence/absence of a biker is something temporary and it’s not shared by all segments on the network.

It is important to note that many, if not all, land use zones, are separated by street segments. Land use assignment will be done gradually considering the hierarchy established in table 10 in the previous chapter. This way, there will be a strategic order to assign land use numerical codes to streets that in close vicinity of land use zone boundaries and there won’t be any biased information.

As an example, figure 25 shows the street segments “within or close vicinity” to the coded areas.

In addition, as a control variable, a new attribute will be added to the full layer, it’s called CHECK. This is a binomial variable that will serve as a control variable assigning a value of ONE to those street segments used in during the survey and ZERO to those that were not used. As a result, this value will come to the BCEQ equation as the breaking point of compatibility between streets and bicyclists, as shown in Figure 23.
FIGURE 22. OKI'S STREET NETWORK SELECTED BY LAND USE
5. Data Analysis

Mapped density on “Predisposition to bike”

As it was explained in the previous chapter, bicyclists’ accessibility to certain items or amenities makes them more or less predisposed to ride a bicycle.

For example, having one or more bicycles in his/her house and may be inviting to using it to commute or for short trips; not having a valid driver’s license will force the person to seek alternative modes of transportation (in this case, a bicycle). Living in an area close to a transit center will make it easier to move around the area in a multimodal way; unlike living in a suburban area, in where access to public transportation or sharrows are non-existing make it necessary to ride a car.

The density maps show areas where there is a higher concentration of riders based on their characteristics. Figures 24A, B, C, D, E and F are the density maps elaborated for this research for each one of the personal characteristic sets of data. Darker coloration represents a higher density.

Figure 24G presents the density map for the CHECK variable.
FIGURE 24C. DENSITY ANALYSIS – HH VEHICLES
FIGURE 24D. DENSITY ANALYSIS – HH INCOME
FIGURE 24F. DENSITY ANALYSIS – HH TYPE
FIGURE 24G. DENSITY ANALYSIS – CHECK
It can be seen from the density maps some specific corridors where bicyclists from the HTS prefer to ride: Downtown Cincinnati, Glenway Ave, Colerain Ave, Beechmont Ave, Montgomery Rd and Alexandria Rd.

Also, there are some specific areas in which the surveyed bikers prefer to ride: Downtown Cincinnati, Union, Colerain, Woodlawn, Madison, Miami, Miller (IN), Oxford, Township,

Most of these locations have a considerable number of bikers, and are interconnected with one another through the corridors, but there are some others like Mason and Miami that are completely isolated, not too dense in terms of bicyclists but there is the presence of, which leads to think that the interconnection between bicycle friendly areas may lead to more bicycle users.

There are four well-defined areas in which the bikers who rode are also licensed drivers: Springdale, Downtown, Union and Goshen.

The areas where bikers who own their own bicycles ride are Oxford, Mason, Union and Downtown Cincinnati.

The concentration areas ridden by surveyed bikers who live in a household where there is at least ONE vehicle is very (if not completely) similar to the one defined by general ridership.

Route ridership distribution considering number of vehicles and income are practically the same, which leads to believe that vehicle ownership increases as the household’s income increases.

The density maps for accessibility and location are very similar, with the slight difference that there is an increase of ridership on streets that serve as connectors from riders that come from areas well connected.
Spatial correlation analysis results

Table 16 shows the compiled results of the performed spatial correlation analysis.

<table>
<thead>
<tr>
<th>Variable Type</th>
<th>Name</th>
<th>Moran’s index</th>
<th>z-score</th>
<th>p-value</th>
<th>Clustered?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal</td>
<td>Gender</td>
<td>0.201508</td>
<td>18.528359</td>
<td>0.000024</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Age</td>
<td>0.224363</td>
<td>20.631039</td>
<td>0.000012</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Licensed driver</td>
<td>0.147912</td>
<td>13.615722</td>
<td>0.000021</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Act @ origin</td>
<td>0.218481</td>
<td>20.088930</td>
<td>0.000005</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Act @ destination</td>
<td>0.275556</td>
<td>25.333720</td>
<td>0.000015</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Bikes on HH</td>
<td>0.0300585</td>
<td>2.884725</td>
<td>0.000018</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Vehicles on HH</td>
<td>0.306364</td>
<td>28.165859</td>
<td>0.000010</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>HH income</td>
<td>0.416620</td>
<td>38.297249</td>
<td>0.000006</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>HH accessibility</td>
<td>0.037597</td>
<td>3.468464</td>
<td>0.000010</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>HH location</td>
<td>0.542789</td>
<td>49.892005</td>
<td>0.000050</td>
<td>Yes</td>
</tr>
<tr>
<td>Street</td>
<td>Speed limit</td>
<td>0.462857</td>
<td>42.545184</td>
<td>0.000008</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>One/two way</td>
<td>0.647561</td>
<td>59.520112</td>
<td>0.000023</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Land Use</td>
<td>0.717769</td>
<td>76.132981</td>
<td>0.000010</td>
<td>Yes</td>
</tr>
</tbody>
</table>

