Dissertations and Theses   Dissertations and Theses

Summer 9-24-2015

An Exploration of Bicyclist Comfort Levels Utilizing Crowdsourced Data

Bryan Philip Blanc
Portland State University

An Exploration of Bicyclist Comfort Levels Utilizing Crowdsourced Data

by

Bryan Philip Blanc

A thesis submitted in partial fulfillment of the

requirements for the degree of

Master of Science
in
Civil and Environmental Engineering

Thesis Committee:
Miguel Figliozzi, Chair
Christopher Monsere
Krista Nordback

Portland State University
2015

ABSTRACT

Bicycle transportation has become a central priority of urban areas invested in improving sustainability, livability, and public health outcomes. Transportation agencies are striving to increase the comfort of their bicycle networks to improve the experience of existing cyclists and to attract new cyclists. The Oregon Department of Transportation sponsored the development of ORcycle, a smartphone application designed to collect cyclist travel, comfort, and safety information throughout Oregon. The sample resulting from the initial deployment of the application between November 2014 and March 2015 is described and analyzed within this thesis. 616 bicycle trips from 148 unique users were geo-matched to the Portland metropolitan area bicycle and street network, and the self-reported comfort level of these trips was modeled as a function of user supplied survey responses, temporal characteristics, bicycle facility/street typology, traffic volume, traffic speed, topography, and weather. Cumulative logistic regression models were utilized to quantify how these variables were related to route comfort level within separate variable groups, and then the variables were used in a pooled regression model specified by backwards stepwise selection.

The results of these analyses indicated that many of the supplied predictors had significant relationships with route comfort. In particular, bicycle miles traveled on facilities with higher traffic volumes, higher posted speeds, steep grades, and less separation between bicycles and motor vehicles coincided with lower cyclist comfort

ratings. User supplied survey responses were also significant, and had a greater overall model variance contribution than objectively measured facility variables. These results align with literature that indicates that built environment variables are important in predicting bicyclist comfort, but user variables may be more important in terms of the variance accounted for. This research outlines unique analysis methods by which future researchers and transportation planners may explore crowdsourced data, and presents the first exploration of bicyclist comfort perception data crowdsourced using a smartphone application.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

Bicycle transportation has become a central priority of urban areas invested in improving

sustainability, livability, and public health outcomes. Metropolitan areas around the

county have set aggressive bicycle mode share objectives for their 2030-2040

transportation plans. The objective of increasing bicycle use for transportation is meeting

at least two interrelated impediments: constrained transportation infrastructure budgets

and the difficulty of successfully converting short automobile trips to bicycle trips by

attracting new cyclists. Constrained infrastructure spending has motivated research into

understanding where bicycle improvements can be made that can yield the maximum net

benefit in terms of increased ridership and safety. The goal of encouraging new bicycle

trips has also motivated research to understand where inadequacies exist in the current

bicycle network that may be barriers to less competent/confident cyclists; thus increasing

the success rate of converting auto trips to bicycle trips. Both of these research interests

have stimulated the development of smartphone applications to crowdsource information

from regional cyclists to understand empirically *where* they ride, *why* they ride, and *what*

improvements could make their cycling experience more safe and comfortable.

In 2014, researchers in the Transportation, Technology, and People Laboratory (TTP

Lab) began working in conjunction with the Oregon Department of Transportation

(ODOT) to develop a smartphone application aimed specifically towards ODOT's desire

to understand Oregon cyclists' bicycle infrastructure preferences and safety issues. While

ORcycle is not the first smartphone application to collect bicycle travel data, it is the first statewide deployment of a smartphone application collecting bicycle specific safety data (in addition to travel data). Besides adding this new data objective, ORcycle also increases the depth to which transportation planners and researchers can understand users' unique characteristics and their preferences/issues with existing bicycle infrastructure.

This thesis will review results from the initial data collection of the ORcycle smartphone application taken between November 2014 and March 2015 in Oregon. Inferences about the relationship between cyclist comfort and explanatory factors will be made using statistical models of the initial data pool. Methodologies will also be outlined for how future data collected using this application may be analyzed to produce increasingly robust and useful results. The goals of this project are to describe the initial sample of ORcycle data and to use that sample in combination with other data sources to make inferences about bicyclist comfort.

1.1   RESEARCH OBJECTIVES

Table 1 outlines the specific research questions this thesis addresses and the methods by which those questions were investigated. The first goal of this research is to describe the expansive initial dataset of ORcycle. The second goal of this research is to use the initial ORcycle sample in combination with other data sources to make statistical inferences about cyclist comfort. Both of these goals are expanded into specific objectives in Table 1.

Table 1: Research Questions

Research Questions	Methods of Analysis and/or Interpretation
What does the initial data sample look like? • What types of users are using the application? • What types of trips are being recorded and where are they being made? • What types of reports are being recorded and where are they being documented? • How do users in the ORcycle sample differ from the Oregon cycling and non-cycling population?	• Descriptive statistics • Histograms • Bar plots • Chi-square tests
What inferences can we make about trip and route characteristics and their relationships with route comfort? • How did other trip questions relate to route comfort? • How did user characteristics relate to route comfort? • How do bicycle facility differences relate to route comfort? • How does topography relate to route comfort? • How does traffic flow relate to route comfort? • How do trip characteristics (speed, distance, duration, time of day) relate to route comfort? • How do weather characteristics relate to route comfort? • How does each of these models change when another group of variables is controlled for?	• Cross-tabulations • Stacked bar plots • Chi-square tests • Ordinal logistic regression

1.2   ORGANIZATION

The remainder of this thesis is organized as follows. First, prior literature pertaining to the following research is reviewed. The methodology behind application development and data cleaning/analysis is then described. Descriptive statistics for the sample of ORcycle data utilized herein are then presented. Statistical models are then utilized to explore the variation in cyclist comfort as a function of potential explanatory factors.

Finally, lessons learned and future research opportunities are discussed and concluding

comments are given.

2.   LITERATURE REVIEW

Before examining the data obtained from the ORcycle smartphone application, it is

pertinent to review related research, data collection methods, and data analysis methods.

Section 2.1 will review methods of evaluating cyclist comfort and relate these methods to

the goals of this project. Section 2.2 will review typologies that have been developed to

classify different types of cyclists. Section 2.3 will review studies that have examined

bicyclist preferences using stated preference surveys. Section 2.4 will review studies that

have examined bicyclist travel choice behavior using GPS devices. Section 2.5 will

review other studies that have examined bicyclist travel behavior using smartphone

applications.

2.1   UNDERSTANDING CYCLIST COMFORT

2.1.1 Bicycle Level of Service

Bicycle Level of Service (BLOS) is a performance measure used to describe the

performance (comfort, safety, operation, etc.) of bicycle facilities and should reflect

travelers' perceptions, be useful to transportation agencies, and be directly measured in

the field (Figliozzi, Blanc, and Johnson 2014). BLOS methods are formulated using

statistical analyses to connect a subjective rating of a bicycle facility's perceived comfort

with geometric, operational, and other characteristics of the bicycle facility. Some BLOS

methods are complex and data intensive. Most BLOS methods are simple, user-friendly,

with readily understandable calculations or scores, and not data intensive. An example of

the former includes the 2010 HCM BLOS; examples of the latter include the Bicycle

Suitability Score (BSS), Bicycle Compatibility Index (BCI) and the Bicycle Suitability

Assessment (BSA). More information on how BLOS methods have evolved over the last

two decades is available in an ODOT report (Figliozzi, Blanc, and Johnson 2014)

BLOS methods fall into three broad analysis tool groups: segment analysis, intersection

analysis, and network analysis. Segment and intersection BLOS are computed using

observed environmental characteristics, while network BLOS measures are computed

using network models. Two tables are presented, which summarize BLOS methods and

the characteristics used to calculate each BLOS measure.

Table 2: Summary of Methods and their Scope (Figliozzi, Blanc, and Johnson 2014)

Method Number	Name	Acronym	Scope	Reference	Reference Year
1	Bicycle Safety Index Rating	BSIR	Segment	(Davis 1987)	1987
2	Bicycle Stress Level	BSL	Segment	(Sorton and Walsh 1994)	1994
3	Road Condition Index	RCI	Segment	Epperson (Epperson 1994)	1994
4	Interaction Hazard Score	IHS	Intersection	(Landis 1994)	1994
5	Bicycle Suitability Rating	BSR	Segment	(Davis 1995)	1995
6	Bicycle Level-of-Service	BLOS	Segment	(Botma 1995)	1995
7	Bicycle Level-of-Service	BLOS	Segment	(Dixon 1996)	1996
8	Bicycle Suitability Score	BSS	Segment	(Turner, Shafer, and Stewart 1997)	1997
9	Bicycle Compatibility Index	BCI	Segment	(Harkey, Reinfurt, and Knuiman 1998)	1998
10	Bicycle Suitability Assessment	BSA	Segment	(Emery and Crump 2003)	2003
11	Rural Bicycle Compatibility Index	RBCI	Rural Segment	(Jones and Carlson 2003)	2003
12	Compatibility of Roads for Cyclists	CRC	Rural Segments	(Noël, Leclerc, and Lee-Gosselin 2003)	2003
13	Bicycle Intersection Safety Index	BISI	Intersection	(Carter et al. 2007)	2007

14	Bicycle Level-of-Service	BLOS	Segment	(Zolnik and Cromley 2008)	2007
15	Bicycle Level-of-Service	BLOS	Segment	(Jensen 2007)	2007
16	Bicycle Level-of-Service	BLOS	Segment	(Petritsch et al. 2008)	2007
17	Bicycle Environmental Quality Index	BEQI	Segment, Intersection	(San Francisco Department of Public Health 2009)	2009
18	Bicycle Quality Index and Cycle Zone Analysis	BQI & CZA	Segment, Network, Zone	(Birk et al. 2010)	2010
19	Bicycle Level-of-Service	BLOS	Segment & Intersection	(Transportation Research Board 2010)	2010
20	Simplified Bicycle Level of Service	BLOS	Segment	(Ali, Cristei, and Flannery 2012)	2012
21	Level of Traffic Stress	LTS	Intersection, Segment, Network, Zone	(Mekuria, Furth, and Nixon 2012)	2012
22	Bicycle Level-of-Service at Intersections	BLOS	Intersection	(Jensen 2013)	2013
23	Protected Bicycle Lane Level of Service	PBL-LOS	Segment	(Foster et al. 2015)	2015

Table 3: BLOS Variables by Category (Figliozzi, Blanc, and Johnson 2014)

Category	Parameter	Data Type	Methods that Utilize Parameter (see Table 2 for a reference).
Bikeway Geometric Design	Facility Type	Categorical	RCI[3], BLOS[7], BCI[9], BISI[13], CZA[18], BLOS[19], BLOS[20], LTS[21], BLOS[22]
	Width of Bicycle Facility	Number (feet)	IHS[4], BLOS[6], BCI[9], BSA[10], RBCI[11], CRC[12], BLOS[15], BEQI[17], BQI[18], BLOS[19], LTS[21]
	Topographic Grade	Number (% grade)	RCI[3], BSR[5], BSA[10], BEQI[17], CZA[18]
Bikeway Environment	Width of MV Buffer (proximity to edge of moving traffic lane)	Number (feet)	BLOS[15], LTS[19], PBL-LOS[23]
	Bicycle marking presence	Categorical	BSA[10], BEQI[17]
	Presence of bicycle signage	Categorical	BEQI[17]
	Presence of trees	Categorical	BEQI[17]
	Presence of bicycle scale	Categorical	BEQI[17]

	lighting		
	Width of Shoulder	Number (feet)	BSS[8], BCI[9], BSA[10], RBCI[11], BLOS[19], LTS[21]
	Presence of Sidewalks	Categorical	BSA[10], BLOS[15]
Roadway Geometric Design	Number of Vehicle Lanes	Number (count)	BSIR[1], RCI[3], IHS[4], BSR[5], BSA[10], BSA[10], BLOS[15], BLOS[16], BEQI[17], BLOS[19], BLOS[20], LTS[21], PBL-LOS[23]
	Width of Outside Lane	Number (feet)	BSIR[1], BSL[2], RCI[3], IHS[4], BSR[5], BLOS[7], BSS[8], BCI[9], BSA[10], RBCI[11], BLOS[15], BLOS[16], BLOS[19]
	Turning Lane Configuration	Categorical	BCI[9], BSA[10], BEQI[17]
	Physical Median	Categorical	RCI[3], BLOS[7], BSA[10]
	Frequent Curves	Categorical	BSA[10]
Bicycling Nuisance/Hazard	Presence of On-Street Parking	Categorical (2)	RCI[3], BSR[5], BLOS[7], BCI[9], BSA[10], BISI[13], BLOS[15], BEQI[17], BLOS[19], LTS[21], PBL-LOS[23]
	Occupancy of On-Street Parking	Number (%)	BCI[9],
	Conflicting Transit Stop Presence	Categorical	BLOS[15],
	Presence of a Curb	Categorical (2)	BSA[10], BLOS[19]
	Storm Drain Grates	Categorical (2)	RCI[3], BSR[5], BSA[10],
	Roadside Hazard Presence (Sand, gravel, vegetation, ditches)	Categorical	CRC[12]
	Restricted Sight Distance	Categorical	BSR[5], BLOS[7], BSA[10], BEQI[17],
	Access point density	Number (# access points per mile)	IHS[4], RBCI[11], CRC[12], BLOS[16], BEQI[17], BLOS[20]
	Numerous Driveways	Categorical	BSA[10]
	Rail Crossings	Number (count)	RCI[3], BSR[5], BSA[10]
	Bike Lane Drop	Number (# times within segment)	BQI[18],
	Difficult Transition	Number per Segment	BQI[18],
Bikeway Condition	Pavement Condition	Location, Picture, Description	BSIR[1], RCI[3], IHS[4], BSR[5], BLOS[7], BSS[8], BSA[10], CRC[12], BLOS[16], BEQI[17], BLOS[19]
Roadway Traffic	Vehicle Traffic Volume	Number (veh/day)	BSIR[1], BSL[2], RCI[3], IHS[4], BSR[5], BSS[8], BCI[9], BSA[10], RBCI[11], CRC[12], BLOS[15], BLOS[16], BEQI[17], BQI[18], BLOS[19], BLOS[22], PBL-LOS[23]
	Right Turning Vehicle	Number (veh per hr or day)	BCI[9]

	Volume		
	Vehicle Speed	Number (mph)	BSIR[1],BSL[2], RCI[3], IHS[4], BSR[5], BLOS[7], BSS[8], BCI[9], BSA[10], RBCI[11], CRC[12], BLOS[15], BLOS[16], BQI[18], BLOS[19], BLOS[20], LTS[21], PBL-LOS[23]
	Percentage of Heavy Vehicles	Number (%)	IHS[4],BCI[9], RBCI[11], CRC[12], BLOS[16], BEQI[17], BLOS[18]
	Motor Vehicle LOS	Categorical (A-F)	BLOS[7]
	Bicycle Lane Blockage	Categorical	LTS[21]
Bikeway Traffic	Average Speed/Acceleration	Number(ft/s or ft/s^2)	BLOS[6], BLOS[19]
	Bicycle Volumes	Number (bikes/hr or day)	BLOS[6], BLOS[19]
	Pedestrian Volume (for multi-use paths)	Number (bikes/hr or day)	BLOS[6], BLOS[19]
Intersection Specific	"No Turn on Red" sign	Categorical	BEQI[17], BLOS[22]
	Intersection Type	Categorical	BISI[13], BLOS[22]
	Intersection Quality	Categorical	BSA[10], CZA[18]
	Crossing Distance	Number (feet)	BISI[13], BLOS[22]
	Number of lanes crossed for cyclist left turn		BISI[13]
	Number of lanes crossed for cyclist right turn		BISI[13]
	Signal Delay	Number (seconds)	BLOS[22]
Built Environment	Activity Density	Number (Pop. + Employment per sq. mile)	IHS[4]
	Adjacent Land Use Type	Categorical	BSR[5], BCI[9], BSA[10], BLOS[15], BEQI[17],
	Multi-modal or TOD Proximity	Categorical	BLOS[7]
	Bicycle parking presence	Categorical	BEQI[17]
Network	Connectivity	Number (connected node ratio)	BEQI[17], CZA[18]
	Presence of Parallel Facility	Categorical	BLOS[7]
	Intersection Density	Number (Intersections per sq. mile)	RBCI[11]
	Road Network Density	Number (Linear Feet per sq. mile)	CZA[18]
	Bicycle Network Density	Number(Linear Feet per sq. mile)	CZA[18]
	Permeability/Barrier	Number ("score" per feet-boundary)	BLOS[7], CZA[18]
	Stops	Number (# stop signs per mile)	BQI[18]
	Route Simplicity	Number (Turns per mile)	BQI[18]
	Detour	% over shortest path distance	LTS[21]

2.1.2 Level of Traffic Stress

In the recent literature, level of traffic stress (LTS) primarily refers to a specific evaluation method developed by Mekuria et al. (Mekuria, Furth, and Nixon 2012). Level of stress is not a new concept, and previous work/methods have utilized similar language (e.g. the Bicycle Stress Level or BSL from 1994 is based on safety levels and physical/mental effort as a function of age).

Unlike BLOS or network BLOS methods, a LTS measure serves as a proxy for measuring the desirability of a bicycle facility for segments of the population with different levels of age, experience or skill. In this report **LTS** is defined as a performance measure that takes into account not only traffic/geometric characteristics of the riding environment but *also the suitability of the environment for different user groups within the population*. LTS can be used to delineate islands of low-stress network connectivity, highlighting disconnections and especially stressful links within a bicycle network.

2.1.3 Bikeability

Another term that is commonly used in the bicycle literature is "bikeability". For example, McNeil (McNeil 2011) proposes a methodology that assigns points to various destination types, such as grocery stores or restaurants, and calculates a score out of one hundred for a given location by totaling up the points for destinations within a twenty minute bike ride. The method is similar to the popular Walk Score®, which calculates a

score out of one hundred for an input address based on the number of destinations within walking distance ("Walk Score" 2014).

Lowry used BLOS in combination with information about relevant destinations to develop another measure of the bikeability of areas (Lowry and Callister 2012). The primary inputs are the bike/street network with BLOS calculated for each link, the locations of destinations, and a weighting scheme outlining the importance or desirability of different destinations.

The Bikeability Checklist (Pedestrian and Bicyle Information Center 2002), developed by the Pedestrian and Bicycle Information Center (PBIC) at the University of North Carolina, is a simple form to be filled out by any citizen to assess the bikeability of their community. The user is asked to take a bike trip to one of their regular destinations and answer a series of questions about the comfort and convenience of their experience.

2.1.4 Bicycle Friendliness

Some bicyclist advocacy groups have developed the concept of "bicycle friendliness". Perhaps the most well known assessment of bicycle friendliness is conducted by the League of American Bicyclists (LAB). Cities or municipalities can submit a paid application biannually to the LAB for potential recognition as a "bicycle friendly community" at either the platinum, gold, silver, or bronze designation; with platinum being the highest designation. The LAB evaluation is based on assessment of the municipality with respect to five categories: engineering, education, encouragement, enforcement, and evaluation.

LAB also has a state level assessment based on five categories: legislation, policies, and programs; infrastructure; education; enforcement; and evaluation. Instead of an application process, LAB assesses every state in the country on an annual basis and ranks them on their statewide bicycle friendliness. LAB also has recently started evaluating bicycle friendly businesses and universities. Other national and state organizations evaluate bicycle friendliness at various scales. Oregon's Bicycle Transportation Alliance (BTA) developed the *Bike Friendly Report Card* to compare the bicycle friendliness of cities throughout Oregon (Bicycle Transportation Alliance 2014).

2.1.5 **Terminology Summary**

In this thesis, **BLOS** is defined as any bicycle performance measure that *can be computed* (based on a formula or score) *utilizing data/variables that are measured or observed in the field* (geometric, environmental, nuisance, or traffic variables).

Network BLOS is a performance measure (or weighted set of performance measures) used to describe the performance of bicycle facilities at the network level. Network BLOS should also reflect bicyclists' perceptions but they *are not measured in the field but using network models* (i.e. in networks defined by sets of nodes and links) and are usually best calculated using software packages (GIS systems or network algorithms).

LTS is defined as a performance measure that takes into account not only traffic/geometric characteristics of the riding environment but *also the suitability of the environment for different user groups within the population*. LTS can be used to

delineate islands of low-stress network connectivity, highlighting disconnections and especially stressful links within a bicycle network.

Bikeability is a macro-level assessment of a network of bicycle facilities in terms of the *accessibility to important destinations*.

Bicycle friendliness is a macro-level assessment *at the community and government level*. Friendliness is related to the degree of acceptance of cycling within the community and with the adoptions of programs, laws, and policies that protect and promote cycling.

These terms and their unique features are outlined in Table 4.

Table 4 : Overview of Terminology and Keywords (unique feature underlined)
(Figliozzi, Blanc, and Johnson 2014)

Term→ Feature/Scope ↓	BLOS	Network BLOS	Level of Stress	Bikeability	Bicycle Friendliness
Segment/Intersection	✓		✓	✓	
Network PMs		✓	✓	✓	
User Group			✓		
Accessibility				✓	✓
Community & Government					✓

2.2   CYCLIST TYPOLOGY

In order to better understand how to design bicycle facilities that will serve a wide population segment, bicycle planners at many different agencies have attempted to categorize utility cyclists in their jurisdictions based on their differential preferences.

Since the introduction of Cycle Atlanta (as presented in section 2.5.3), many of the smartphone applications collecting cyclist travel behavior data have been using a modified version of Roger Geller's typology (see section 2.2.1) as outlined in the four cyclist categories in Table 5.A "Comfortable, but cautious" category was added in Cycle Atlanta. The "No Way, No How" category is also removed, likely assuming that non-cyclists will not be using the application to track rides. The goal of this self-reported typology is to estimate distribution of difference cyclist types within the application user population.

Table 5: Cycle Atlanta Rider Type Table Schema

Answer ID	Rider Type Category
0	No data
1	Strong and fearless
2	Enthused and confident
3	Comfortable, but cautious
4	Interested, but concerned

However, this language was considered confusing by the researchers and may mean something different to transportation researchers than it does to the general population. Therefore, the ORcycle project team proposed that more specific questions be utilized to elicit precise information about different user groups. These questions are outlined in section 3.1.2.2. A brief literature review of cyclist typology and market segmentation research applied to bicycling populations is presented below.