The preliminary analysis (map density) and the statistical analysis seem to point that the selected variables are clustered. The low p-values indicate significance.
6. Results

Ordinary Least squares

Giving an “ideal” BCE value of 100, in which particular individuals and the streets they rode on match is possible to estimate those coefficients. Table 17 presents the first OLS analysis, with ALL suggested variables.

Filtered variables

From these results, considering VIF values and, as recommended by the readings, variables with VIF lower than 7.5 are excluded from the model/equation. In this particular case: GENDER, AGE, LOCATION and ACCESSIBILITY, VEHICLES PER HOUSEHOLD and HOUSEHOLD INCOME.

A second attempt is made considering the remaining variables and the results are, as presented on table 18, where all the variables meet the VIF requirements, being less than 4.5. In this case, the variable SPEEDLIMIT has a high p-value (0.177501) and it should be removed from the equation. The corresponding Multiple R² for this combination is 0.819189

A third attempt is made without the SPEEDLIMIT variable. The results are presented on table 19:

In this combination, the entire variable group meets the VIF requirements, all of them are less than 4.5, all probabilities and robust index are below 0.05 which makes them significant. The corresponding Multiple R² for this combination is 0.819185.

This R value is a 0.0001% lower than the previously estimated, but still is acceptable.

Additional variable combinations were made with lower estimated Multiple-R² results.
Table 17. OLS Analysis Results.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>StdError</th>
<th>t-Statistic</th>
<th>Probability</th>
<th>Robust_SE</th>
<th>Robust_t</th>
<th>Robust_Pr</th>
<th>VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.305286</td>
<td>0.037491</td>
<td>8.142936</td>
<td>0.000000</td>
<td>0.035468</td>
<td>8.607272</td>
<td>0.000000</td>
<td>------</td>
</tr>
<tr>
<td>SPEEDLIMIT</td>
<td>-0.004348</td>
<td>0.000901</td>
<td>-4.826974</td>
<td>0.000002</td>
<td>0.000876</td>
<td>-4.965699</td>
<td>0.000001</td>
<td>1.265364</td>
</tr>
<tr>
<td>ONEWAY</td>
<td>-0.001765</td>
<td>0.000190</td>
<td>-9.271080</td>
<td>0.000000</td>
<td>0.000207</td>
<td>-8.511912</td>
<td>0.000000</td>
<td>1.272383</td>
</tr>
<tr>
<td>OACT</td>
<td>0.357602</td>
<td>0.010696</td>
<td>33.432279</td>
<td>0.000000</td>
<td>0.075910</td>
<td>4.710886</td>
<td>0.000004</td>
<td>4.211427</td>
</tr>
<tr>
<td>DACT</td>
<td>-0.075156</td>
<td>0.010844</td>
<td>-6.930908</td>
<td>0.000000</td>
<td>0.063190</td>
<td>-1.189382</td>
<td>0.234293</td>
<td>3.733090</td>
</tr>
<tr>
<td>GENDER</td>
<td>8.497419</td>
<td>0.093851</td>
<td>90.541369</td>
<td>0.000000</td>
<td>0.550784</td>
<td>15.427866</td>
<td>0.000000</td>
<td>9.631773</td>
</tr>
<tr>
<td>AGE</td>
<td>5.057577</td>
<td>0.029364</td>
<td>172.239827</td>
<td>0.000000</td>
<td>0.216816</td>
<td>23.326536</td>
<td>0.000000</td>
<td>23.336287</td>
</tr>
<tr>
<td>LICENSE</td>
<td>0.022423</td>
<td>0.000258</td>
<td>86.986661</td>
<td>0.000000</td>
<td>0.001201</td>
<td>18.663709</td>
<td>0.000000</td>
<td>1.694070</td>
</tr>
<tr>
<td>STRATA</td>
<td>8.502260</td>
<td>0.076281</td>
<td>111.460191</td>
<td>0.000000</td>
<td>0.680683</td>
<td>12.490783</td>
<td>0.000000</td>
<td>19.479814</td>
</tr>
<tr>
<td>AREATYPE</td>
<td>6.726030</td>
<td>0.092619</td>
<td>72.620634</td>
<td>0.000000</td>
<td>0.552752</td>
<td>12.168269</td>
<td>0.000000</td>
<td>33.061435</td>
</tr>
<tr>
<td>HHVEH</td>
<td>-0.182653</td>
<td>0.066550</td>
<td>-2.744616</td>
<td>0.006060</td>
<td>0.476302</td>
<td>-0.383482</td>
<td>0.701379</td>
<td>9.628358</td>
</tr>
<tr>
<td>HHBIKES</td>
<td>0.180501</td>
<td>0.005725</td>
<td>31.526954</td>
<td>0.000000</td>
<td>0.028078</td>
<td>6.428452</td>
<td>0.000000</td>
<td>1.243316</td>
</tr>
<tr>
<td>HHINCOME</td>
<td>0.599414</td>
<td>0.041392</td>
<td>14.481428</td>
<td>0.000000</td>
<td>0.259869</td>
<td>2.306598</td>
<td>0.021063</td>
<td>8.968970</td>
</tr>
<tr>
<td>LU_CODE</td>
<td>0.000911</td>
<td>0.001697</td>
<td>0.537017</td>
<td>0.591267</td>
<td>0.001255</td>
<td>0.726423</td>
<td>0.467573</td>
<td>1.004208</td>
</tr>
</tbody>
</table>
Table 18. OLS Analysis Results. First Filter