2.2.1 **Geller 2006 (Portland, OR)**

One of the most publicized rider categorizations is that of Portland Bicycle Coordinator Roger Geller, who estimated that residents of the city of Portland could be categorized into four distinct groups (Geller 2006). Geller's estimates were based largely on his

extensive experience in working for the Portland Bureau of Transportation. These
categories' estimated proportions within the Portland population were substantially
validated by research conducted at Portland State University (Dill and McNeil 2012). A
summary table of both the estimated proportions by Geller and the surveyed proportions
by Dill and McNeil in the Portland area is shown in Table 6.

Table 6: Distribution of Rider Types within Portland Area (Dill and McNeil 2012)

Rider Type	City of Portland	Rest of Portland Metro Area	All	Geller's estimate for City of Portland
Strong & Fearless	6%	2%	4%	<1%
Enthused & Confident	9%	9%	9%	7%
Interested but Concerned	60%	53%	56%	60%
No Way, No How	25%	37%	31%	33%

Other typologies have also been proposed, generally with a more empirical basis
predicated on self-segmenting survey results or utilizing cluster or factor analyses.

2.2.2 **Reid J. 2011 (Victoria, Australia)**

In 2010, a market segmentation analysis was conducted in the state of Victoria, Australia
by a market research agency at the behest of VicRoads, an organization striving to make
cycling safer in Victoria in the midst of rapidly increasing cycling rates (Reid 2011).
Both factor and cluster analyses – statistical techniques for quantitatively detecting
distinct groups within a dataset – were utilized to identify the key cycling segments in
Victoria. The three segments identified were given the following names: "Let's go for a
ride" (LGFAR), "This cycling life" (TCL), and "Catch me if you can" (CMIYC).
LGFAR represented the largest proportion of cyclists within Victoria (75%), and
contained cyclists who generally only ride for recreational purposes and are averse to

15

high levels of stress while cycling. TCL corresponded to 20% of cyclists in Victoria, and contained cyclists that regularly ride their bicycles in a variety of conditions, on nearly all facility types, and for various purposes. Cycling is central to their identity, but they are generally respectful of road rules. In opposition, the CMIYC group (representing the remaining 5% of Victorian cyclists) is generally not respectful of road rules, and will ride anywhere, anytime, and during any condition. These findings helped VicRoads to design plans for outreach about increasing cycling safety to each market segment separately, likely increasing their outreach success.

2.2.3 **Damant-Sirois et al. 2014 (Montréal, QC, Canada)**

In 2013, a team of researchers from McGill University in Montréal, QC analyzed survey data from over 2,000 Montréal cyclists using factor and cluster analyses to detect four distinct groups of cyclists in the city (Damant-Sirois, Grimsrud, and El-Geneidy 2014). The goal of the group identification was to obtain a greater understanding of the differing needs and wants (in terms of infrastructure provision and advocacy efforts) of the distinct cycling groups. The four groups were given the following names: "dedicated cyclists", "path-using cyclists", "fairweather utilitarians", and "leisure cyclists". The analysis was based on seven groups of variables identified to effect cyclist preference: weather and effort, time efficiency, dislike cycling near cars, bicycle route infrastructure, peer and institution encouragement, cycling identity and enjoyment, parental encouragement. The groups were determined by their responses related to these variable groups, with different relationships emerging in each group. For example, dedicated cyclists' decision to cycle is "not strongly impacted by weather conditions", and they are motivated to cycle by the

"speed, predictability, and flexibility of bicycle trips" (Damant-Sirois, Grimsrud, and El-Geneidy 2014).

2.2.4 **Fernanadez-Heredia et al. 2014 (Ciudad Universitaria, Madrid, Spain)**

Fernandez-Heredia et al. conducted a survey at a large university in Madrid, Spain to quantify "psycho-social" factors related to bicyclist perspectives and reasons for bicycle use (Fernández-Heredia, Monzon, and Jara-Díaz 2014). Using factor analyses, four major groups of variables were identified within the results of the survey. The four factor groups are the following:

1. Convenience: This group includes measures of the efficiency and flexibility of the bicycle as a mode of transportation.
2. Pro-Bike Interests: Set of ideas associated with bicycle riding, such as cost savings, environmental friendliness, healthy lifestyles, and enjoyment.
3. External Restrictions: Exogenous variables effecting ones decision to bike; such as perception of danger or the availability of comfortable bicycle infrastructure.
4. Physical Determinants: Variables related to the physical fitness of the user in riding a bicycle.

While this study did not explicitly outline a typology distribution for cyclist categorization, it lent empirical basis to the psychological factors related to bicycle use.

2.2.5 **Overview**

Cycling typologies are important to transportation researchers and planners as they clarify how and where divides exist in the population when considering existing and potential cyclists. This segmentation of groups is helpful in understanding how cyclists of different "types" may respond differently to infrastructure, programming, or other

interogations aimed at increasing cycling rates. Three of the reviewed cyclist typology

population breakdowns are outlined in Table 7.

Table 7: Comparison of three different cyclist typologies

Typology Author and Reference	Geographical Location	Segment Name	Estimated percentage of area population
Geller (Geller 2006)	Portland, Oregon, United States	Strong & Fearless	<1%
		Enthused & Confident	7%
		Interested but Concerned	60%
		No Way, No How	33%
Reid (Reid 2011)	Victoria, Australia	Catch me if you can	5%
		This cycling life	20%
		Let's go for a ride	75%
(Damant-Sirois, Grimsrud, and El-Geneidy 2014)	Montréal, QC, Canada	Dedicated cyclists	24%
		Path-using cyclists	36%
		Fairweather utilitarians	23%
		Leisure cyclists	17%

2.3   STUDIES USING STATED PREFERENCE SURVEYS

Studies of stated cyclist comfort preferences using surveys are reviewed in the following

section.

2.3.1 Stated Preference Survey in Portland, OR

In 2005, researchers at Portland State University conducted a random phone survey in

Portland, OR to explore the relationship between cycling rates and demographics,

measurable characteristics of the built environment (e.g. bicycle infrastructure

availability), perceptions about the environment, and attitudes (Dill and Voros 2007).

Findings from the survey are summarized briefly below.

1) Demographics

    a) Men and younger adults cycled more and were more likely to want to cycle more.

b) Other demographic relationships (household income, vehicle ownership etc.) with rates of cycling were inconclusive.

2) Built environment characteristics

a) No significant relationship between bicycle infrastructure availability and rates of cycling.

3) Built environment perceptions

a) Respondents who positively perceived the cycling environment (e.g. felt bicycle network was safe and accessible) were more likely to be regular cyclists.

b) The most common deterrent to cycling rates was the perception of "too much traffic".

4) Attitudes

a) People living in households with other adults who cycled regularly, had coworkers who cycled regularly, or who saw adults cycling on their street frequently were more likely to be regular cyclists themselves.

2.3.2 Cycling in Cities Survey in British Columbia, CA

In 2006, researchers at the University of British Columbia (Vancouver, BC, CA) conducted a survey of over 1,400 current and potential cyclists in the Vancouver metropolitan area. The survey evaluated motivators and deterrents to cycling among the sample. Several factors had significant impacts on the stated likelihood of cycling: safety, ease of cycling, weather conditions, route conditions, and interactions with motor vehicles (Winters et al. 2011).

The researchers also analyzed cyclists' preferences for specific route types and estimated the likelihood of cycling based on facility type and cycling frequency within the sample. Cycling frequency affects cyclist stress tolerance and thus impacts facility preferences and route choice (Teschke and Winters 2013). The results of the analysis are illustrated in Figure 1, from which a primary conclusion is that only frequent (in this case being at least once per week) cyclists would ride on busy streets without physical traffic separation. Occasional and potential cyclists desired facilities on quiet streets or entirely separated facilities.

16 Route Types, Average Likelihood of Choosing by Cyclist Type
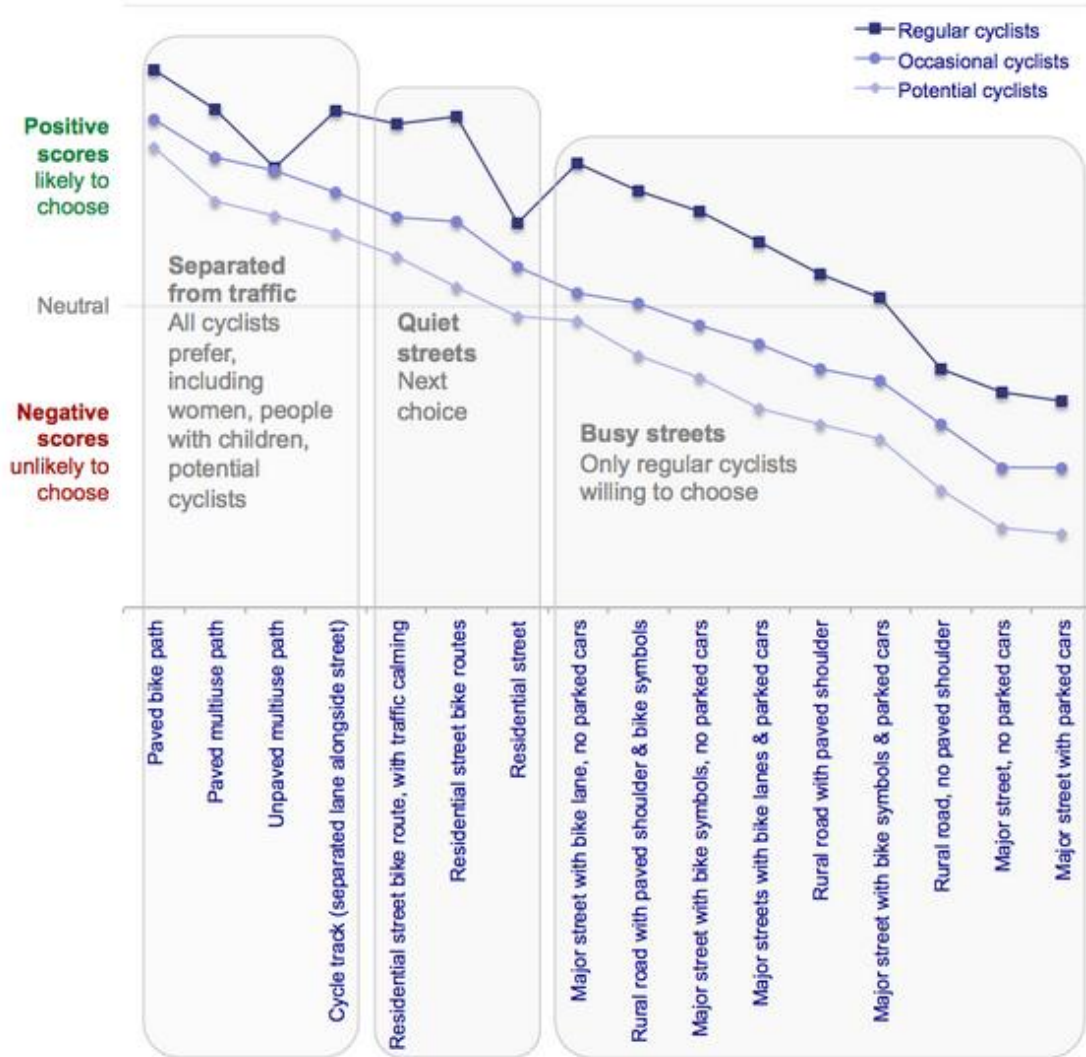
Figure 1: Likelihood of choosing facility type vs. cycling frequency (Teschke and Winters 2013)

2.3.3 An analysis of bicycle route choice in Texas using a web-based survey

In 2008, researchers in Austin, TX administered a statewide web-based survey that

elicited stated preference information on the characteristics informing bicycle route

choice (Sener, Eluru, and Bhat 2009). The survey functioned by first collecting

information about the participant (demographics, bicyclist type) and then asking a user to pick from among a set of three hypothetical routes with listed characteristics theorized to affect bicyclist route choice. Six groups of variables were analyzed for their contribution to bicyclist route choice: (1) bicyclist characteristics, (2) on-street parking characteristics, (3) bicycle facility characteristics, (4) roadway characteristics, (5) traffic characteristics, and (6) travel time. The route choices and their respective characteristics were then incorporated into a multinomial logit model to estimate the relative utility (or disutility) of the variables examined. Parking related attributes were found to be significant; bicyclists preferred routes with minimal on-street parking. Continuous bicycle facilities, lower traffic volumes and speeds, and fewer intersections were all found to increase the relative utility of a bike route.

2.3.4 **Stated Preference Survey in Waterloo, Ontario**

A stated preference web survey was conducted by researchers in the Waterloo, Ontario area in 2010 (J. Casello et al. 2011). The survey was administered along with a GPS bicycle travel study (further discussed in 2.4.3) on the same sample of 100 cyclists. The survey focused on seven categories of information:

1. Demographics and auto ownership
2. Characteristics of regular cycling route
3. Cycling behavior
4. Specific cycling hazards
5. Cycling economics

6. Necessary cycling infrastructure

7. Use of GPS/cell phone while bicycling and other questions about cycling deterrents

The sample consisted of self-identified "winter cyclists", which likely resulted in a substantial bias of the sample away from the general cycling population; especially when considering Waterloo's harsh winter climate. The sample was also somewhat biased towards higher income cyclists, with 57% of cyclists reporting their personal incomes as greater than $50,000. This indicated that cycling for transportation was likely a choice, rather than a necessity, within the sample.

The researchers were able to estimate the importance of a number of factors that influenced cycling travel behavior by calculating a weighted average of the ordinal survey responses. Each survey response was a rating of the importance of some factor proposed to effect bicycling behavior on a scale of 1 (least important) to 5 (most important). The results are outlined in Table 8. Convenience was the top motivation for cycling, while safety was the primary obstacle and consideration used in route selection. Accompanying this survey was GPS data collection discussed in section 2.3.4.

Table 8: Relative ranking of influences on cycling behavior (J. Casello et al. 2011)

Motivations for cycling	
Variable	Importance
Convenience compared to other modes	4.26
Contribution to environment	4.19
Lower cost compared to other modes	3.80
Allows for recreation	3.42
Improves health	3.40
Obstacles to cycling	
Feels unsafe	2.70
Poor motorist behavior	2.66
High traffic volumes	2.65
Poor road conditions	2.32
Travel time is long	1.88
Poor weather	1.63
Many stops	1.62
Distance travelled is long	1.54
Lack of bike parking	1.44
Route not scenic	1.23
Factors influencing route choice	
Feels safe	2.91
Shortest by time	2.90
Low amount of traffic	2.83
Best road conditions	2.64
Shortest by distance	2.55
Fewest stops	2.10
Route is scenic	1.99

2.4   STUDIES USING GPS DEVICES

Studies of revealed and stated cyclist comfort preferences using survey questions and

GPS devices are reviewed in the following section.

2.4.1 GPS Data Collection in Minneapolis, MN

In the spring of 2006, researchers in Minneapolis, MN used GPS units to study the

cycling behavior of 55 cyclists over the course of three weeks; focusing specifically on

commute trips (Harvey, Krizek, and Collins 2008). The study compared the preferred

route of each cyclist with the calculated shortest route based on trip distance and bicycle

facility type (off-street path, on-street bike lane, or a road with no designated bicycle facility). Participants also reported demographic characteristics and their "cycling comfort" on a 1 through 5 scale; with 1 indicating that the cyclist was only comfortable riding on off-street paths, and 5 indicating the cyclist was comfortable on urban streets with heavy traffic. A linear regression model was constructed to examine the relationship between the distance traveled out of the way (the difference between the distance of the chosen route and the shortest route) and several predictive variables, including bicycle facility type, historical route safety, traffic control type, number of intersections along route, cycling comfort level, gender, and age. The only independent variable found to be statistically significant was the rider's reported cycling comfort level. The authors posit that this finding indicates that cyclists with lower comfort levels are more willing to travel out of their way to use a preferred route rather than the shortest one.

The other variables measured were not significant, which could be partly due to a small sample size (55 cyclists), but is also likely due to the research design; which instructed cyclists to use a single preferred route over the study period rather than allowing the cyclists to choose their route based on their individual circumstances, resulting in a restricted range for the variables. In later studies, these other variables were found to be significant predictors, but riders were able to use a larger number and variety of facilities, resulting a much wider range of variability, which likely increased statistical significance of the predictive variables.

2.4.2 GPS Data Collection in Portland, OR

2.3.2.1   *Data Collection*

In 2007, researchers in Portland, OR used handheld GPS units to examine the cycling behavior of 164 participants from March through November of that year in the Portland area (Broach, Dill, and Gliebe 2012). Trip purpose was reported by the user for each trip, and only utilitarian (non-exercise) trips were kept the in the dataset used for analysis. After additional data cleaning, the GPS coordinate traces for the remaining 1,449 trips were matched to the bicycle and street network in the Portland area supplied by the local MPO (Oregon Metro) and modified to include additional bicycle links observed in the GPS data. Attached to the bicycle/street network were facility characteristics (e.g. bicycle facility type), environmental variables (e.g. topography), and traffic volumes (e.g. AADT), and these were used in a route choice model (multinomial logit) to compare the characteristics of the route chosen by the cyclist with those of the shortest route.

The results indicated the relative utility (or disutility) the cyclists in the sample attributed to the predictive variables. For example, cyclists were willing to travel significantly out of their way (estimated 17.9% of trip distance) to use bicycle boulevards, while they were willing to travel even farther out of their way (estimated 72.3% trip distance) to *avoid* a path with a 2-4% upslope. Overall, the results indicated several significant characteristics associated with route choice within the sample; namely distance, turn frequency, slope, intersection control type, traffic volumes, and bicycle facility type (e.g. Bicycle Boulevard vs. arterial road).

Metro Bicycle Model

Oregon Metro is the Portland area's metropolitan planning organization and manages the regional travel demand model, including a bicycle travel demand model. Using the results of the study conducted by Broach, Gleibe, and Dill, Metro incorporated route characteristics into the travel demand model to more accurately predict what routes bicyclists would use – as opposed to the standard motor vehicle approach of simply considering trip distance and/or travel time (Stein 2011). They incorporated the following variables into the model, most of which are cataloged geographically in Metro's Regional Land Information System (RLIS):

Table 9: Route characteristics incorporated into Metro's regional bicycle demand model

Variable	Impact on bicycle utility of route (+/-)
Proportion of route on off-street paths, bike boulevards, bike lanes	+
Proportion of route on links with grade > 2%	-
Turns, traffic signals, stop signs per mile	-
Traffic volumes of on-street travel and opposing links at left turns	-
Bridge bike facility type	+ or -
Distance	-

2.4.3 **GPS Data Collection in Waterloo, Ontario**

2.3.3.1 *Initial Sample*

GPS data was collected for 100 cyclists in Waterloo, Ontario in 2010 (J. Casello et al. 2011) in combination with the survey data collection discussed in section 2.3.4. A total of 1,232 trips were recorded over a five-week period. The study applied the GPS and survey data collected to four-step travel demand model. Using a regression model, the researchers calculated trip rates based on population and employment density. Trips were then distributed throughout the region using an observed distribution of trip lengths. Self-reported mode and path choice were modeled over the survey responses and the observed GPS traces. The development of these models enabled the researchers to highlight variables that were predicted to increase cycling rates.

2.3.3.2 *Final Sample and Route Choice Model*

After filtering the initial trip set for very short trips and inaccurate GPS traces, the research team had 724 trips from which to construct a route choice model. Five route characteristics were used as predictive variables in the route choice model: (1) the length of each link in the network, (2) the posted auto speed of each link, (3) the auto volume of each link, (4) the gradient (elevation change) of each link, and (5) the presence or absence of a cycling lane (J. M. Casello and Usyukov 2014). Using these five attributes, the resulting route choice model was able to select the observed route for 65% of the trips, and an additional 13% of trips were very close to the observed selection.

2.4.4 GPS Data Collection in Zurich

In 2009 researchers in Zurich, Switzerland analyzed GPS data sourced from a private sector data collection effort (Menghini et al. 2010). Unlike the other studies reviewed herein, this required the mode of the GPS trace to be imputed so that the bicycle trips could be analyzed separate from trips made by other modes. The travel mode imputation procedure was outlined in the white paper. After removing non-bike trips, and filtering out GPS traces that could not be matched to the Zurich bicycle and street network, 636 GPS traces remained. After matching the GPS traces, feasible alternative routes were generated from each trace's origin and destination. A descriptive analysis compared the chosen routes with the alternative routes on a number of characteristics, and exposed differences in the grade of the route chosen (routes chosen were less steep) and the proportion of the route chosen along dedicated bicycle facilities (routes chosen included a higher portion of dedicated bicycle facilities). These traces were applied to a multinomial logit route choice model, which found that topography had a statistically significant negative impact on cyclists' utility (leading them to choose routes with more gentle topography). Route length was also found to have a statistically significant impact on cyclist utility, though this is typical of nearly all route choice models.

2.5   STUDIES USING SMARTPHONE GPS

Studies of revealed and stated cyclist comfort preferences using online computer surveys are reviewed in the following section.

2.5.1 CycleTracks (San Francisco, CA - 2009)

2.5.1.1 *The Application*

In 2009, the San Francisco County Transportation Authority developed and released a smartphone application (CycleTracks for Android and iOS) to collect GPS and travel survey data about cyclists in the San Francisco area. CycleTracks uses a smartphone device's Global Positioning System (GPS) sensor to track a user's time and space trajectory.   It can also provide some (optional) user demographic information; the demographic information is collected to study self-selection and overrepresentation of some user groups. The application is available for download free of charge on the iTunes app store or the Android Play app store.

The development team had several critical criteria to guide the application development (Schwartz and Hood 2011):

1. It must be free and quick to download and install
2. It must be as easy to use as possible, with minimum tapping/clicking necessary to get started, so even causal cyclists can use it
3. It must upload every track data immediately to [SFCTA's] central database using the phone's built-in data plan, so the user doesn't have to manually intervene, sync, or upload anything
4. It must not run down the user's battery
5. It needs a catchy name

The application recorded GPS coordinates which could later be geo-matched to road and bicycle networks. Trip purpose is recorded at the end of each trip, with the following trip

purposes given as options: (1) commute, (2), school, (3), work-related, (4) exercise, (5) social, (6) shopping, (7) errand, and (8) other.

If a trip purpose was considered an "Other", the user can enter more details about their trip purpose into the comments field associated with each trip. A comments field was optionally filled in for each trip, and could supplement SFCTA's information about a route or trip. Users could then view their trip on a map. Users also had the option of inputting demographic information within the "Settings" sub-menu; this only had to be done once. The optional additional information fields available are: age, e-mail address, gender, home ZIP code, work ZIP code, school ZIP code, and cycling frequency.

The basic application functionality is illustrated in Figure 2. More information about the application functionality can be found on CycleTracks's website[1] or a 2011 Transportation Research Board (TRB) paper (Schwartz and Hood 2011).

---

[1] http://www.sfcta.org/modeling-and-travel-forecasting/cycletracks-iphone-and-android
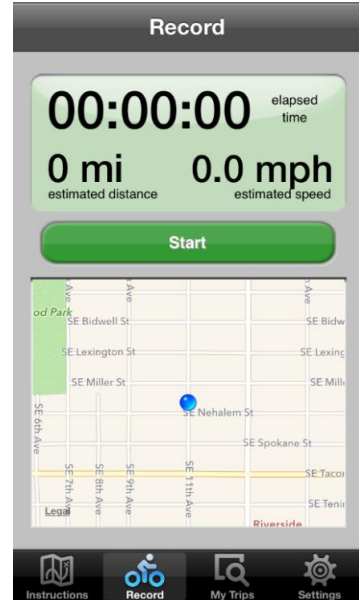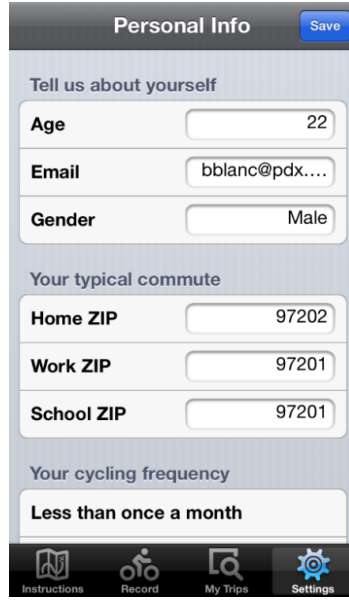
1. The application is opened

2. The user can optionally enter in demographic information and cycling frequency

3. The user presses "Start" to begin recording a trip. GPS coordinates are now being recorded.
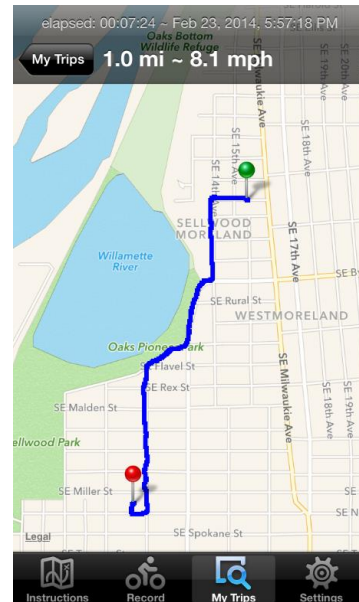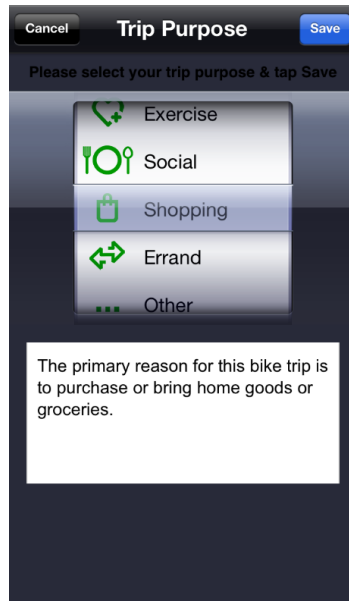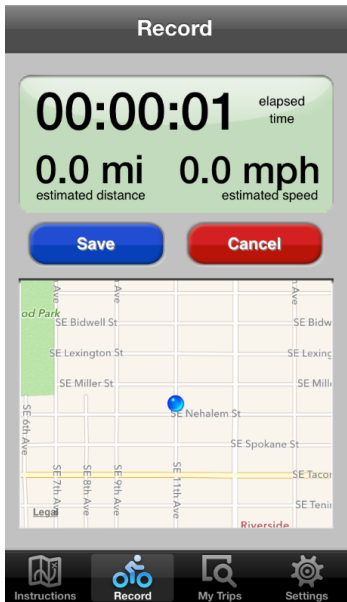
4. When the user arrives at their destination, the trip can be recorded by pressing "Save".

5. The trip purpose is then entered, and the trip is then transferred to the server.

6. The user then can review their trip on the Google Maps API.

Figure 2: CycleTracks User Interface and Functionality (iOS version shown)

2.5.1.2   *Initial Results*

The initial survey sample collected over 7,000 trips from 1,083 users between November 2009 and April 2010. SFCTA then developed a bicycle route choice model from this sample and incorporated it into their SF-CHAMP regional travel demand model (Zorn, Sall, and Bomber 2011). The CycleTracks source code is open source and available free to the public[2]. All subsequent smartphone applications reviewed herein are built upon the original CycleTracks source code.

The San Francisco County Transportation Authority manages the San Francisco area's regional travel demand model, including a bicycle demand model. Prior to the data collected from CycleTracks (see section 2.5), the SFCTA had assumed cyclists would choose the shortest path from their origin to destination. The SFCTA was able to improve the route choice portion of the demand model to incorporate cyclists' differential preferences as revealed by the route choice model built from the CycleTracks data. They incorporated the following variables into the model (outlined in Table 10).

---

[2] https://github.com/sfcta

Table 10: Route characteristics incorporated into SFCTA's regional bicycle demand model

Variable	Impact on bicycle utility of route (+/-)
Route distance	-
Turns per unit distance	-
Proportion of the route going the wrong way on a one way street	-
Proportion of the route on dedicated bike facilities	+
Proportion of the route on signed bike routes (shared with motor vehicles)	+
Average up-slope	-

2.5.2 CycleTracks (Austin, TX - 2011)

After the initial success of CycleTracks in San Francisco, the application was deployed in

Austin, TX in 2011 by researchers at the Texas Transportation Institute (TTI) between

May and October 2011. Over 3,600 GPS traces were collected from over 300 users, many

of which provided demographic data so the researchers could evaluate sample bias. The

results were summarized in a report (Hudson et al. 2012), which cross tabulates many of

the demographic characteristics with bicycling environment variables, such as bicycle

facility type. The report also adds valuable information about data cleaning and

processing procedures.

The goal of the project was to test if using CycleTracks was feasible to apply in another

region and would provide useful information for decision making in planning bicycle

networks and infrastructure. At the conclusion of the study, the researchers remarked that "the amount of information provided by the use of CycleTracks far exceeds what would be available using other data collection methods" (Hudson et al. 2012).

2.5.3 Cycle Atlanta (Atlanta, GA – 2012)

2.5.3.1   *The Application*

In 2012, researchers at Georgia Tech worked with the City of Atlanta and the Atlanta Regional Commission to modify CycleTracks for deployment in the Atlanta, GA region. Cycle Atlanta includes all of the functions performed by CycleTracks but adds several additional features and uses a different user interface. Screenshots of the user interface are presented in Figure 3.
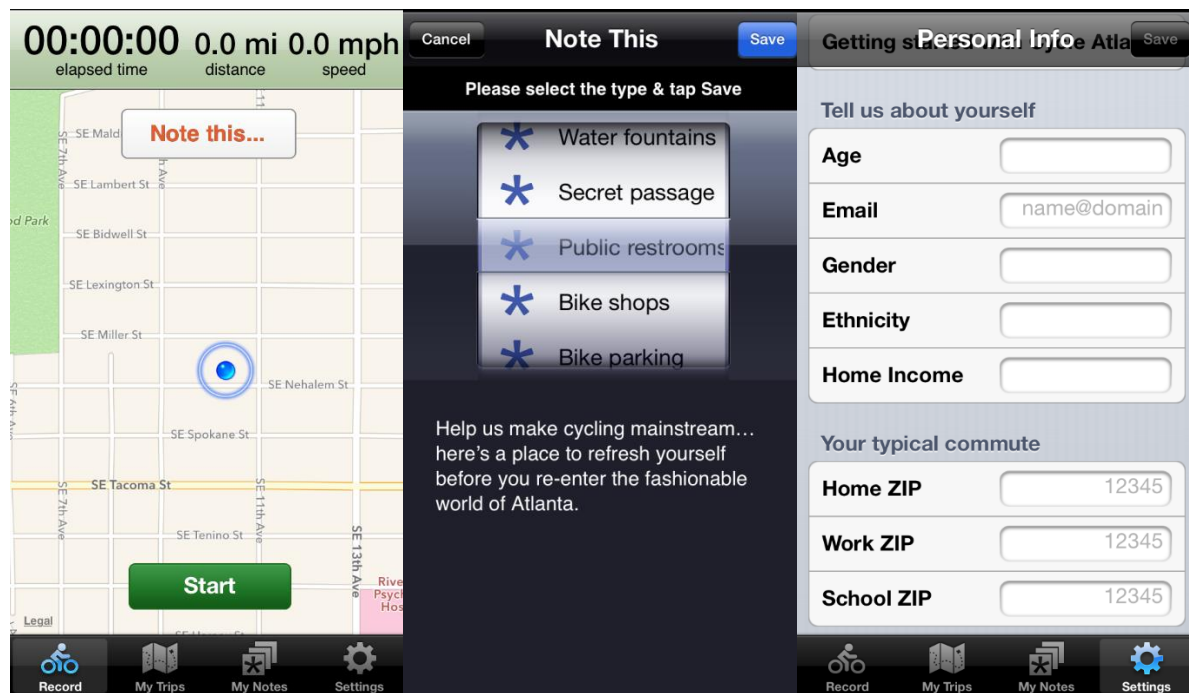
1. Google Maps API fronts user interface
2. "Notes" can be made about assets or issues
3. Demographic information is entered in

Figure 3: Cycle Atlanta screenshots (iOS version shown)

In addition to collecting GPS bicycle route data, Cycle Atlanta can also crowdsource information about geo-located bicycle deterrents (e.g. pavement issues, traffic signal deficiencies, etc.) or amenities (e.g. bicycle parking, water fountains, etc.) (Misra et al. 2014). These deterrents and amenities (called "notes") are selected from a categorical list and can be supplemented with descriptive text and/or a photo. The following notes are available for selection:

Table 11: Cycle Atlanta Note Selection

Issues/Deterrents	Assets/Amenities
Pavement issues	Water fountain
Traffic signal issue	"Secret Passage" [3]
Enforcement request	Public restroom
Bicycle parking request	Bicycle shop
Bicycle lane design issue	Bicycle parking
Custom entry	Custom entry

Cycle Atlanta also can collect additional (optional) user socio-demographic information: ethnicity and household income (both categorized). It also breaks the age field into categories, instead of requesting a numerical entry. The categories for each field are listed below:

---

[3] "Secret Passage" identifies bicycle-navigable paths that are not on map

Table 12: Cycle Atlanta Demographic Categories

Ethnicity	Household Income	Age
White	Less than $20,000	Less than 18
African American	$20,000 to $39,999	18-24
Asian	$40,000 to $59,999	25-34
Native American	$60,000 to $74,999	35-44
Pacific Islander	$75,000 to $99,999	45-54
Multi-racial	$100,000 or greater	55-64
Hispanic/Mexican/Latino		65+
Other		

Finally, Cycle Atlanta also collects (optional) data about the type of cyclist using the application. The rider can indicate their type using a modified version of Geller's (Geller 2006) cyclist typology: "Strong & fearless", "Enthused & confident", "Comfortable, but cautious", or "Interested, but concerned". The rider can indicate its level of experience/years riding by choosing among these options: "Since childhood", "Several years", "One year or less", "Just trying it out/just started".

A website[4] was developed to display the information as a live feed coming from the application; displaying trips, notes, and aggregated user statistics to the public. The Cycle Atlanta application is also available to the public as an open source codebase[5].

---

[4] http://cycleatlanta.org/version2/CATLMaps.php

[5] https://github.com/cledantec?tab=repositories

2.5.4 **Mon ResoVelo (Montreal, QC- 2013)**

2.4.1.1 *The Application*

In 2013, researchers at McGill University worked with the City of Montréal in Montréal, QC, Canada to develop Mon RésoVélo, which was built off of the CycleTracks and Cycle Atlanta open source codebases. Mon RésoVélo does not include the "deterrent and amenity reporting" present in Cycle Atlanta and RenoTracks (see section 2.5.5) but their authors claim that the app improves several other application functions (Jackson et al. 2014).

The first difference between Mon RésoVélo and prior applications is a difference in user interface design. User interface screenshots are shown in Figure 4. The application comes with a complete French language interface option.
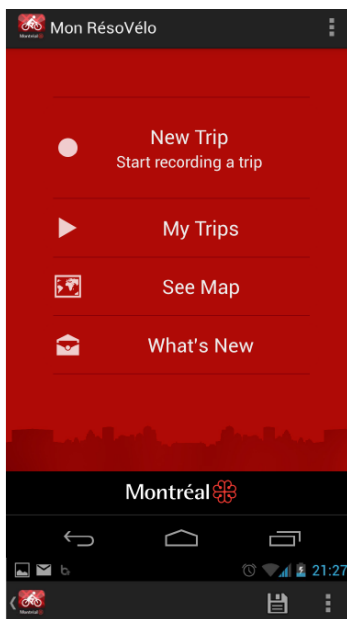
Home navigation screen
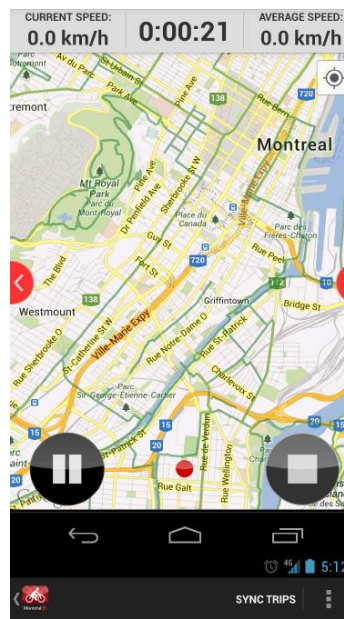
View trip through Google Maps API
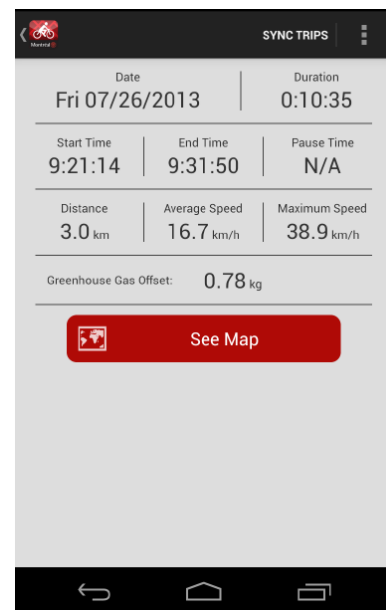
End of trip summary

Figure 4: Example screenshots of Mon RésoVélo interface (Android version shown)
(Jackson et al. 2014)

The app developers indicate that Mon RésoVélo also restructures the underlying GPS

data collection model to "break single trips into a series of segments to manage more

easily stopping, pausing, GPS connection loss, and forgetting to turn off GPS collection

when finishing a trip" (Jackson et al. 2014). Finally, Mon RésoVélo adds a greenhouse

gas emissions calculator based on local conditions (Jackson et al. 2014). A calorie

counter is also included that corrects for cyclist weight. Mon RésoVélo's codebase is not

available open source, but was later adapted by Brisk Synergies[6] to be folded into their

Brisk Cycle platform, which could be re-branded for other regional deployments, as was

done in Toronto, ON.

2.5.4.2   *Initial Results*

A paper was presented at the 2014 Transportation Research Board meeting summarizing

some preliminary results (Jackson et al. 2014). As reported in the other smartphone

application studies, the sample of users analyzed was biased towards young (24-44)

males, with ages 24-34 comprising 46% and ages 35-44 comprising 23% of the

participants, while 73% of the users were male. While no specific data were cited, the

paper also stated that the relative popularity of different bike routes in Montréal were

comparable with the proportions observed by the city's bike counters.

---

[6] http://www.brisksynergies.com/briskcycle/

2.5.4.3 *Mon RésoVélo and Safety Data*

An innovative use of the data collected by Mon RésoVélo was presented at the 94th

Transportation Research Board Annual Meeting (Strauss, Miranda-Moreno, and Morency

2015) GPS traces sourced from Mon RésoVélo were combined with point bicycle counts

in the city of Montreal to represent network wide bicyclist exposure rates. These

exposure rates were combined with geocoded safety and injury data to create an injury

risk model. This injury risk model can highlight areas of considerable risk taking into

account both injury rates and exposure rates.

2.5.5 Other Bicycle Data Smartphone Applications

Many other smartphone applications were created based on CycleTracks, some

improving or expanding upon its features (including ORcycle, which this thesis centers

around).

Table 13 outlines all known Cycletracks derived applications that have been deployed in various cities across North America.

Table 13: CycleTracks and its derivative applications

Year-Month First Released	City/Region	Re-branded or Improved App?	Application Name	Project Link
2009-11	San Francisco, CA	-	CycleTracks	http://www.sfcta.org/modeling-and-travel-forecasting/cycletracks-iphone-and-android
2011	Lane County, OR	Yes	Cycle Lane	http://www.thempo.org/611/CYCLELANE---Bike-routes
2011-05	Austin, TX	No	CycleTracks	http://ntl.bts.gov/lib/45000/45700/45731/Hudson_11-35-69.pdf
2012-06	Minneapolis/St. Paul, MN	No	CycleTracks	http://www.minneapolismn.gov/www/groups/public/@publicworks/documents/images/wcms1p-094499.pdf
2012-10	Atlanta, GA	Yes	Cycle Atlanta	http://cycleatlanta.org/
2012-Summer	Fort Collins, CO	No	CycleTracks	http://today-archive.colostate.edu/story.aspx?id=7744
2013-07	Montréal, QC	Yes	Mon RésoVélo	http://ville.Montréal.qc.ca/portal/page?_pageid=8957,112451619&_dad=portal&_schema=PORTAL
2014-01	Reno, NV	Yes	RenoTracks	http://renotracks.nevadabike.org/
2014-05	Lexington, KY	No	CycleTracks	http://www.kentucky.com/2014/05/04/3227486/lexington-bicyclists-help-sought.html
2014-06	Philadelphia, PA	Yes	CyclePhilly	http://www.cyclephilly.org/
2014-11	Toronto, Ontario	No	Toronto Cycling App	http://www1.toronto.ca/wps/portal/contentonly?vgnextoid=5c555cb1e7506410VgnVCM10000071d60f89RCRD&vgnextchannel=6f65970aa08c1410VgnVCM10000071d60f89RCRD&appInstanceName=default
2014-11	State of Oregon	Yes	ORcycle	http://www.pdx.edu/transportation-lab/orcycle
Not Available	Monterey, CA	No	CycleTracks	http://www.cycletracksmonterey.org/home.html
Not Available	Raleigh, NC	No	CycleTracks	http://www.creativisibility.com/westernblvd/CycleTracks.html

Not Available	Seattle, WA	No	CycleTracks	http://www.psrc.org/transportation/bikeped/cycletrack/
Not Available	Salt Lake City, UT	No	CycleTracks	Not Available
Not Available	Los Angeles, CA	No	CycleTracks	Not Available
Not Available	College Station, TX	Yes	AggieTracks	Not Available
Not Available	Charlottesville, VA	Yes	C-Vill Bike mAPP	http://www.tjpdc.org/cvillebikemapp/
Not Available	Hampton Roads, VA	Yes	Not Available	Not Available

2.5.6 **Cyclist Comfort and Smartphone Data Collection Opportunities**

One of the key advantages of smartphone data is the collection of some user demographic data and Global Positioning System (GPS) data. The collection of GPS points can be matched into segments and intersections of the road and bicycle network. For each trip, detailed paths can be constructed.

As discussed in the previous sections, BLOS and other cyclist comfort evaluation methods rely on data collected or measured in the field. Hence, smartphone detailed route data will not provide data that can be input directly into BLOS methods. However, BLOS methods have been calibrated or estimated in most cases finding statistical relationships between variables that can be measured or observed in the field and users' perceptions of the facilities. Users' perceptions are usually stated preference data and elicited utilizing video or surveys. Many of the videos of facilities and biking conditions may not compare well with Oregon facilities and biking conditions since they are from Florida. The smartphone data is revealed preference data that can be potentially used to

43

calibrate or estimate Oregon specific cyclist comfort evaluation methods based on field

data (Figliozzi, Blanc, and Johnson 2014).

3.   METHODOLOGY

This chapter will review the design elements of the ORcycle smartphone application utilized for collecting the data this thesis analyzed. First, the development process of the application is briefly reviewed. All of the data types collected by the app are then reviewed. Data processing methods used are then summarized. Information about the sampling techniques and final sample used in analysis is then reviewed. Finally, the data analysis methodology is then briefly reviewed.

3.1   ORCYCLE SURVEY TOOL DESIGN

The ORcycle smartphone application was the primary data collection tool in this research. The application was developed to collect cyclist user, trip, and safety data across the State of Oregon. The following sections will review the development of the smartphone application.

3.1.1 Smartphone Application Development

Development of the applications (Android and iOS) and the web server/interface took place primarily between May 2014 and January 2015. The initial public deployment of the application took place in November, with the basic features desired available in all platforms (e.g. GPS recording, survey questions). However, shortly thereafter, upon feedback from users and refinement of our analysis goals, several features were changed slightly or added. These included the addition of a reminder feature, an app tutorial, and minor bug fixes. The major feature change relevant to this thesis was the changing of the

"route comfort" question (see sections 3.1.3.2 and 4.3.3) from optional to mandatory upon completion of a trip. This question was necessary to impute in our analysis if the user did not respond, but after this feature was changed so the question was mandatory (about a month after application release) the question would no longer be necessary to impute. This issue informed some of the limitations of our analysis presented in this thesis, and will likely become less of a problem in the future as the ORcycle sample grows.

The U.S. smartphone market is (as of late 2014[7]) dominated by Android (53%), with iOS comprising 42% of the market and competitors like Microsoft and Blackberry holding the remainder. Android's majority market share and open source development environment encouraged the research team to develop the Android application as the first priority, but it was later deemed feasible to also develop an iOS version as well. With the addition of an iOS version, an estimated 95% of the smartphone market could be reached with the ORcycle application.