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>StdError</th>
<th>t-Statistic</th>
<th>Probability</th>
<th>Robust_SE</th>
<th>Robust_t</th>
<th>Robust_Pr</th>
<th>VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.957504</td>
<td>0.112310</td>
<td>8.525582</td>
<td>0.000000</td>
<td>0.106973</td>
<td>8.950903</td>
<td>0.000000</td>
<td>----</td>
</tr>
<tr>
<td>SPEEDLIMIT</td>
<td>-0.003639</td>
<td>0.002698</td>
<td>-1.348535</td>
<td>0.177501</td>
<td>0.002668</td>
<td>-1.363887</td>
<td>0.172619</td>
<td>1.263488</td>
</tr>
<tr>
<td>ONEWAY</td>
<td>-0.006815</td>
<td>0.000570</td>
<td>-11.949446</td>
<td>0.000000</td>
<td>0.000637</td>
<td>-10.695604</td>
<td>0.000000</td>
<td>1.270821</td>
</tr>
<tr>
<td>OACT</td>
<td>5.457290</td>
<td>0.024745</td>
<td>220.536714</td>
<td>0.000000</td>
<td>0.105384</td>
<td>51.784804</td>
<td>0.000000</td>
<td>2.508175</td>
</tr>
<tr>
<td>DACT</td>
<td>4.805704</td>
<td>0.027194</td>
<td>176.718851</td>
<td>0.000000</td>
<td>0.110286</td>
<td>43.574782</td>
<td>0.000000</td>
<td>2.612578</td>
</tr>
<tr>
<td>LICENSE</td>
<td>0.042091</td>
<td>0.000607</td>
<td>69.350427</td>
<td>0.000000</td>
<td>0.003218</td>
<td>13.080373</td>
<td>0.000000</td>
<td>1.045040</td>
</tr>
<tr>
<td>HHBIKES</td>
<td>1.199392</td>
<td>0.016278</td>
<td>73.682118</td>
<td>0.000000</td>
<td>0.154680</td>
<td>7.753999</td>
<td>0.000000</td>
<td>1.118378</td>
</tr>
<tr>
<td>LU_CODE</td>
<td>0.017049</td>
<td>0.005087</td>
<td>3.351446</td>
<td>0.000821</td>
<td>0.003502</td>
<td>4.868267</td>
<td>0.000002</td>
<td>1.003798</td>
</tr>
</tbody>
</table>