---

[7] http://www.statista.com/statistics/266572/market-share-held-by-smartphone-platforms-in-the-united-states/

ORcycle was developed for Android using Eclipse[8], an open-source Android Integrated

Development Environment (IDE). Android software is written primarily in the Java

coding language.
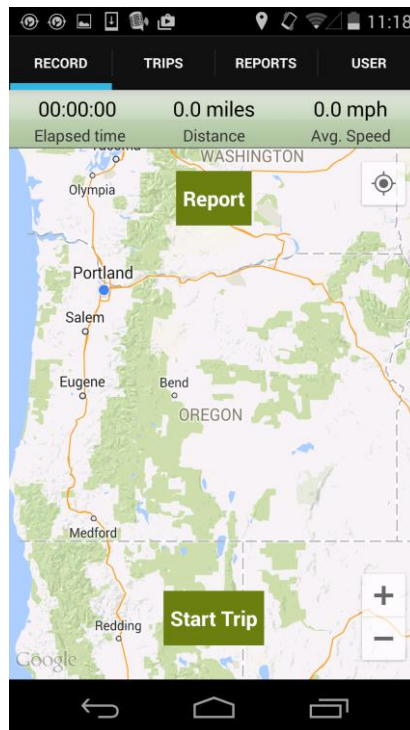
Figure 5 : Home Screen of the Android Version of ORcycle

The Android version of ORcycle was built off of the open-source Android version of

Cycle Atlanta (see section 2.5.3), which was built off of CycleTracks (see section 2.5.1).

The application was re-branded as ORcycle, and then features were modified and added.

---

[8] Eclipse website: https://www.eclipse.org/

The application was tested numerous times on different Android devices throughout the

development and debugging process.

3.1.1.2   *iOS Application*

ORcycle was developed for iOS using XCode, Apple's proprietary IDE. ORcycle was

written primarily in the Objective-C coding language.
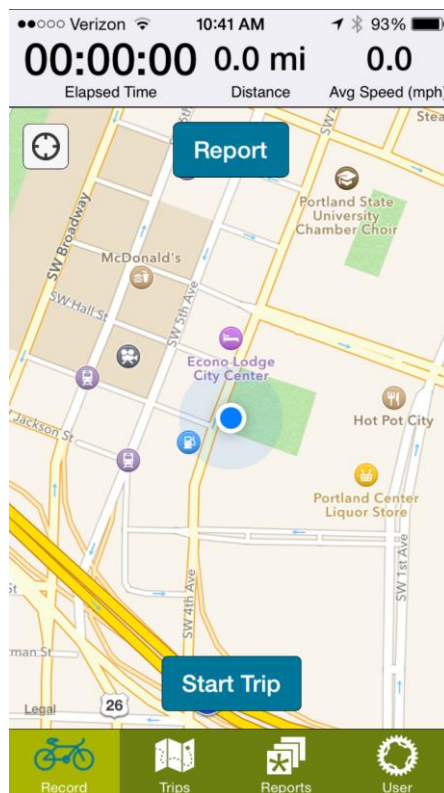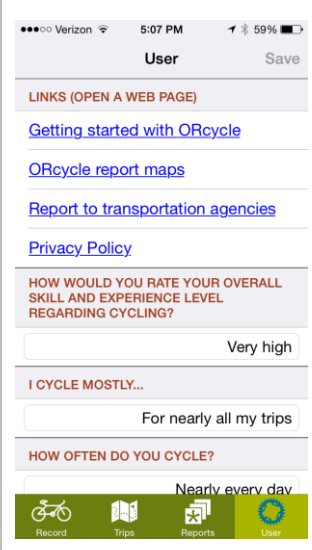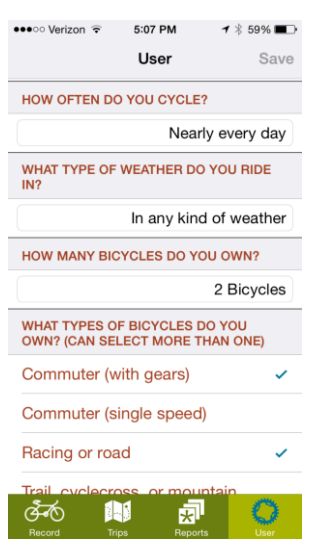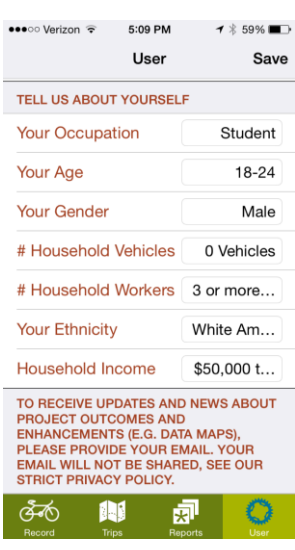
Figure 6: Home screen of the iOS version of ORcycle

ORcycle was built off of the iOS version of RenoTracks (see section 2.5.5), which was

built off of the iOS version of Cycle Atlanta (see section 2.5.3), which was built off of the

iOS version of CycleTracks (see section 2.5.1).

3.1.2 User Attribute Data Collection

User group questions were used to control for differences in behavior and preferences over different user grouping factors. The user questions were asked through the screens presented in Table 14. The user questions utilized in the final version of the application are outlined below. The questions are broken up into two groups: (1) questions about a user's demographics and (2) questions about a user's biking attitude and cyclist type. All user group questions were optional to answer.

Table 14: User Screens (iOS version)

Screen #	1	2	3
iOS			

3.1.2.1 *Demographic Information*

Demographic indicators are often significant covariates with cycling travel behavior (see literature review). The demographic data collected by ORcycle includes age, ethnicity,

gender, household income, occupation, number of household workers, and number of
household vehicles.

Age

Age was considered in all of the previous CycleTracks-derived smartphone applications
and is considered in most travel surveys. Cyclists and smartphone users are both
generally on the younger end of the age spectrum, so it was important to control for this
factor when making inferences from the application results. The age group stratification
used in ORcycle is outlined in Table 15.

Table 15: Age Group Responses (question 1)

Age Category
No data
Less than 18
18-24
25-34
35-44
45-54
55-64
65+

Ethnicity

Ethnicity was another major demographic variable to control for and was considered in
several of the cycling apps as well as in most travel surveys. Cyclists are generally less
diverse than the population at large (Pucher, Buehler, and Seinen 2011; Pucher, Dill, and
Handy 2010; Dill and Voros 2007), and so it was important to control for this factor
when making inferences from the application results. The ethnicity selection categories
used in ORcycle are outlined in Table 16.

Table 16: Ethnicity group responses (question 4)

Ethnicity Category
No data
African American
Asian American
Hispanic
Native American
White American
Other

Gender

Gender was also important to control for and has been included in many of the other cycling applications and in most travel surveys. Bicycling mode share differs considerably by gender, with more males cycling than females on average. The proposed categorization schema for gender selection is outlined in Table 17.

Table 17: Gender group responses (question 3)

Gender Category
No data
Female
Male
Other

Household Income

Taking into account the income level distribution of the survey group was important, as it has been shown that middle to high income groups have so far been more likely to commute by bicycle within the U.S. (Pucher, Buehler, and Seinen 2011). The income category selection was created to match the Oregon Household Activity Survey (OHAS) categories. The proposed categorization schema for income range selection is listed in Table 18.

Table 18: Income group responses (question 6)

Income Category
No data
Less than $14,999
$15,000 to $24,999
$25,000 to $34,999
$35,000 to $49,999
$50,000 to $74,999
$75,000 to $99,999
$100,000 to $149,999
$150,000 or more

Occupation

It was also proposed by the ODOT Technical Advisory Committee (TAC) that

information about rider occupation be collected. The available choices for this question

are outlined in Table 19.

Table 19: Occupation responses (question 5)

Occupation Category
No data
Employed
Student
Retired
Homemaker
Other

Household Workers

Household size is a typical question for travel surveys, as it is often indicative of the

number of trips a household makes. Instead of assessing household size, it was decided in

conjunction with the ODOT TAC that asking for the number of household workers would

be more pertinent. The proposed categorization schema for household workers is listed in

Table 20.

Table 20: Household workers responses (question 7)

Household Workers Category
No data
0 Workers
1 Worker
2 Workers
3 Workers or more

Household Motor Vehicle Ownership

Cycling trip characteristics and preferences likely depend somewhat on the alternative

travel options of the user. Therefore, it was decided that the application document

household motor vehicle ownership. The proposed categorization schema for number of

household vehicles is listed in Table 21.

Table 21: Household vehicles responses (question 8)

Income Category
No data
0 vehicles
1 vehicle
2 vehicles
3 vehicles or more

3.1.2.2   *Cyclist Typology*

Several questions were asked that attempt to evaluate the "type" of cyclist using the

application; see section 2.2 for more information on cyclist typology. As mentioned in

that section, the researchers decided to ask users for several pieces of information related

to their cyclist type, rather than directly asking them to sort themselves into cyclist types.

Questions about bicycle ownership, biking preferences, and biking attitudes were asked.

Bicycle Ownership

Similar to vehicle ownership, it was also proposed that the number of bicycles the rider owns be quantified. This may indicate a user's proclivity towards bicycling. The available choices for this question are given in Table 22.

Table 22: Number of bicycles owned responses (question 9)

Income Category
No data
0 bicycles
1 bicycle
2 bicycles
3 bicycles
4 or more bicycles

Bicycle Types

Knowing a user's bicycle type(s) may reveal information about relationships between facility preferences, user characteristics, and different bicycle types. This question was asked as the following:

What types of bicycles do you own? (can select more than one)

The available responses are listed in Table 23.

Table 23: Bicycle Type (question 10) Responses (select multiple)

Bicycle Type Response
No data
Commuter (with gears)
Commuter (single speed)
Racing or road
Cycle Cross or mountain
Cargo Bike
Recumbent
Other

General Comfort/Ability Level When Cycling

Asking for a user's self-reported general comfort/ability level with riding a bicycle can reveal information about the user's baseline level of comfort, which should be taken into account when analyzing comfort/stress level on specific routes and facilities. A Likert-type scale was used to measure this item. This question was asked as the following:

How would you rate your overall skill and experience level regarding cycling?

The available responses for this question are outlined in Table 24.

Table 24: Cycling Ability responses (question 16)

General Cycling Comfort Category
No data
Very Low
Low
Average
High
Very High

Cycling Dedication

The user's dedication to cycling can elicit information about a user's general attitude about bicycle use. This information will relate to both cycling frequency and trip purposes, which are also asked explicitly. This question is asked as the following:

I cycle mostly …

The available responses for this question are outlined in Table 25.

Table 25: Cycling Dedication responses (question 17)

Cycling Dedication Category
no data
For nearly all my trips
To and from work
For recreation and/or exercise
For shopping, errands, or visiting friends
Mainly to and from work, but occasionally for exercise, shopping, etc.
Other

Weather Tolerance

The user's tolerance for adverse weather is important in calibrating their general

tolerance for external stressors as well as their specific tolerance for weather conditions

while cycling. This question is asked as the following:

What type of weather do you ride in?

The available responses for this item are outlined in Table 26.

Table 26: Weather Tolerance responses (Question 15)

Cycling Dedication Category
no data
In any kind of weather
When it does not rain
Usually warm and dry weather
Only with warm and dry weather

Cycling Frequency

Cycling frequency affects cyclist stress tolerance and thus impacts facility preferences

and route choice (Teschke and Winters 2013). ORcycle collects cycling frequency

information to see how cycling frequency relates to the other information collected. The

cycling frequency question is asked as the following:

How often do you cycle?

The available responses for this question are given in Table 27.

Table 27: Cycling Frequency responses (Question 14)

Cycling Frequency Category
no data
A few times per year
A few times per month
A few times per week
Nearly every day

3.1.3 Trip Data

Table 28: Trip Screens

Screen #	1	2	3	4
iOS				
Description	Users can begin recording a trip by pushing "start trip".	Users can then respond to trip questions including trip purpose, route frequency, and route comfort.		Trips can then be reviewed with summary statistics and saved responses.

3.1.3.1 *Time-Space Trajectory*

Obtaining the time-space trajectories of cyclists utilizing the application was one of the primary objectives of the application. Knowing empirically when and where cyclists chose to ride provides a wealth of revealed preference information about cyclist preferences. These time-space trajectories were obtained using the Android and iPhone devices' built-in GPS units. Within the application, a user can start recording GPS coordinates by pressing the "Start Trip" button on the "Record" screen, as shown in Table 28. This initializes the GPS coordinate recording, which continues until the user indicates that they have finished traveling and/or recording GPS coordinates. For the remainder of the document, this GPS coordinate trajectory will be referred to as a "Trip".

3.1.3.2 *Trip Questions*

These questions are asked after each trip to gain more stated preference information about the user's trip characteristics and preferences. Trip questions included trip purpose, route frequency, route comfort, and route stressors.

Trip Purpose

A majority of cycling advocacy and encouragement focuses on converting motor vehicle trips to bicycle trips for work or school commutes. However, commutes are only one possible trip purpose, and trip purpose likely contributes to a user's bicycle facility preferences or route choice. For example, a study by Haworth and Schramm (Haworth and Schramm 2011) concluded that utilitarian riders were more likely to ride on

sidewalks than other users, and they were also the most likely to utilize separated bicycle paths.

Trip purpose was useful in stratifying trip types to see if there is differentiation in geographic and temporal bicycle travel patterns for different trip purposes. Following the implementation of prior applications, trip purpose can be selected from categories after a trip is completed. The available trip purpose categories, descriptions, and corresponding icons are outlined in Table 29.

Table 29: Trip Purpose (Question 20) Responses, Descriptions, and Icons (select one)

Trip Purpose	Description	Visual Icon
Commute	This bike trip was primarily to get between home and your main workplace.	
School	This bike trip was primarily to go to or from school or college.	
Work related	This bike trip was primarily to go to or from a business related meeting, function, or work-related errand for your job.	
Exercise	This bike trip was primarily for exercise, or biking for the sake of biking.	
Social or Entertainment	This bike trip was primarily for going to or from a social activity, e.g. at a friend's house, the park, a restaurant, the movies.	
Shopping or Errands	This bike trip was primarily to attend to personal business such as buying groceries, banking, a doctor visit, going to the gym, etc.	
Transport Access	The primary reason for this bike trip was to access public transit or some other vehicle (private vehicle, car share, etc.)	
Other	If none of the other reasons applied to this trip, you can enter comments below to tell us more.	

Route Frequency

How often a user rides a route will likely have an effect on their perception of the route. For example, a user will likely be more comfortable with a route if they ride it to work every day, rather than if it is brand new to them. This question is asked as the following:

How often do you ride this route?

The available answers for this question are given in Table 30.

Table 30: Route Frequency (Question 19) Responses (select one)

Route Comfort Response
No data
Several times per week
Several times per month
Several times per year
Once per year or less
First time ever

Route Choice Preferences

Having self-reported route choice characteristics can help in understanding the reasons for route choice from among a set of viable alternatives. This sort of perspective can give bicycle planners greater insight into how route choice decisions are made by bicyclists. This question is asked as the following:

I chose this route because … (can select more than one)

The available responses are listed in Table 31.

Table 31: Route choice preferences (Question 21) responses (select multiple)

Route Preferences Response
No data
It is direct/fast
It has good bicycle facilities
It is enjoyable/has nice scenery
It is good for a workout
It has low traffic/low speeds
It has few intersections
It has few/easy hills
It has other riders/people (I'm not alone)
I do not know/have another route
I found on my phone/online
Other (indicate in comments)

Route Comfort

Route comfort is meant to be analogous to Level of Traffic Stress. Route comfort is an

ordinal, Likert type rating of a user's self-reported comfort on a route. It is the dependent

variable modeled over many other independent variables in section 4.4.2. This question is

asked as the following:

In terms of comfort, this route is…

The available responses for route comfort are given in Table 32.

Table 32: Route Comfort (Question 22) Responses (select one)

Route Comfort Response
No data
Very bad (unacceptable for most riders)
Bad (only for confident riders)
Average
Good (for most riders)
Very Good (even for families/children)

It is important to know which characteristics of a rider's route may have caused them to feel some level of traffic stress. This question is asked as the following:

 Along this route, you are concerned about conflicts/crashes with… (can select more than one)

The available responses are listed in Table 33.

Table 33: Route stressors (Question 27) responses (select multiple)

Route Stressors Response
Not concerned
Auto traffic
Large commercial vehicles (trucks)
Public transport (buses, light rail, streetcar)
Parked vehicles (being doored)
Other cyclists
Pedestrians
Other

Custom Additional Details

Having an additional details entry gave users a place to write something specific about their trip that may not be described by the trip questions available.

3.1.4 Crash and Safety Issues Reports

The ability to record "issues" and "assets" (referred to as "notes") was one of the most significant improvements to Cycle Atlanta (as discussed in 2.5.3) This functionality combines the uses of a bicycle trip tracking application like CycleTracks (section 2.5.1) with the infrastructure crowdsourcing functionality of applications like Citizens Connect and PDX Reporter (Figliozzi, Blanc, and Johnson 2014).

It was decided that ORcycle would remove the asset recording functionality (deemed to
be of minimal value to ODOT) in favor of recording crash events and
infrastructure/safety issues. We chose to call these data objects "Reports" instead of
"Notes". There were two types of reports: (1) crash or near-crash events and (2) location
specific infrastructure/safety issues.

Both types of reports were uploaded with a location, which could be submitted as either
the user's current location or a custom location selected on a map. Reports were also
uploaded with a date, which could either be the current date or a custom-selected date.

3.1.4.1 *Crash or Near-Crash Events*

Crash event reports were submitted using the screens shown in Table 34. Crash event
reports asked four mandatory questions: (1) crash severity, (2) vehicle or object related to
event, (3) crash event actions, and (4) crash event reasons.

Table 34: Crash Report Screens

Screen #	1	2	3	4
iOS				

Crash Event Severity

When documenting a crash event report, this was the first question to be answered. The user could indicate the relative severity of their crash event. The question was asked as the following:

Severity of the crash event: (choose one)

The available answers for this question are given in Table 35.

Table 35: Crash event severity (Question 28) responses (select one)

Severity Category	Report Icon
Major injuries (required hospitalization)	
Severe (required a visit to ER)	
Minor injury (no visit to ER)	
Property damage only (bicycle damaged but no personal injuries)	
Near-miss (no damage or injury)	

Vehicle or Object related to the event

We also asked the user what transportation mode or physical object they may have had a crash or conflict with. This question was asked as the following:

Vehicle or object related to the event… (can select more than one)

The available answers for this question are given in Table 36.

Table 36: Vehicle or object (Question 29) responses (select multiple)

Vehicle or object category
Small/medium car
Large car/Van/SUV
Pickup truck
Large commercial vehicles (trucks)
Public transport (buses, light rail, streetcar)
Another bicycle
Pedestrian
Pole or fixed object
Cyclist fell (or almost fell)
Other

Crash Event Actions

The user also reported what particular traffic movements led to the crash event they experienced. The corresponding question was asked as the following:

Actions related to the event… (can select more than one)

The available answers for this question are given in Table 37.

Table 37: Crash event actions (Question 32) responses (select multiple)

Vehicle or object category
Right-turning vehicle
Left-turning vehicle
Parking or backing up vehicle
Person exiting a vehicle
Cyclist changed lane or direction of travel
Vehicle changed lane or direction of travel
Cyclist did not stop
Driver did not stop
Cyclist lost control of the bike
Other

Crash Event Reasons

The user also reported what environmental, traffic, or personal conditions may have

contributed to the crash event. The corresponding question was asked as the following:

What contributed to the event? (can select more than one)

The available answers for this question are given in Table 38.

Table 38: Crash event reasons (Question 33) responses (select multiple)

Vehicle or object category
Debris or pavement quality
Poor lighting or visibility
Cyclist was outside the bike lane or area
Vehicle entered the bike lane or area
Cyclist did not follow stop sign or red light
Vehicle did not follow stop sign or red light
Cyclist did not yield
Vehicle did not yield
Cyclist was distracted
Careless driving or high vehicle speed
Other

3.1.4.2   *Location Specific Infrastructure/Safety Issues*

Crash event reports were submitted using the screens shown in

Table 39. Crash event reports asked four mandatory questions: (1) issue type and (2) issue urgency.

Table 39: Issue Report Screens

Screen #	1	2	3	4
iOS				

Issue Type

The first question asked when a user reported a "safety/infrastructure issue" was a description of the issue type. This question was asked as the following:

Location specific infrastructure/safety issues… (can select more than one)

The available "issue types" for documentation are given in Table 40.

Table 40: Issue Type (Question 30) responses (select multiple)

Issue Type
Narrow Bike Lane
No bike lane or separation
High vehicle speeds
High traffic volumes
Right/left turning vehicles
Traffic signal timing
No traffic signal detection
Truck traffic
Bus traffic/stop
Parked vehicles
Pavement condition
Other

Urgency

When documenting a safety/infrastructure issue report, the user was asked to indicate the urgency level of the location specific infrastructure or safety issue. The user could indicate the relative urgency of the issue on a scale of 1 to 5. The question was asked as the following:

Urgency of the problem: (choose one)

The available answers for this question are given in Table 41.

Table 41: Issue urgency (Question 31) responses (select one)

Severity Category	Report Icon
1 (not urgent)	
2	
3 (somewhat urgent)	
4	
5 (urgent)	

3.2   DATA PROCESSING

The majority of data cleaning and processing took place in the R Project for Statistical Computing environment. Where mentioned, some other software or coding environments were utilized.

For modeling purposes, missing survey responses were filled in using a multiple imputation algorithm from the R package missForest[9]. The multiple imputation algorithm utilized a case's (which could be a user, trip, or report) other survey responses to predict the response to the missing question based on the response distributions of other cases in the overall sample. Filling in the missing responses allowed users, trips, and reports that were missing optional responses to still be used in statistical models.

---

[9] http://cran.r-project.org/web/packages/missForest/index.html

70

3.2.1 **Users**

Users filled out optional survey questions about their demographic characteristics and cycling preferences/attitudes. Missing survey responses were imputed for the statistical modeling in section 5, but the sample description in section 4 describes the raw data.

3.2.2 **Trips**

Trip data came in two groups: the GPS coordinate traces and the survey question responses. Raw GPS coordinates would not allow us to connect transportation link characteristics with the cyclists' routes, so the coordinates needed to be geo-matched to a network. The GPS coordinate traces were matched to the Portland metropolitan area bicycle and street network where possible. An example of the difference between a geo-matched route and its raw GPS coordinates is given in Figure 7, where a cyclist crossed the Willamette River in Portland, OR using the lower deck of the Steel Bridge and proceeded onto the Eastbank Esplanade. This kind of detail in the path of the cyclist would be difficult to entangle without an accurate and topologically correct cycling network and a robust geo-matching script.

Figure 7: Example of raw GPS coordinates (red circles) compared with matched route (red polyline)

The network used for geo-matching was a modified version of Metro's bicycle and street

network improved by John Gliebe and Joseph Broach[10] in 2012[11] to include additional

links in the network utilized by bicyclists. Geo-matching was carried out using a group of

Python scripts developed for the bicycle GPS study conducted by Jennifer Dill and John

Gliebe at Portland State University in 2007 (Broach, Dill, and Gliebe 2012) and slightly

---

[10] Joseph Broach; E-mail: jbroach@pdx.edu

[11] There are new bicycle network links in the Portland area not included in this network, and so in some cases these links could not be matched to the correct link (resulting in matching to a nearby link).

modified to work with ORcycle's data structure. The scripts were developed based on

algorithms from Schuessler and Axhuasen (Schuessler, Axhausen, and Zurich 2009;

Schuessler and Axhausen 2009) and took several factors into account:

1. Proximity of the GPS points to eligible links in the network
2. Topological connectivity of the precedent and antecedent links in the bicycle network
3. Rejection of "spurious" u-turns

More in-depth description of the geo-matching algorithm is available in (Dill and Gliebe

2008). Only trips that took place within Metro's jurisdiction could be matched to the

bicycle network, and only trips that met filtering criteria (minimum trip length, maximum

trip speed) were left in the final match set. Of 780 potential trips, only 616 (79%) were

geo-matched given the above criteria. The 616 geo-matched trips were made by 148

unique users.

In addition to the GPS trace, users filled out mandatory and optional survey questions

after recording a trip. Missing survey responses were imputed for the statistical modeling

in section 5, but the sample description in section 4 uses the raw data.

3.2.3 **Reports**

Report data came in three primary groups: location, survey question responses, and a

photo. A report was located either using the GPS location of the smartphone device or by

allowing users to optionally input a custom location by panning and zooming to the

position of the report on an interactive map. Users filled out mandatory and optional

survey questions after recording a report. Users could also optionally upload a photo

along with their report, which could potentially provide researchers and planners with more information about the nature of the reported problem.

3.3   SAMPLE SELECTION

ORcycle was publicized shortly after its release through internet and e-mail campaigns led by ODOT and the project PI. Due to time constraints, only users created and trips/reports recorded between the initial deployment (November 3rd, 2014) and March 31st, 2014 were used for analysis within this thesis. However, the application is still collecting data, so the conclusions herein only apply to this specific sample of users, trips, and reports.

3.4   DATA ANALYSIS

Data analysis also took place primarily within the R coding environment. Data could be pulled directly from the remote MySQL database where it is securely stored and then statistics and spatial analyses could be automated and conducted repeatedly as new information flowed into the SQL database from users.

4.   SAMPLE CHARACTERISTICS

4.1   USER CHARACTERISTIC DATA

Users were asked several optional questions that they could answer upon first opening the application or anytime thereafter. The questions consisted of two main groups: one group evaluated a user's bicycling attitudes, and the other documented a user's demographic characteristics.

Upon downloading ORcycle, each installation was given a unique "user" identity. Associated with that user identity were the responses to all the user-related survey questions explored below. The user sample considered herein included users that were "created" (i.e. downloaded the application and uploaded at least one trip or report) between the application release on November 1st, 2014 and March 31st, 2015. User creation rates and the cumulative number of users created over the study period are graphed in Figure 8 and Figure 9. There was an initial surge in user participation just after the application release with 226 users by December 1st, but the number of new users slowed to a nearly constant rate (~1.4 users per day) of creation within a month of the release. There were a total of 381 users in the sample considered herein.
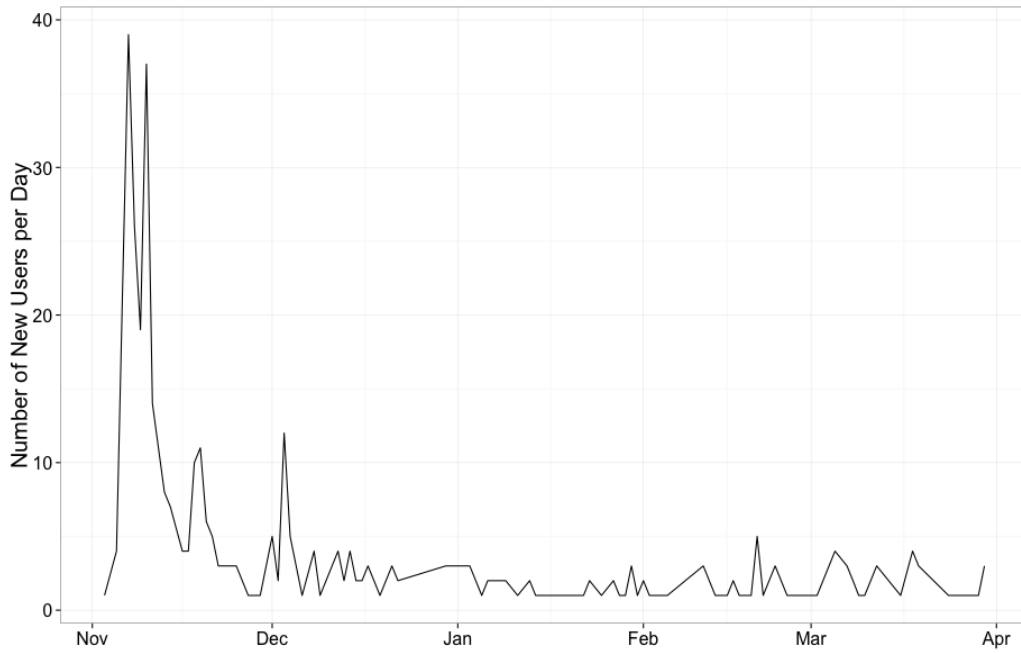
Figure 8: Users created per day during study period
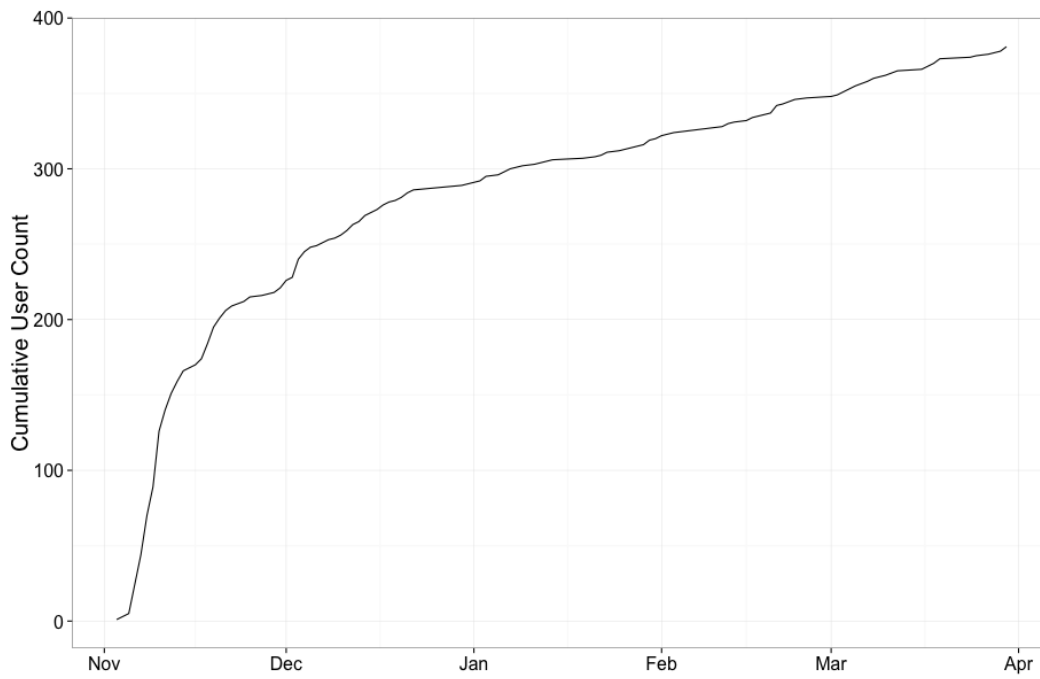
Figure 9: Cumulative user count over study period

4.1.1 **Smartphone platform**

Users could download and operate the ORcycle application for either iOS (e.g. iPhone) or Android (e.g. Samsung Galaxy, Google Nexus) operating system platforms. When a user downloaded ORcycle and submitted at least one piece of data (either a trip or a report), we could differentiate between those using Android or iOS devices. Figure 10 indicates that the majority of users (67%) used ORcycle on Android devices. The U.S. smartphone market is (as of late 2014[12]) marginally led by Android (53%), with iOS comprising 42% of the market and competitors like Microsoft and Blackberry comprising the remainder of smartphone devices. Among the initial sample of users of ORcycle, the proportion of Android users was higher than the nationwide market average.
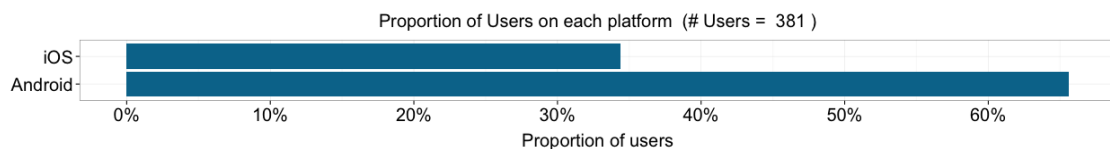
Figure 10: User distribution by platform

4.1.2 **Age**

Users were asked to indicate which age group they belonged to from among seven options. Age category distribution within the sample is illustrated in Figure 11. Within

---

[12] http://www.statista.com/statistics/266572/market-share-held-by-smartphone-platforms-in-the-united-states/

the sample, the majority of users (52%) are between 25 and 44. There was a negligible

amount of under-18 users. 17% of users chose not to provide information about their age.
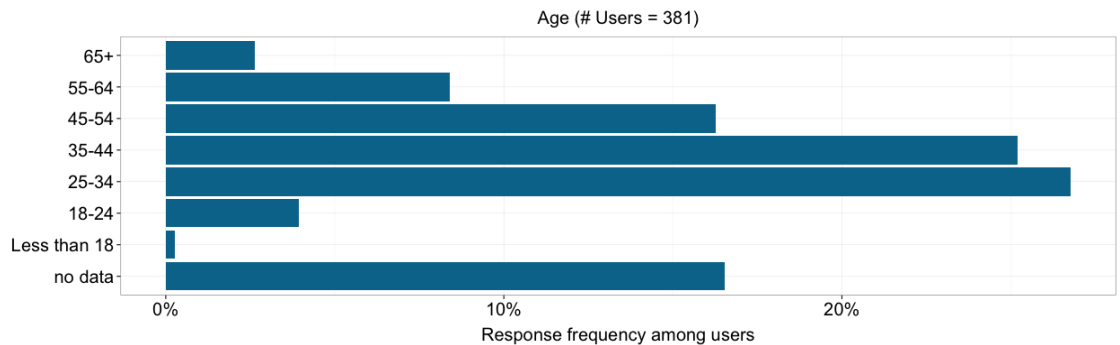
Figure 11: Age distribution of users

4.1.3 **Gender**

Users were asked to indicate which gender group they identified with from among three

options. Gender distribution among the user sample is illustrated in Figure 12. 68% of

users identified as males and 15% as females. These results indicated a sample bias

towards males, which is typical for studies of cycling behavior (see literature review).

17% of users declined to provide information about their gender group.
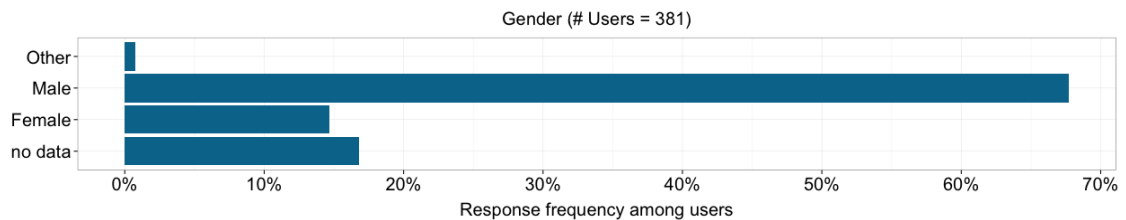
Figure 12: Gender distribution among users

4.1.4 **Ethnicity**

Users were asked to indicate which ethnic group they identified with from among six

options. The ethnicity distribution among the user sample is illustrated in Figure 13. 70%

of users identified as "White American", with less than 5% each of the other available

ethnicity categories. 20% of users declined to provide information about their ethnicity.

Portland (where many of the users are located – see sections 4.3.6, 5.3.1.3, and 4.4.1.5)

has a substantial white population (76% in 2010[13]). Oregon also has a large white

population (84% in 2010). Though cycling studies are typically biased towards white

demographics (see literature review), the proportion of ORcycle users that are white

seems to be in order with the ethnicity makeup of Portland and Oregon.
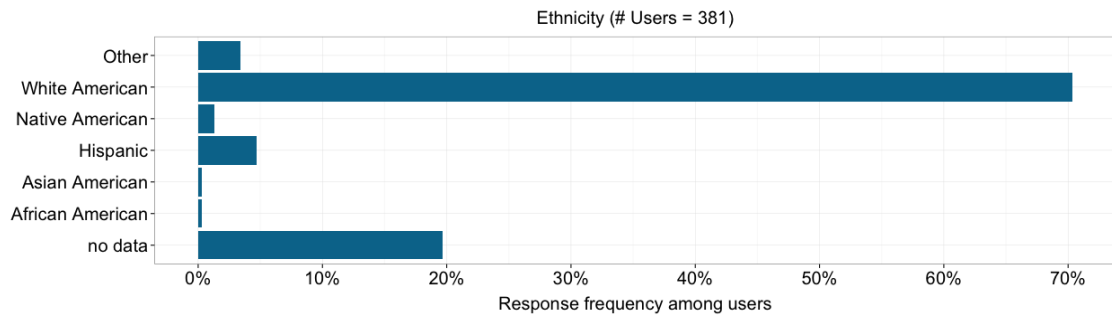
Figure 13: Ethnicity distribution among users

4.1.5 **Occupation**

Users were asked to indicate their occupation from among five choices. The occupation

distribution among the user sample is illustrated in Figure 14. 68% of users indicated that

they were employed and 8% of users indicated they were students. 18% of users declined

to provide information about their occupation.

---

[13] http://quickfacts.census.gov/qfd/states/41/4159000.html

Occupation (# Users = 381)
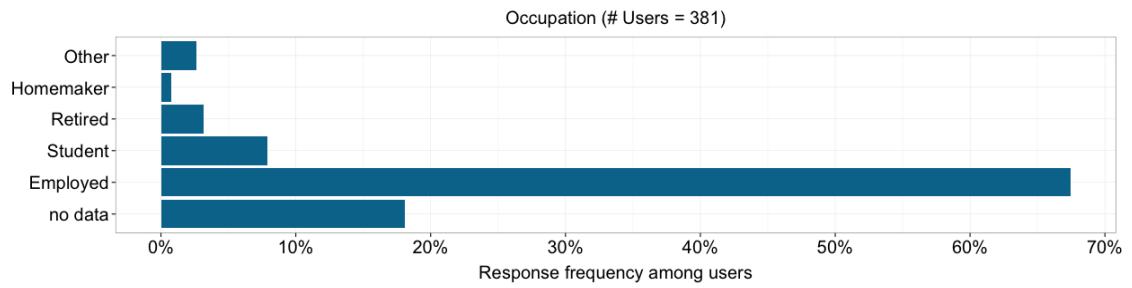
Figure 14: Occupation distribution among users

4.1.6 Household Income

Users were asked to indicate which income group their household fell into from among

eight options. The household income group distribution among the user sample is

illustrated in Figure 15. The majority of users fell into the middle to high-income

categories. This indicates a potential sample bias towards higher income households.

25% of users declined to provide information about their household income.

Household Income (# Users = 381)
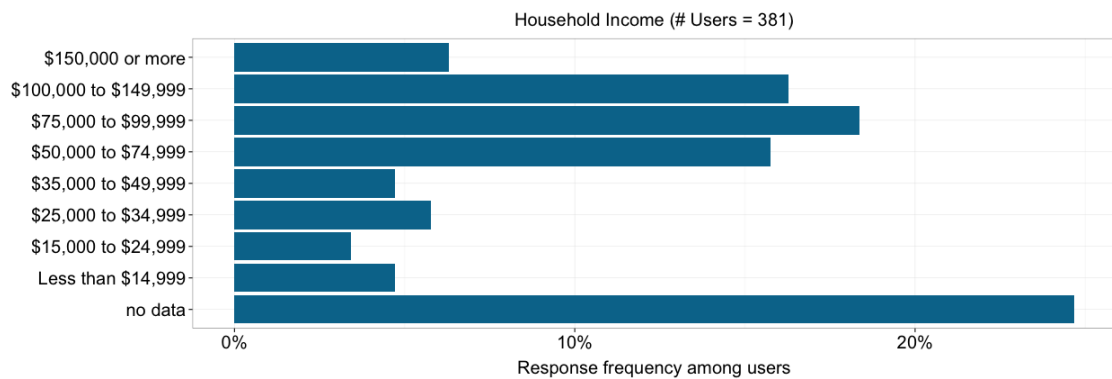
Figure 15: Household income distribution among users

4.1.7 Household Workers

Users were asked to indicate the number of workers in their household from among four

options. The household worker category distribution among the user sample is illustrated

in Figure 15. The majority of users (72%) indicated that they lived in one or two worker

households. 18% of users declined to provide information about the number of workers in

their households.

Household Workers (# Users = 381)
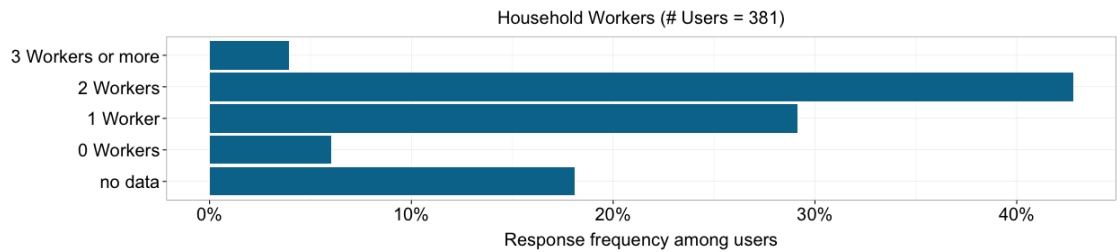
Figure 16: Household workers distribution among users

4.1.8 Household Vehicles

Users were asked to indicate the number of vehicles their household owned from among

four categories. The household vehicle category distribution among the user sample is

illustrated in Figure 17. The majority of users (64%) indicated that they lived in one or

two vehicle households. A substantial proportion of users (12%) indicated that they lived

in zero vehicle households, which may indicate captive users or a sample bias towards

those very invested in a "bicycling lifestyle". 16% of the sample declined to provide

information about the number of vehicles owned in their household.

Household Vehicles (# Users = 381)
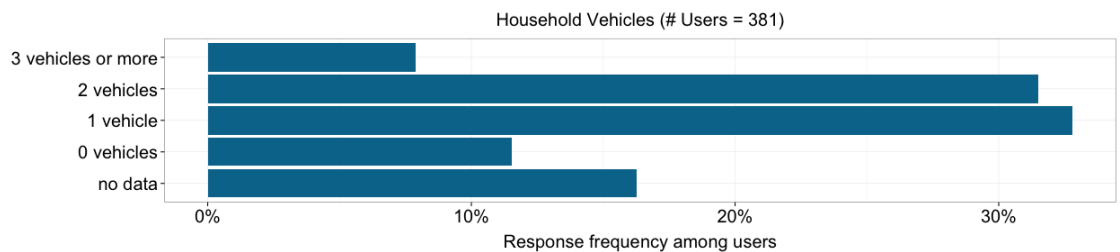
Figure 17: Household vehicles distribution among users

4.1.9 **Household Workers to Vehicles Ratio**

The number of household vehicles was divided by the number of household workers to calculate a vehicles/workers ratio. This ratio could be used as an indicator of the vehicle accessibility within a household. The mean vehicles/workers ratio was close to one, but there were a number of users with ratios below one (104 users). The distribution of the vehicles/workers ratio is summarized in Table 42 and Figure 18.

Table 42: Vehicles/Workers Distribution Summary

Statistic	N	Mean	St. Dev.	Min	Max
Vehicles/Workers Ratio	314	1.054	0.579	0.250	4.000

Figure 18: Vehicles/workers ratio distribution among users

4.1.10 Number of Bicycles

Users were asked to indicate the number of bicycles that they personally owned from among five options. The number of bicycles distribution among the user sample is illustrated in Figure 19. Proportions among the choices were fairly evenly spread (with the exception of those who owned zero bicycles). 15% of users declined to provide this information.
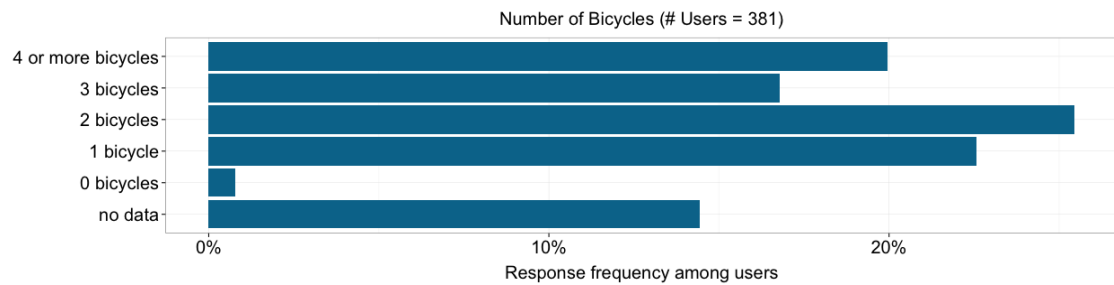
Figure 19: Number of bicycles among users

4.1.11 Bicycle Types

Users were asked to indicate the types of bicycles that they owned from among seven options, with the ability to select multiple choices. The bicycle type distribution among the user sample is illustrated in Figure 20. 61% of the sample indicated they owned a commuter bicycle (with gears), while 39% of the sample indicated they owned a racing/road bike and/or a trail/cyclocross/mountain bike. 18% of the sample indicated they owned other types of bicycles not available within the selection set. 15% of the sample declined to provide any information about their bicycle types.