Table 19. OLS Analysis Results. Second Filter

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>StdError</th>
<th>t-Statistic</th>
<th>Probability</th>
<th>Robust_SE</th>
<th>Robust_t</th>
<th>Robust_Pr</th>
<th>VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.823997</td>
<td>0.053030</td>
<td>15.538288</td>
<td>0.000000</td>
<td>0.049037</td>
<td>16.803487</td>
<td>0.000000</td>
<td>----</td>
</tr>
<tr>
<td>ONEWAY</td>
<td>-0.006469</td>
<td>0.000510</td>
<td>-12.696774</td>
<td>0.000000</td>
<td>0.000597</td>
<td>-10.833647</td>
<td>0.000000</td>
<td>1.014380</td>
</tr>
<tr>
<td>OACT</td>
<td>5.457174</td>
<td>0.024745</td>
<td>220.532365</td>
<td>0.000000</td>
<td>0.105382</td>
<td>51.784565</td>
<td>0.000000</td>
<td>2.508145</td>
</tr>
<tr>
<td>DACT</td>
<td>4.805095</td>
<td>0.027190</td>
<td>176.720002</td>
<td>0.000000</td>
<td>0.110320</td>
<td>43.556053</td>
<td>0.000000</td>
<td>2.611859</td>
</tr>
<tr>
<td>LICENSE</td>
<td>0.042061</td>
<td>0.000607</td>
<td>69.346856</td>
<td>0.000000</td>
<td>0.003216</td>
<td>13.080091</td>
<td>0.000000</td>
<td>1.043664</td>
</tr>
<tr>
<td>HHBIKES</td>
<td>1.199491</td>
<td>0.016278</td>
<td>73.688598</td>
<td>0.000000</td>
<td>0.154668</td>
<td>7.755272</td>
<td>0.000000</td>
<td>1.118355</td>
</tr>
<tr>
<td>LU_CODE</td>
<td>0.016821</td>
<td>0.005084</td>
<td>3.308423</td>
<td>0.000955</td>
<td>0.003484</td>
<td>4.827541</td>
<td>0.000002</td>
<td>1.002689</td>
</tr>
</tbody>
</table>
**BCE equation coefficients**

The result of this research is the Bicycle Compatibility Evaluator (BCE) applicable to every street segment of the OKI area: Hamilton, Clermont, Butler, Warren (Ohio), Boone, Kenton, Campbell (Kentucky) and Dearborn (Indiana):

\[
BCE = \sum_{i=1}^{n} \alpha_i D_i + \sum_{j=1}^{n} \beta_j S_j
\]  

(13)

Where \(\alpha\) and \(\beta\) are the corresponding estimated coefficients for each one of the proposed Demographic (D) and Street (S) characteristics, presented in tables 20A and 20B.

**Table 20A. Demographic Variables and Their Coefficients**

<table>
<thead>
<tr>
<th>Demographic i-values</th>
<th>(\alpha_i)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 = Rider is a licensed driver</td>
<td>0.042061</td>
</tr>
<tr>
<td>2 = Rider’s activity at origin</td>
<td>5.457174</td>
</tr>
<tr>
<td>3 = Rider’s activity at destination</td>
<td>4.805095</td>
</tr>
<tr>
<td>4 = Number of bicycles at rider’s HH</td>
<td>1.199494</td>
</tr>
<tr>
<td>5 = Rider’s gender</td>
<td>0.0</td>
</tr>
<tr>
<td>6 = Rider’s age</td>
<td>0.0</td>
</tr>
<tr>
<td>7 = Household location</td>
<td>0.0</td>
</tr>
<tr>
<td>8 = Household accessibility to network</td>
<td>0.0</td>
</tr>
<tr>
<td>9 = Number of vehicles at rider’s HH</td>
<td>0.0</td>
</tr>
<tr>
<td>10 = Household Income</td>
<td>0.0</td>
</tr>
</tbody>
</table>

**Table 20B. Street Variables and Their Coefficients**

<table>
<thead>
<tr>
<th>Street j-values</th>
<th>(\beta_j)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 = One/two way street</td>
<td>-0.006469</td>
</tr>
<tr>
<td>2 = Land Use</td>
<td>0.016821</td>
</tr>
<tr>
<td>3 = Speed Limit</td>
<td>0.0</td>
</tr>
</tbody>
</table>

With an intercept value of 0.823997 the equation becomes:

\[
BCE = 0.042061D_1 + 5.457174D_2 + 4.805095D_3 + 1.199494D_4 - 0.006469S_1 + 0.016821S_2 + 0.823997
\]  

(14)
It can be seen that the activities at origin and destination have the most considerable impact in the biker’s decision to ride on a certain street.