Bicycle Types (# Users = 381)
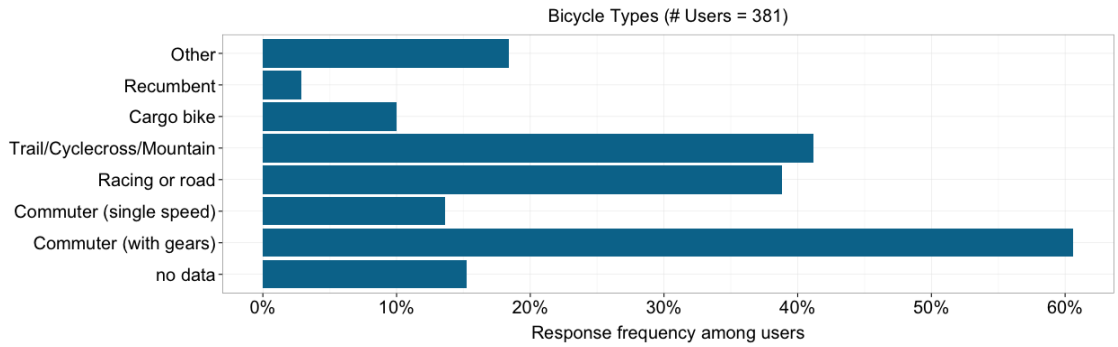
Figure 20: Bicycle type distribution among users

4.1.12 Cycling Frequency

Users were asked to indicate how often they ride a bicycle from among four choices. The cycling frequency distribution among the user sample is illustrated in Figure 21. 50% of users indicated that they bike "nearly every day" while 22% of users indicated that they biked "a few times per week". This indicates there may be sample bias towards frequent cyclists. 15% of users declined to provide information about their cycling frequency.

Cycling Frequency (# Users = 381)
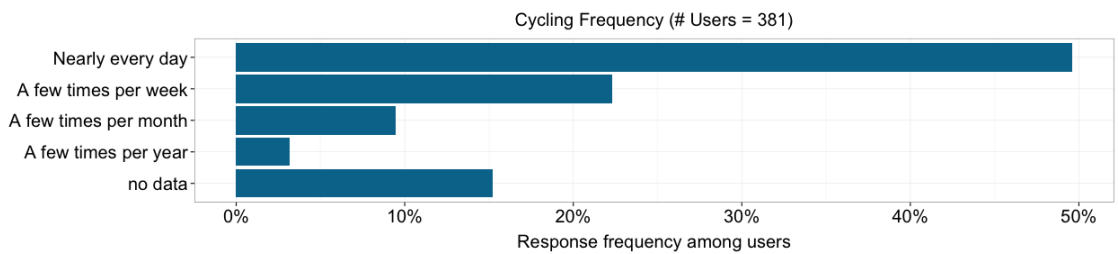
Figure 21: Cycling frequency distribution among users

4.1.13 Preferred Cycling Weather

Users were asked to indicate their preferred cycling weather from among four choices. The cycling weather distribution among the user sample is illustrated in Figure 22. The majority of users (67%) indicated that they would bicycle "In any kind of weather". This

may indicate a bias towards more "serious" cyclists. 14% of users declined to provide information about their preferred cycling weather.
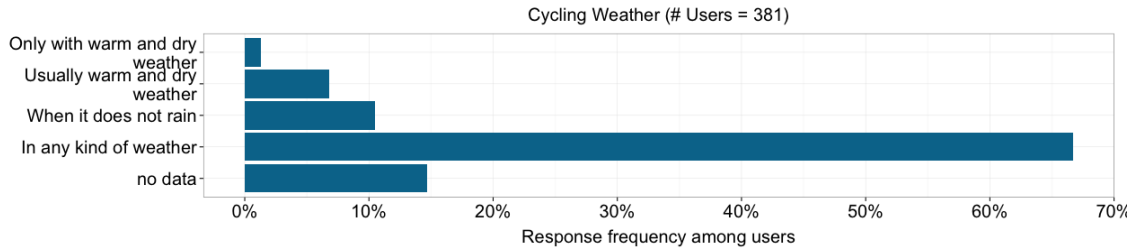
Figure 22: Preferred cycling weather distribution among users

4.1.14 Cycling Ability

Users were asked to indicate their cycling ability from among five choices. The cycling ability distribution among the user sample is illustrated in Figure 23. 33% of users indicated they had "Very High" cycling abilities 32% indicated they had "High" cycling abilities. Less than 2% of users indicated they had "Low" or "Very Low" cycling abilities. These results indicate a sample biased towards more skilled and/or experienced cyclists. 17% declined to provide information about their cycling ability.
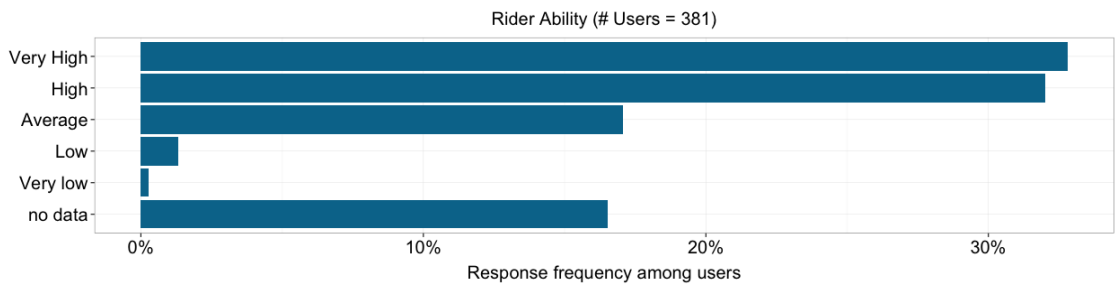
Figure 23: Cycling ability distribution among users

4.1.15 Rider Type

Users were asked to indicate why they rode a bicycle from among six choices. The rider type distribution among the user sample is illustrated in Figure 24. 28% of users indicated that they rode a bicycle "For nearly all my trips" and 19% of users indicated that they rode a bicycle "To and from work". 15% of users declined to provide information about what sort of bicycle rider they are.
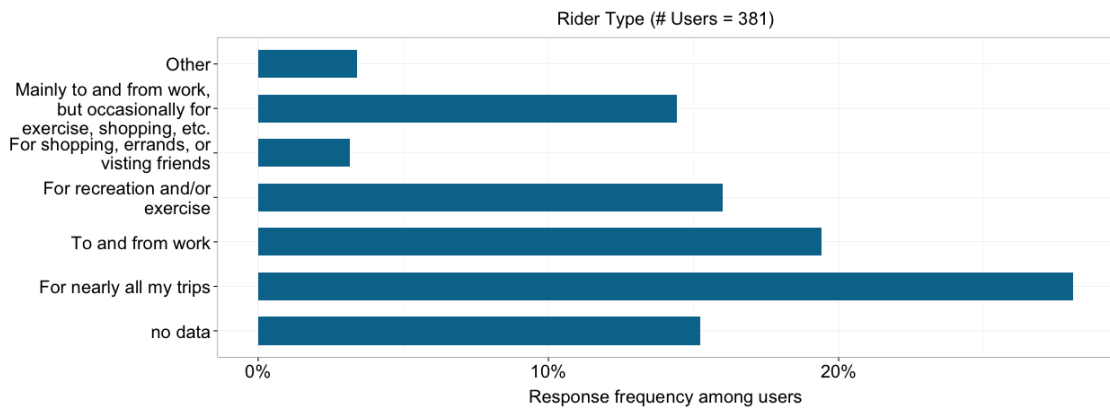
Figure 24: Rider type distribution among users

4.2 USER SAMPLE BIAS

Where possible, the ORcycle sample was compared with the Oregon Household Activity Survey (OHAS) sample to detect statistically significant differences. The OHAS sample is assumed to be more representative of the Oregon cycling and general population due to a more rigorous sampling methodology. While the OHAS sample could still err from a "true" representation of the Oregon population, comparing the two samples can still help to estimate where biases exist in the ORcycle sample and how large they are. The entire OHAS sample and a subsample of bicycle commuters were compared against to gauge

ORcycle's sample bias. The full tables comparing the two samples are supplied in

Appendix 9.1. The results of the chi-square comparisons are presented below in

Table 43.

Table 43: Chi-square testing of user sample bias

Demographic Characteristic	Reference Table	ORcycle vs. OHAS Bike Commuters			ORcycle vs. OHAS Entire Sample		
		Chi-Square	DF	Significance	Chi-Square	DF	Significance
Age	Table 65	89.4	6	p<0.001	592	6	p<0.001
Gender	Table 66	28.4	1	p<0.001	157	1	p<0.001
Ethnicity	Table 67	33.3	5	p<0.001	47.5	5	p<0.001
Household Income	Table 68	15.5	7	p<0.05	57.6	7	p<0.001
Household Workers	Table 70	61.4	3	p<0.001	67.9	3	p<0.001
Household Vehicles	Table 69	39.5	3	p<0.001	123	3	p<0.001

All of the tests resulted in statistically significant differences, though some had greater

differences than others (as gauged by the chi-square statistic). However, the ORcycle

sample was less different from the OHAS bike commuter sample than it was from the

entire OHAS sample, which indicates that ORcycle was reaching Oregon's cycling

population to some degree.

4.3   TRIP DATA

Trip data came in two distinct types: the GPS coordinate trace of the trip and the

responses to the post-trip survey questions. All the trips considered herein were logged

between the application release on November 1st, 2014 and March 31st, 2015. The trip

recording rate and the cumulative number of trips recorded are graphed in Figure 25 and

Figure 26. As with user creation, there was an initial surge in trip recording following the

release of the app, but trip recording activity leveled off to a slower nearly constant rate

by the end of December 2014. Overall, the average trip-recording rate was 5.6 trips per

day. 780 trips are considered in the following sample description.
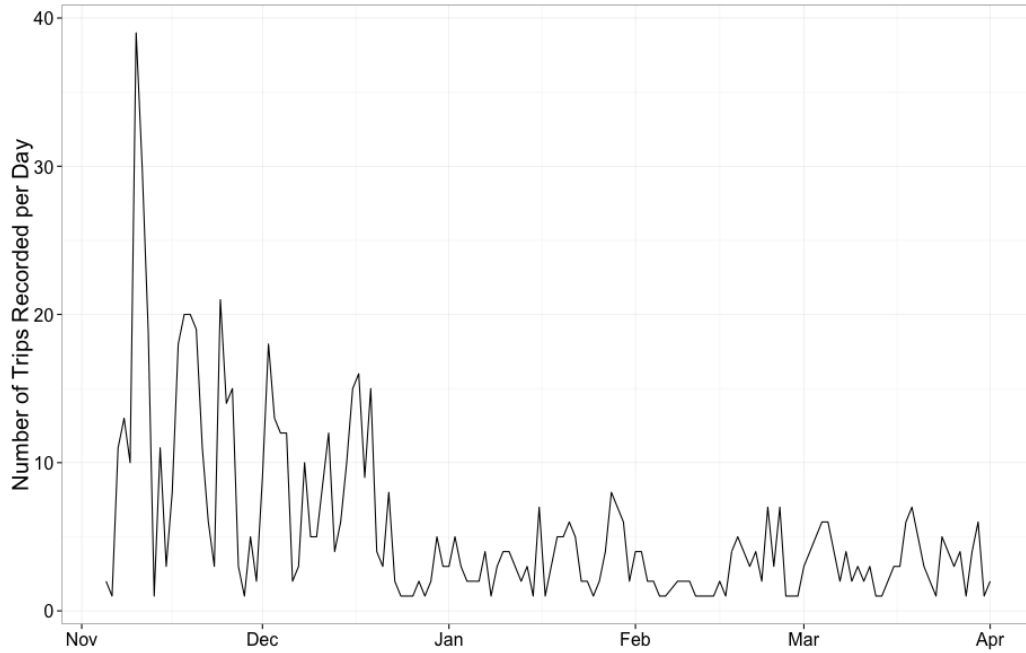
Figure 25: Rate of trip recording over study period
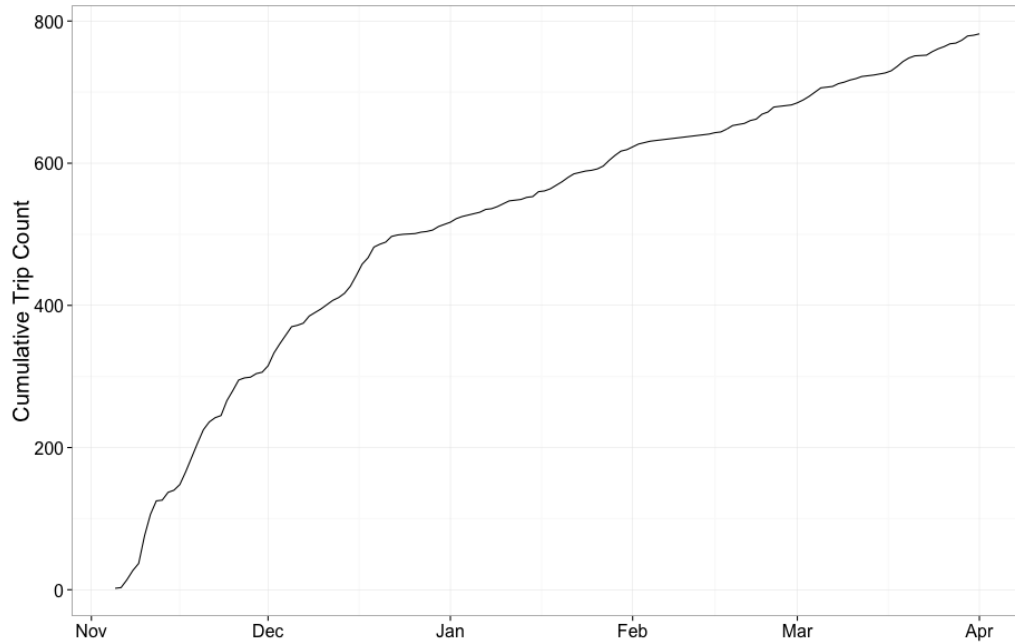
Figure 26: Cumulative number of trips recorded over study period

4.3.1 Trip Purpose

Users were asked to indicate for each trip they recorded what their primary trip purpose was from among eight choices. This question was mandatory upon recording a trip. The trip purpose distribution among the trip sample is illustrated in Figure 27. 55% of trips were indicated to be commuting trips, with the next highest category being "shopping/errands" at 14%. This indicates that most of the trips in the sample were taken for utilitarian purposes.
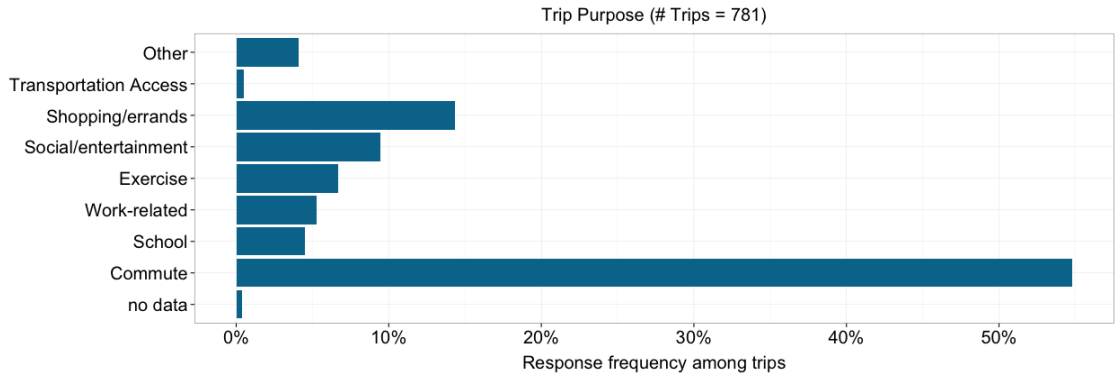
Figure 27: Trip Purpose Distribution among Trips

4.3.2 Route Frequency

Users were asked to indicate for each trip they recorded how often they rode that particular route from among six choices. This question was mandatory upon the recording of a trip. The route frequency distribution among the trip sample is illustrated in Figure 28. 47% of the routes taken on trips were indicated as being ridden "several times per week" by the user. Other trips were indicated to be ridden several times per month (22%) and several times per year (18%).
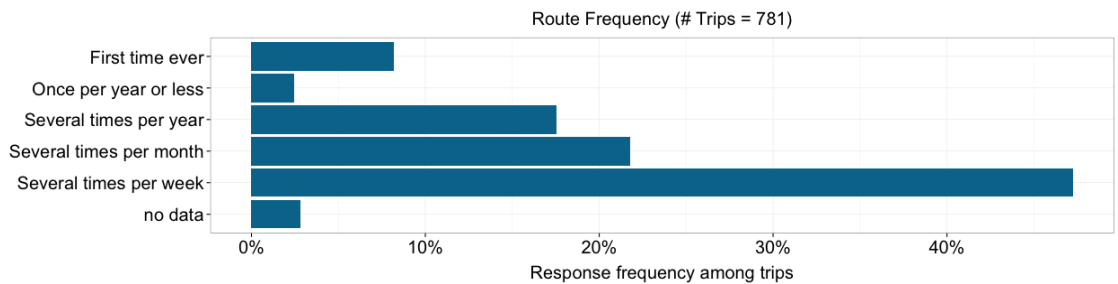
Figure 28: Route Frequency Distribution among Trips

4.3.3 Route Comfort

Users were asked to indicate how comfortable they were with the route they had taken upon finishing recording each trip. This question was mandatory. The route comfort

90

distribution among the trip sample is illustrated in Figure 29. 29% of trips did not have an indicated comfort level (user declined to provide this information). 24% of trips were indicated to have an "average" comfort level, while 28% of trips were indicated to have a "Good (for most riders)" comfort level.
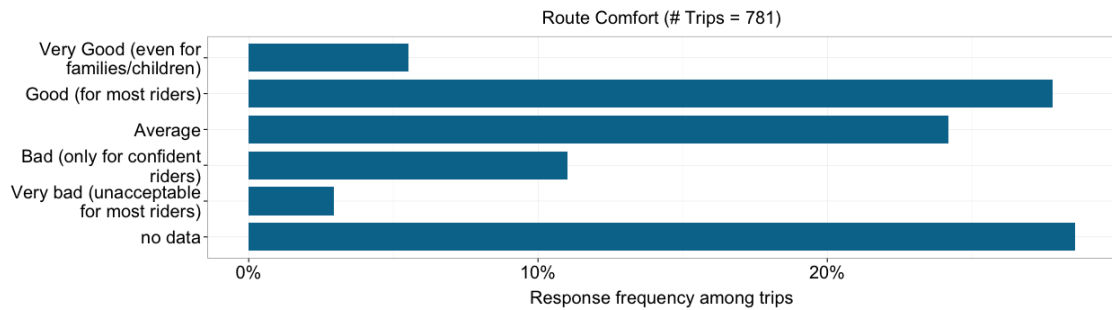
Figure 29: Route Comfort Distribution among Trips

4.3.4 Route Preferences

Users were asked to indicate why they chose their particular route for each trip they recorded. This question was mandatory and could have been answered with multiple responses from among the twelve available responses. The route choice preferences distribution among the trip sample is illustrated in Figure 30. 59% of trips were indicated to have been taken on routes that were chosen because they were "direct/fast". Other popular choices were "It has good bicycle facilities" (37% of trips), and "It has low traffic/low speeds" (30% of trips).
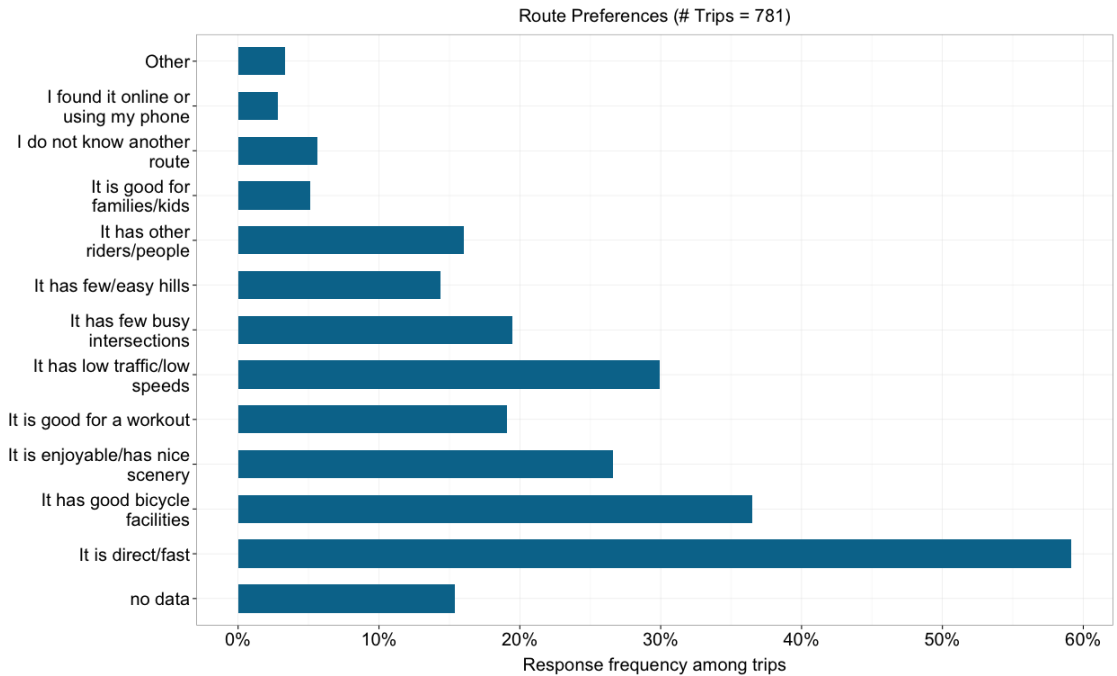
Figure 30: Route Preferences Distribution among Trips

4.3.5 Route Stressors

Users were asked to indicate what objects or other transportation modes they were concerned about conflicts with along the route they had ridden for their recorded trip. This question was optional. The route stressors distribution among the trip sample is illustrated in Figure 31. 16% of trips did not have any route stressors indicated (users declined to provide this information). On 57% of trips, users indicated that they were concerned about conflicts with auto traffic. Other high categories of concern included large commercial vehicles (27%) and parked vehicles (32%).
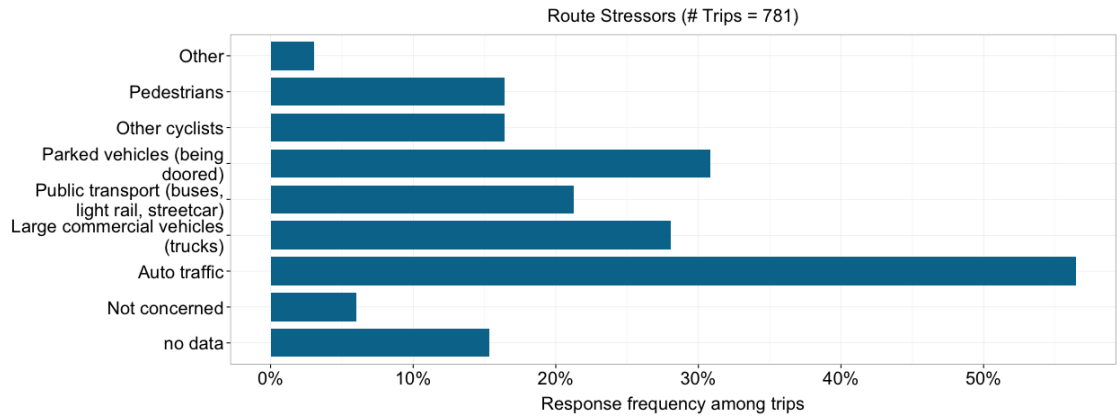
Figure 31: Route Stressors Distribution among Trips