Despite public belief that most cyclists do not have a driver’s license, according to this research results, even though the impact is not considerable, it is a positive one.

The model (equation) has been tested by applying it to the original data. Figures 25A and 25B are a map of the estimated values and its corresponding Histogram. In the map, the closest to cero values are represented by lighter colors and those close to 100 (maximum value) are darker. The Histogram (25 bins) will help visualize the distribution of the values.

Lastly, it can be seen that land use has a positive impact in ridership and two-way streets have a negative impact, which matches the literature review.

From the distribution graphics and the maps, the results show that when pairing streets and groups of people (infrastructure and land use AND demographics), when the equation results

Figures 26A and 26B are the map and Histogram of all residuals from the BCE equation to evaluate discrepancies between the original and estimated data.
FIGURE 25A. ESTIMATED BCE - MAP
FIGURE 25B. ESTIMATED BCE – HISTOGRAM
FIGURE 26B. BCE RESIDUALS – HISTOGRAM
**BCE validation**

It has been proven that and demographic backgrounds are decisive factors in the bicycle route choosing process. In the particular case of the City of Cincinnati and the series of data used for this project, there isn’t sufficient information to determine if the individual economic background of the rider has a strong influence in the process.

Bicycle riders from the dataset prefer one-way streets despite the speed limit, again, according to the existing data.

According to tabulated results, the maximum estimated value from streets and individuals that don’t match is 0.996745. 100% of mismatched combinations are described by and estimated less than 1.00 BCE values.

The minimum estimated value for streets and individuals who rode them is 10.87659 and the maximum is 169.764598. Figure 27 shows the percent distribution of matching values. It can be seen that “all but 75%” of matching values are described by estimated values less than 55.00, “all but 50%” of matching values are described by estimated values less than 90.00; and “all but 29%” matching values are described by estimated values less than 100.00.

It will be stated than an estimated BCE value higher than 10.00 will represent a match between a person and the street that person can potentially ride on.
FIGURE 27. BCE VALIDATION – INVERSE CUMULATIVE FREQUENCIES
7. Conclusions

The estimated Bicycle Compatibility Evaluator for the City of Cincinnati is the first approximation to a more elaborated tool, considering the sample size and the distribution of the trips.

By working with a more elaborated and detailed household survey it will be possible to validate another variables in a more precise manner. The more available data, the more precise the BCE will be. Unfortunately, the distribution of the trips around the area and the none-frequency of those did not allowed to include a factor or “repetition of usage”. Each one of the streets was used once, or twice by the same rider, which didn’t add to the purpose of this research.

It should be specified that the variables used in this particular case are not the only once to be used on a general case. Each metropolitan area is a particular case study and according to its punctual characteristics (topography, weather, bicycle culture, for example) the development of a household travel survey should include variables that fit the characteristics of the population.

In the particular case and results of the Cincinnati case study under the OKI household travel survey, it can be seen that the type of activities taking place at the trip’s origin and destination have a considerable impact on the result. Unlike, again, for this particular case: the age, gender and income of the rider.

Previous findings concluded that most cyclists had a driver’s license. According to the results of this particular case study, this variable’s impact although not considerable, it is a positive one.

Land use hierarchy organization was determining in estimating the impact of the variable in the final equation for this particular case. The results show a positive impact on ridership.

As it was found on the literature review, riding on two-way streets have a negative impact on a good perception and matching between individuals and infrastructure.
8. **Directions for future research**

The way this equation is developed, it seeks to create groups of people with similar characteristics, all obtainable from a demographic database. In this particular case, the Cincinnati household travel survey did not include personal experiences as part of the research. Based on the literature review, the perception of safety is different in each individual; which will make two people from the same demographic group have different perspectives towards willingness to ridership and traffic behavior. As it has been suggested during communication and meeting with Mr. Jim Coppock and Melanie from the Cincinnati Transportation Engineers office it is necessary to put together a survey to assess the level of expertise, perception of safety, satisfaction and personal history.

Additionally, from the considerably high “activities at O/D” coefficients it is possible to assess the opportunity to develop a new Activity-based model analysis and forecaster.

Also, Mr. Andrew Rohne from OKI has suggested to continue this process, maybe at a personal research level analyzing specific routes taken by riders. By acknowledging personal preferences by type of users it will be possible to do a parallel analysis of possible routes from point of origin and destination and seek improvements by comparison.
References


OKI, 2010. GPS-based household interview survey for the Greater Cincinnati, Ohio region. State job number 134421.