4.3.6 **Geography**

Geographical analysis was used to determine which state, county, and city the majority of the coordinates of a trip fell inside. The geographic distribution of trips among states is illustrated in Figure 32. 98% of the trips took place within the state of Oregon. This was expected, since the application was marketed to Oregon users.
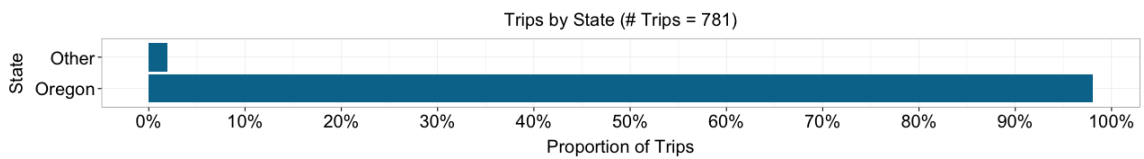
Figure 32: State Distribution among Trips

The geographic distribution of trips among counties is illustrated in Figure 33. 80% of the trips were taken within Multnomah County.
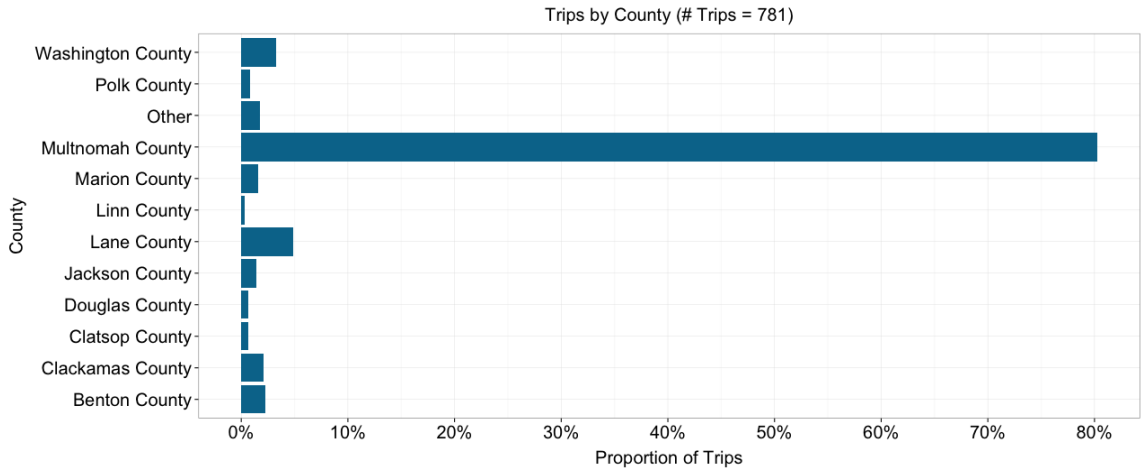
Figure 33: County Distribution among Trips

The geographic distribution of trips among cities is illustrated in Figure 34. 80% of the

trips were taken within the city of Portland.
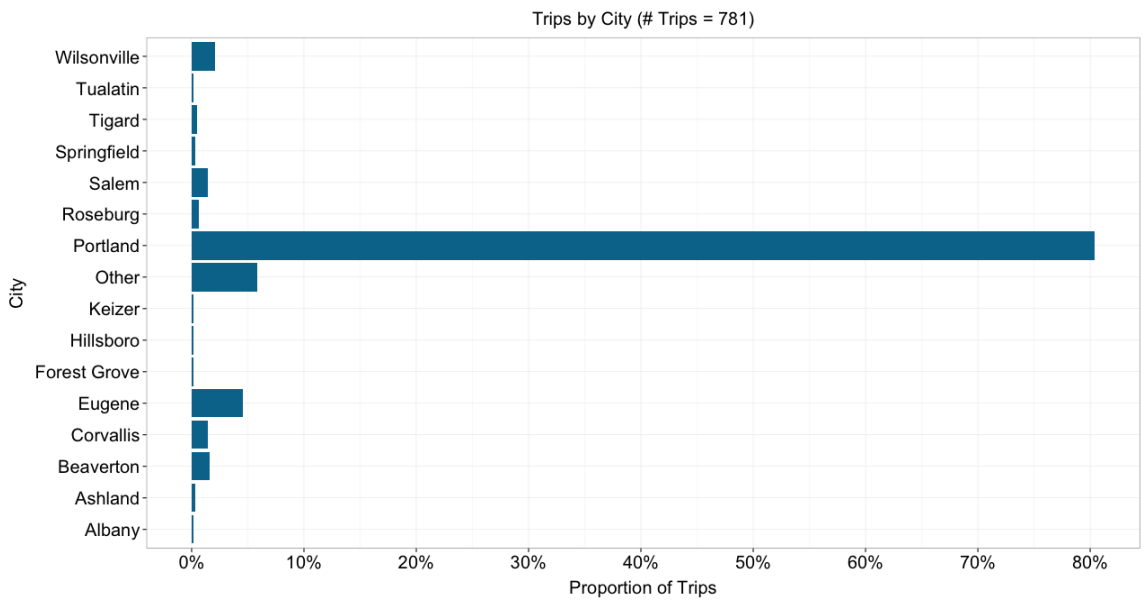
Figure 34: City Distribution among Trips

The high concentration in Multnomah County and the City of Portland indicates a bias in

the sample towards Portland area users.

4.3.7 **Trip Statistics**

Basic statistics for trip times and distances were calculated and separated by trip purpose. Several boxplots are presented below, where the solid black line in the middle of the box indicates the median value, the box itself indicates the inter-quartile range, and the dotted lines indicate the overall range excluding outliers; which are indicated as open circles.

Trip duration was first calculated, with distributions varying substantially among different trip purposes. These distributions are presented in Figure 35. The overall median trip time was 29 minutes. Exercise trips had the highest median trip duration with 57 minutes, while transit access trips had the lowest median trip duration with 11 minutes.
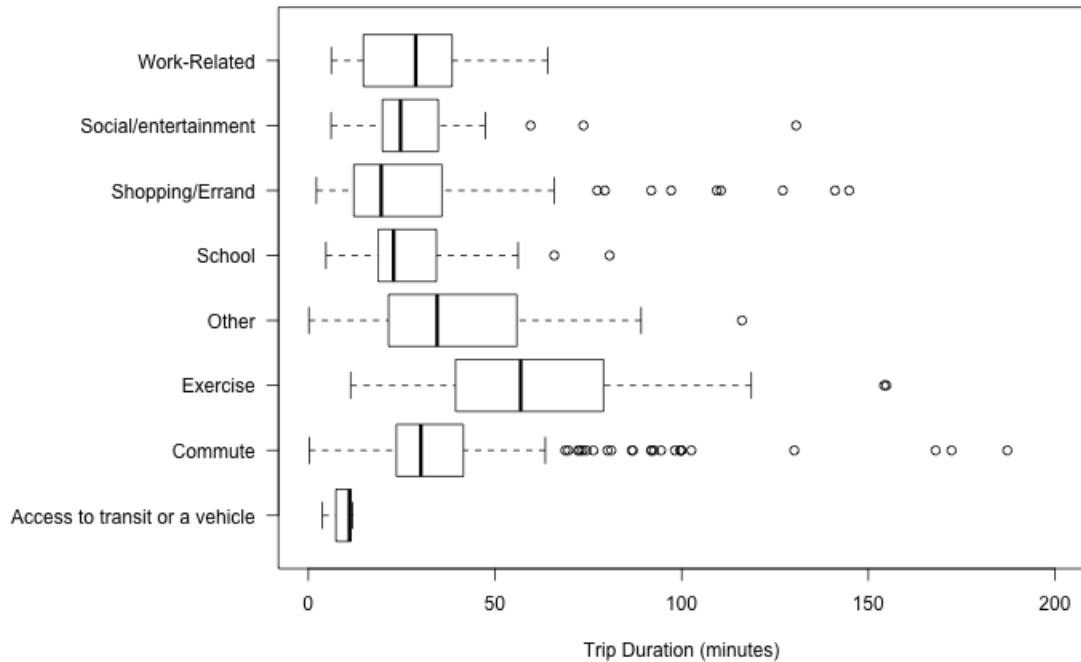
Figure 35: Boxplots of Trip Duration distribution by Trip Purpose

Trip distance was then calculated, with distributions varying among different trip purposes. These distributions are presented in Figure 36. The overall median trip distance was 4.7 miles. Exercise trips had the highest median trip distance with 11.1 miles, while transit access trips had the lowest median trip distance with 1.8 miles.
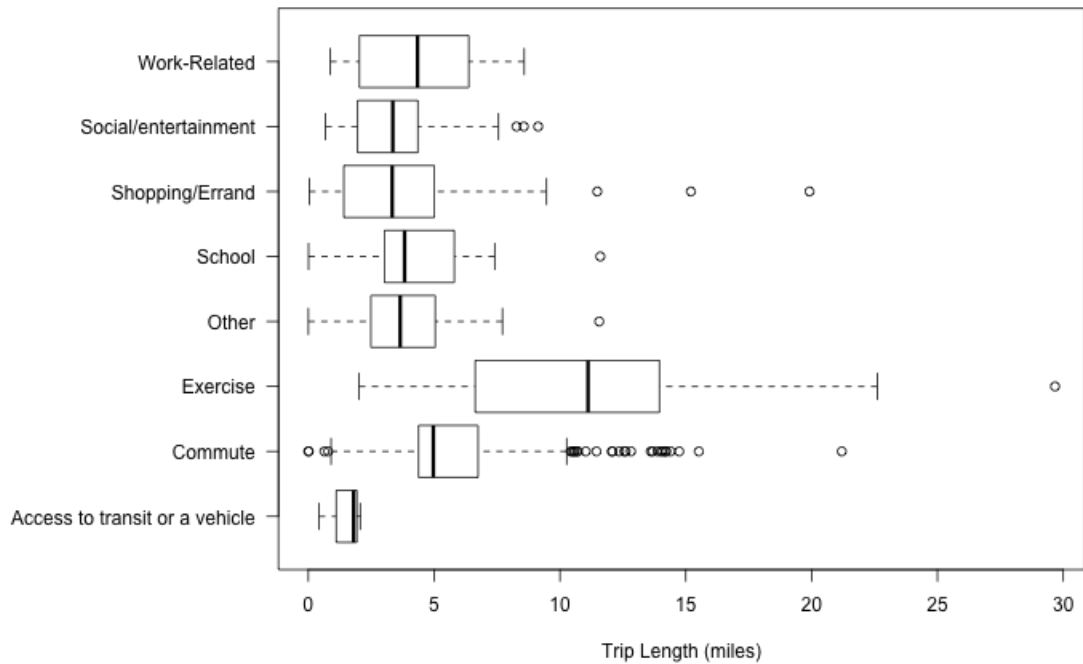
Figure 36: Boxplots of Trip Distance distribution by Trip Purpose

Average trip speed was then calculated by dividing trip distance by trip duration. The distribution of the average speed by trip purpose is presented in Figure 37. The overall median average speed was 9.7 miles per hour. The highest median average speed was for commute trips at 10.6 miles per hour, and the lowest median average speed was for "other" trips at 7.3 miles per hour.
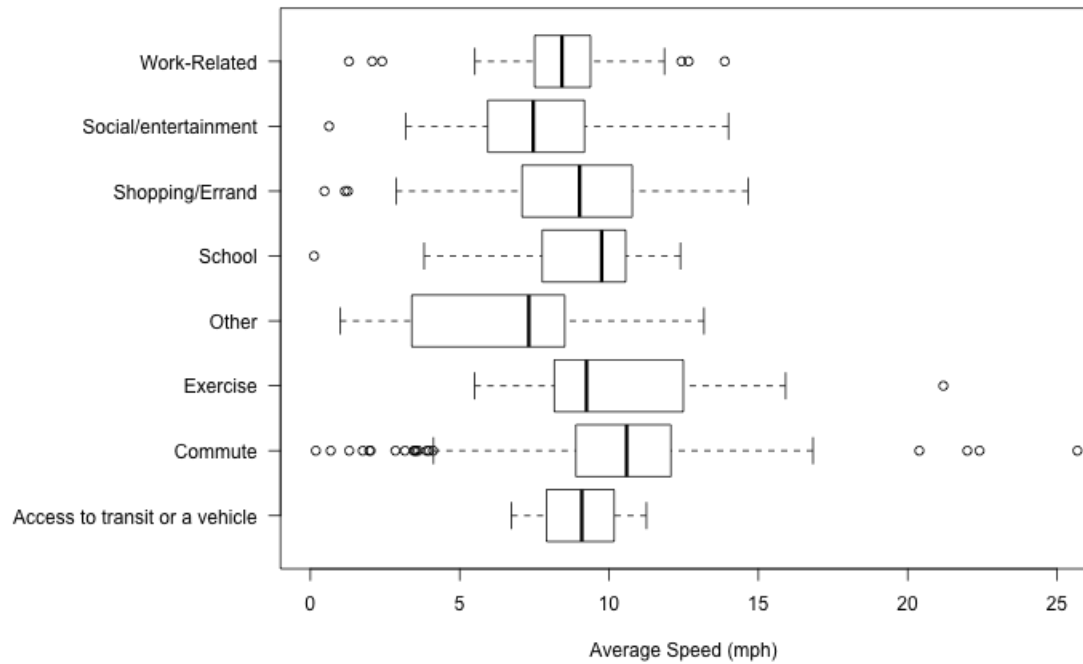
Figure 37: Boxplots of Average Speed by Trip Purpose

Start time distributions for the different trip purposes are presented in Figure 38, with a

higher concentration of points indicating more trips starting around that time. The

commute trip distribution was bi-modal, with many trips starting around 8 AM or 5 PM.

The other trip purpose start times were more evenly distributed throughout the day.
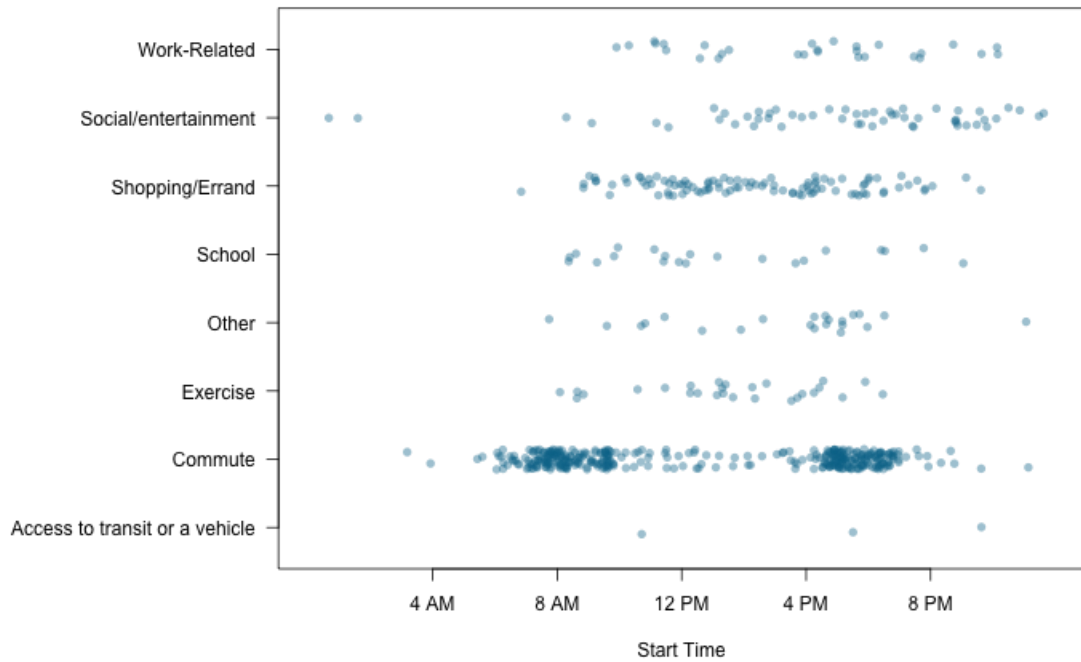
Figure 38: Trip Start Time Distribution by Trip Purpose

Finish time distributions for the different trip purposed are presented in Figure 39 with a higher concentration of points indicating more trips finishing around that time. The commute trip distribution was multi-modal, with many trips finishing around 8 AM or 5 PM. The other trip purpose finish times were more evenly distributed throughout the day.
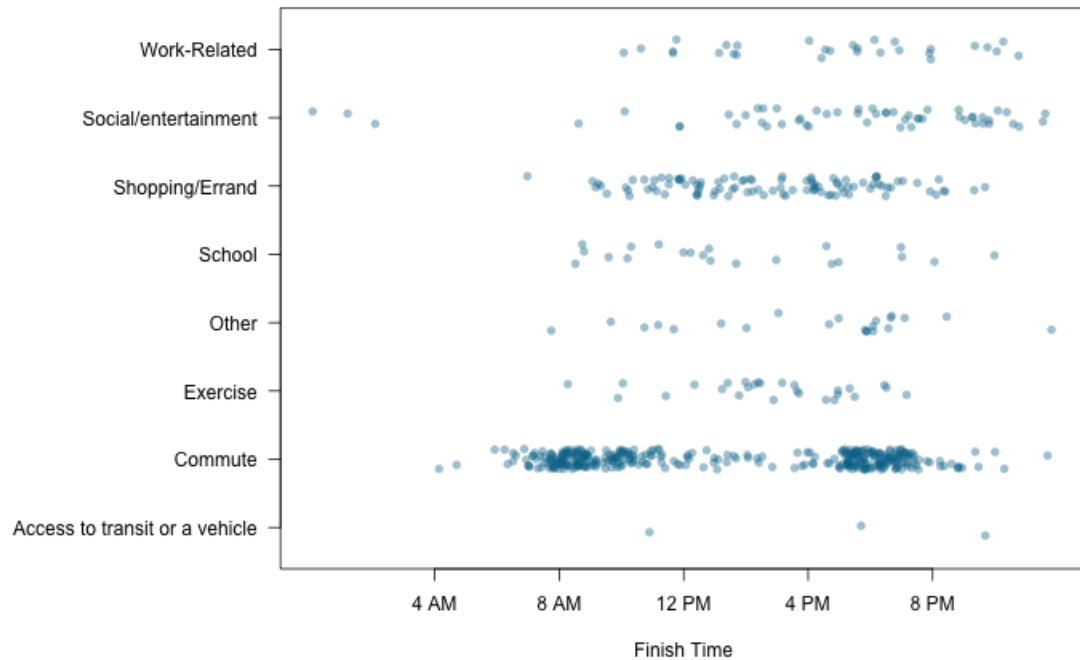
Figure 39: Trip Finish Time Distribution by Trip Purpose

4.4 CRASH EVENT AND SAFETY ISSUE REPORT DATA

Report data was collected in three distinct pieces: the report time and location, the report question responses, and an optionally included report photo. Reports were also divided into two categories: safety/infrastructure issues and crash events. The rate of report recording and the cumulative number of reports recorded over the study period are graphed in Figure 40 and Figure 41. Like users and trips, the rate of report recording initially surged with the release of the app but leveled off to a nearly constant rate shortly after. The average report recording rate was 1.7 reports per day. There were 215 reports considered in this study, with 153 of them being safety/infrastructure issue reports and 62 of them being crash even reports.

Figure 40: Report recording rate over study period

Figure 41: Cumulative report count over study period

4.4.1 Crash Event Reports

All questions asked for the crash event reports were mandatory upon reporting a crash event.

4.4.1.1 *Severity*

When documenting a crash event report, users were asked to indicate the severity of the crash event. Users could indicate, on a 1-5 scale, that the crash event was a "near-miss" or that it resulted "major injuries". The distribution of severity among crash event reports is illustrated in Figure 42. The majority of crash event reports (62%) were indicated to be near misses.
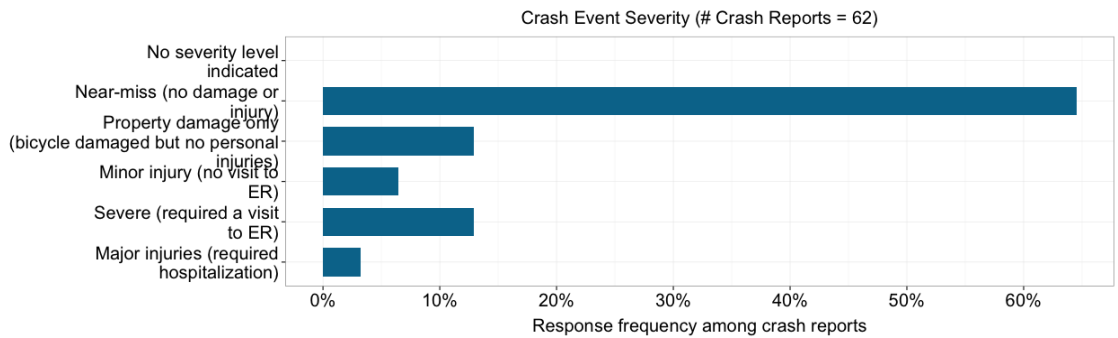
Figure 42: Severity Distribution among Crash Reports

4.4.1.2 *Conflict With*

When documenting a crash report, users were asked to indicate what vehicle or object conflicted with them during the crash event from among ten options (with an "other" option to indicate a custom response). The conflicting vehicle/object distribution among crash event reports is illustrated in Figure 43.
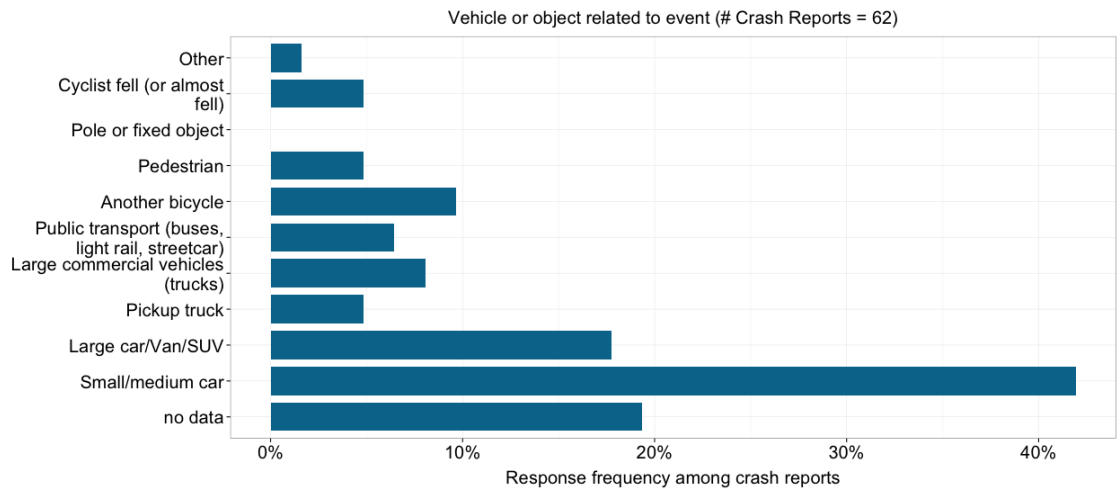
Figure 43: Conflict Type Distribution among Crash Reports

4.4.1.3 *Actions*

Upon reporting a crash event, users were asked to indicate the actions of themselves or another road user that they felt contributed to the crash event. Users could select from among ten options (including an "other" option with custom text input). The crash action distribution among crash events is illustrated in Figure 44.
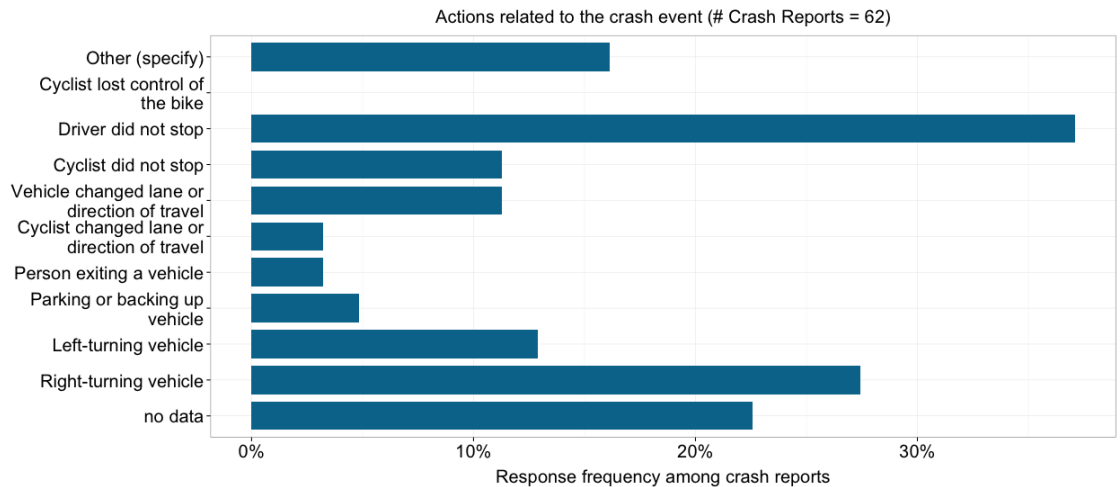
Figure 44: Crash Actions among Crash Reports

4.4.1.4 *Reasons*

Upon reporting a crash event, users could indicate what reasons they felt contributed to the crash event from among eleven options (including an "other" option with custom text input). The crash reason distribution among crash events is illustrated in Figure 45.
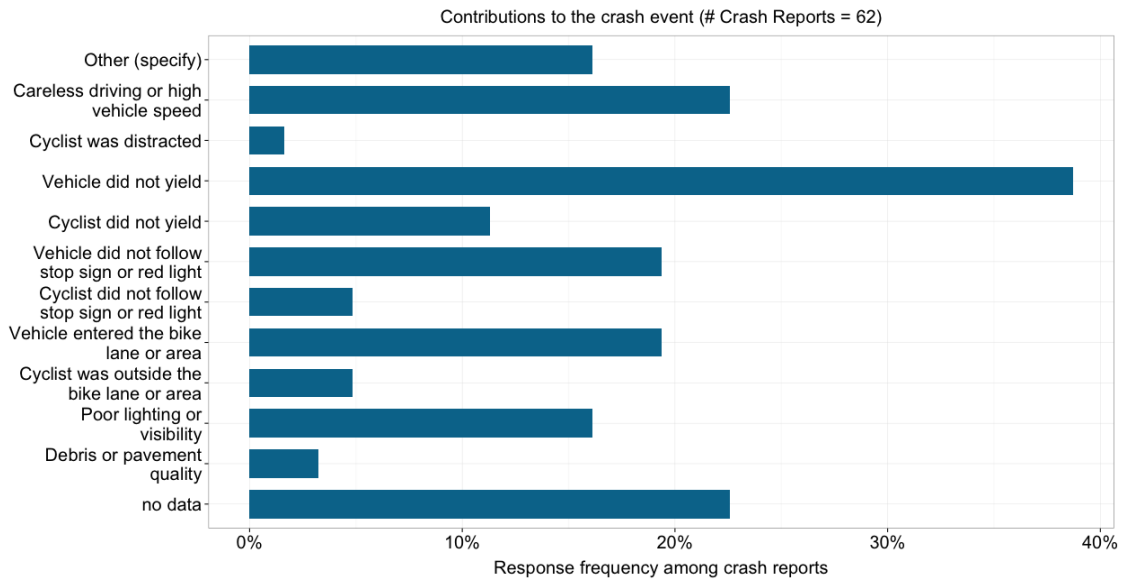
Figure 45: Crash Reasons among Crash Reports

4.4.1.5 *Geography*

Geographic analysis was used to separate crash event reports by state. 95% of the crash event reports were located within Oregon. This was expected as the application was marketed to Oregon users. The geographic distribution of crash reports among states is illustrated in Figure 46.

Crash Reports by State (# Reports = 62)
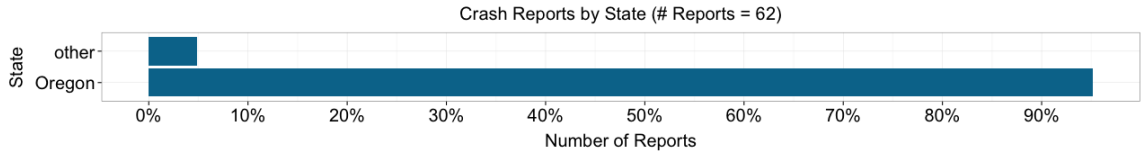
Figure 46: State Distribution among Crash Reports

Geographic analysis was used to separate crash event reports by county. 92% of the

reports were located in Multnomah County. The geographic distribution of crash reports

among counties is illustrated in Figure 47.

Crash Reports by County (# Reports = 62)
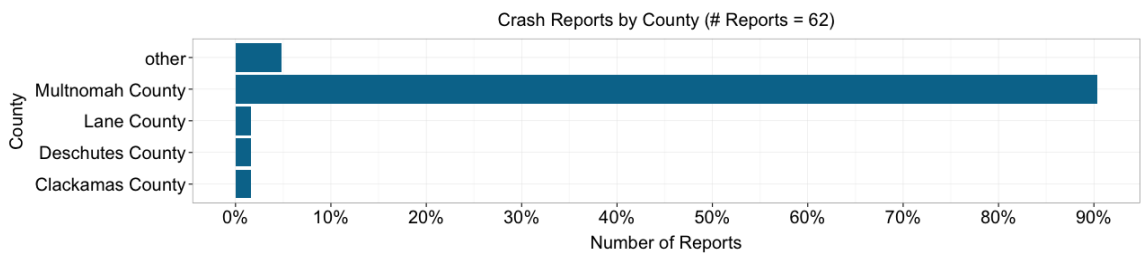
Figure 47: County Distribution among Crash Reports

Geographic analysis was used to separate crash event reports by city. 92% of the reports

were located within the city of Portland. The geographic distribution of crash reports

among cities is illustrated in Figure 48.

Crash Reports by City (# Reports = 62)
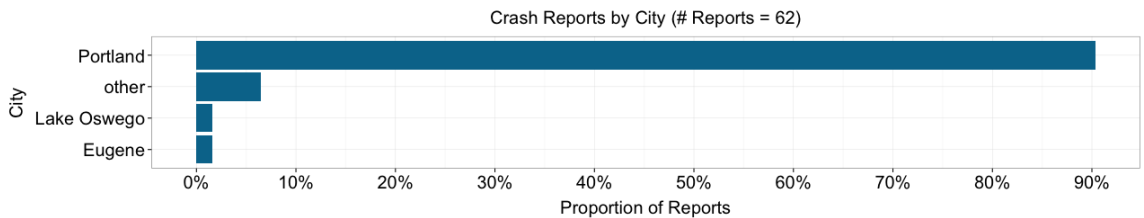
Figure 48: City Distribution among Crash Reports

The high concentration of reports in Multnomah County and the City of Portland

indicates a bias in the sample towards the Portland area.

4.4.2 Safety Issue Reports

Both of the questions asked when reporting a safety issue were mandatory to answer.

5.3.1.1 *Issue Type*

When users reported a safety/infrastructure issue, they were asked to identify what type of issue they were reporting. Users could select one or more of fourteen options and provide custom text input for the "other" option. The issue type distribution among safety issue reports is illustrated in Figure 49. 33% of the reports had "High traffic volume" indicated, and 32% of the reports had "other" indicated.
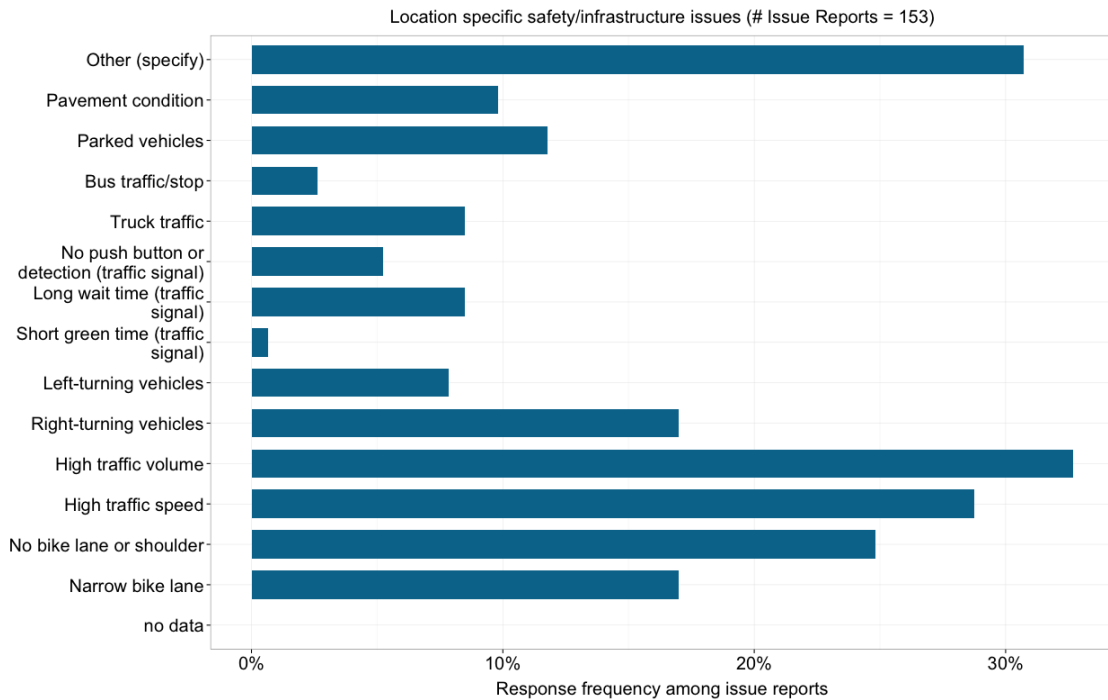
Figure 49: Issue Type Distribution among Safety Issue Reports

5.3.1.2 *Urgency*

When reporting a safety issue, users were asked to indicate the urgency of that issue. Users could select one option on a 1-5 scale of urgency, with 1 being the least urgent and 5 being the most urgent. The urgency distribution among safety issue reports is illustrated in Figure 50.The majority of issues were concentrated in the 3 and 4 categories (53%).
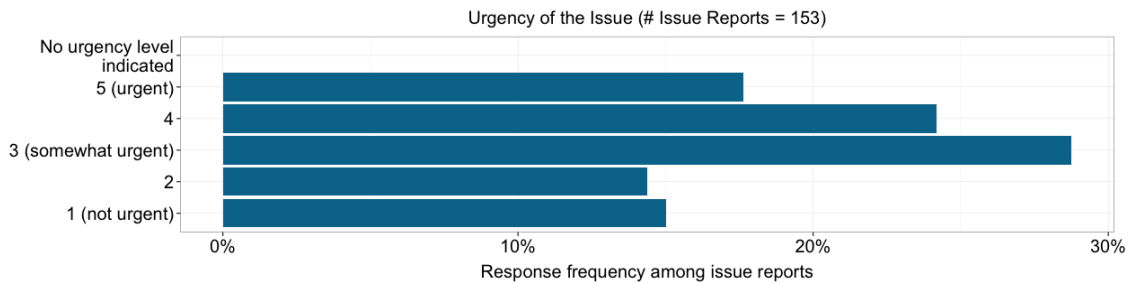
Figure 50: Urgency Distribution among Safety Issue Reports

Geographic analysis was used to separate issue reports by state. The geographic distribution of safety issue reports among states is illustrated in Figure 51. ORcycle was developed to collect data in Oregon, but nearly 10% of the issue reports came from other states.

5.3.1.3 *Geography*
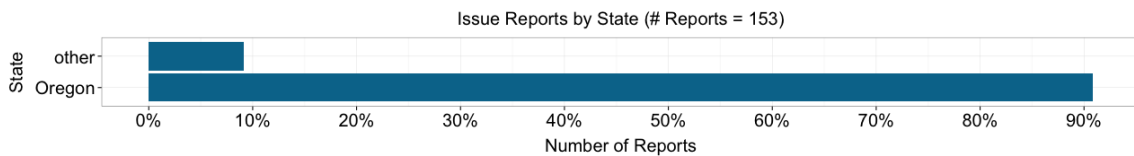
Figure 51: State Distribution among Safety Issue Reports

Geographic analysis was also used to separate issue reports by county. The geographic distribution of safety issue reports among counties is illustrated in Figure 52. The majority (67%) of reports were made in Multnomah County (where Portland is located).

Nearly 10% of the reports were made in "other" counties, which included reports outside of Oregon.
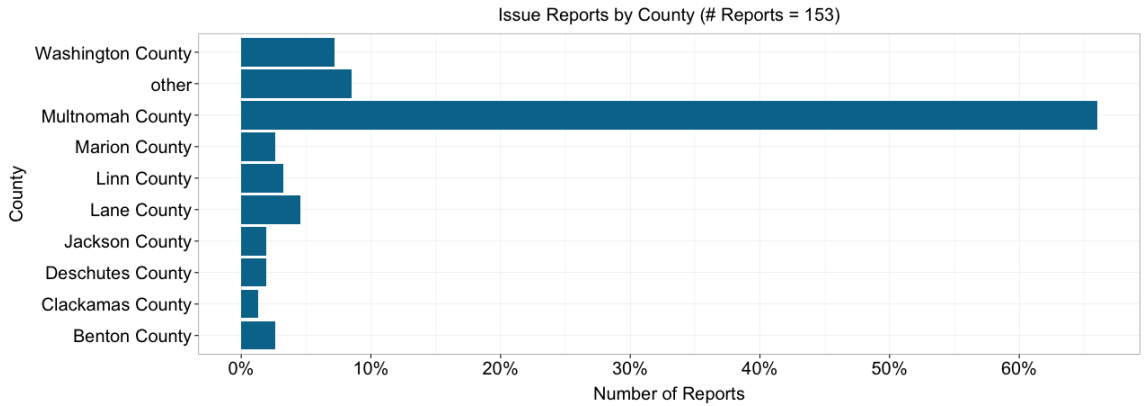
Figure 52: County Distribution among Safety Issue Reports

Geographic analysis was also used to separate issue reports by city. The geographic distribution of safety issue reports among states is illustrated in Figure 53. The majority of reports (67%) were made within the city of Portland. 18% of the reports were made in the "other" category, which was comprised of both reports outside of Oregon and reports in unincorporated areas of Oregon.
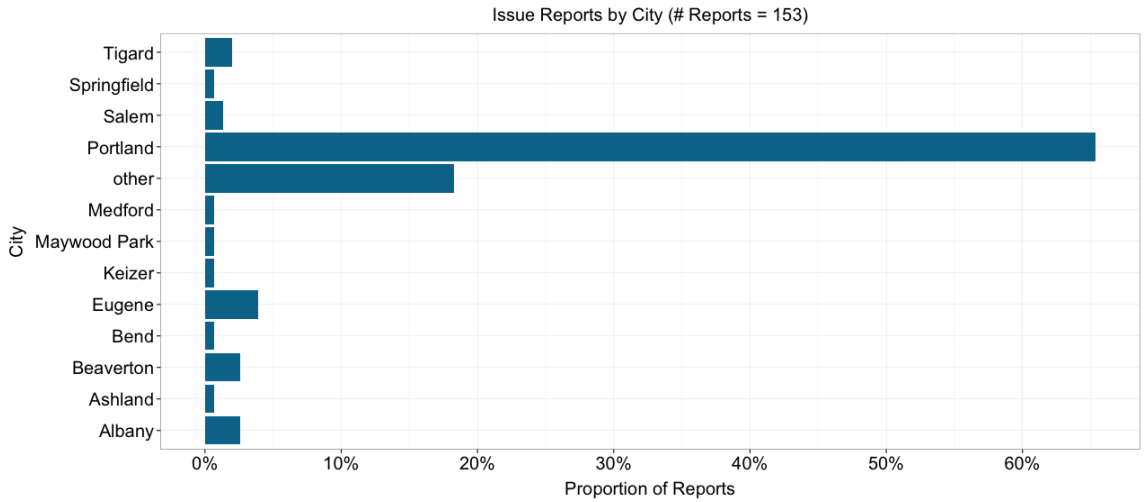
Figure 53: City Distribution among Safety Issue Reports

The high concentration of reports in Multnomah County and the City of Portland

indicates a bias in the sample towards the Portland area.

5. TRIP COMFORT ANALYSIS

5.1 MODEL FORMULATION

After finishing a trip, users could[14] indicate how comfortable they felt the route was (see section 4.3.3 for more information on this question) on a one (very bad) through five (very good) scale. Where the user did not answer this question, the response was imputed as a function of their other survey responses (see section 3.2.2 for more info on multiple imputation procedure). The response to the "route comfort" question serves as the dependent variable for the following analyses and is meant to be roughly analogous to Bicycle Level of Service and Level of Traffic Stress (see section 2.1) measures. Using a cumulative logistic regression approach as is used in several level of service models (Jensen 2007; Ali, Cristei, and Flannery 2012; Foster et al. 2015), route comfort is modeled based on several groups of variables outlined below.

Logistic regression models are used to model categorical dependent variables, whereas standard linear regression models are used to model continuous dependent variables. Cumulative logistic regression models (also known as ordinal logistic regression models) are used to model categorical dependent variables of an ordered nature. Route comfort is clearly ordered, with "very bad" representing a condition worse than "bad", "bad"

---

[14] In ORcycle version 2.2.0, released on March 7th, 2015, this question was made mandatory to answer upon finishing a trip.

representing a condition worse than "average", and so on. The cumulative logistic regression model enables interpretations of the direction and magnitude of the change in route comfort with respect to some independent variable. Using a standard multinomial logistic regression model would result in a loss of information pertaining to the ordered nature of the dependent variable, and using a linear regression model may provide an incorrect model given the different distribution of residuals[15] associated with continuous data. The cumulative logistic regression models presented herein were constructed using the R package "ordinal"[16], which offers many tools for statistically modeling ordinal outcome variables.

A geo-matching script was used that filtered out trips that did not meet minimum criteria (trip length and speed) and then matched the remaining trips to a network model. Only trips that passed initial filtering and could be geo-matched to Metro's bicycle/street network were considered within the following models; resulting in a final sample of 616 trips from 148 unique users. More details on the geo-matching process are given in section 3.2.2.

Given that this thesis investigates a unique dataset with limited related research, variable groups are first explored separately before investigating the use of a pooled regression

---

[15] Categorical data typically have distributions better modeled by logistic functions (as opposed to linear functions). However, with more ordered categories added, the distribution begins to approach normal and is thus more eligible for a linear model.

[16] http://www.cran.r-project.org/package=ordinal/

model. The single variable group model results are presented in section **Error!**

eference source not found.. A pooled regression model is considered in section 0.

Finally, interpretations of the model exploration are given in section 5.4.

5.2   SINGLE VARIABLE GROUP MODELS

In all of the models tested, the route comfort rating was the dependent variable. The

following independent variable groups were first explored separately: (1) trip attributes

(length, duration, and average speed), (2) trip temporal characteristics, (3) user-reported

trip characteristics (e.g. trip purpose), (4) user attitudes and socio-demographics (5)

bicycle facility and street typology, (6) topography, (7) traffic volume, (8) posted traffic

speed, and (9) weather characteristics.

For each variable group, the variable definitions are first presented. Within the presented

variable definition tables, the variable type (e.g. continuous or categorical) is designated

and the range of possible values of the variable within the model is described. Measures

of central tendency are then presented (i.e. median for continuous and mode for

categorical). Stacked bar plots are then referenced that illustrate the relative proportions

of route comfort ratings among different levels of each independent variable (see

Appendix 9.2). For continuous variables, a single variable cumulative logit model was

tested for each variable to assess the relationship of that variable to route comfort (in

terms of significance, magnitude, and direction) alone. For categorical variables, the Chi-

Square test of independence was used to test for a statistically significant relationship

between the variable of interest and route comfort. In this test, the null hypothesis would

be that the variable of interest has no relationship with route comfort; which would be rejected in the case of the Chi-Square statistic being statistically significant.

Cumulative logistic regression models incorporating all potential variables within each variable group were then explored; these exploratory models are presented in Appendix A.3. After running backwards stepwise regression (using the Akaike Information Criterion) and removing insignificant variables, final cumulative logistic regression model specifications are then presented. For all of the models, a description of the significant coefficients is presented along with regression tables.

Odds ratios are also presented along with the model coefficients. Odds ratios are more readily interpretable than the model coefficients, as they describe the odds of an increase in the independent variable corresponding to an increase in route comfort. For example, an odds ratio of 2 is interpreted as "an increase of one unit in the independent variable results in twice the odds that a route will be rated more comfortably than a given comfort rating". Conversely, an odds ratio of 0.5 is interpreted as "an increase of one unit in the independent variable results in half the odds that a route will be rated more comfortably than a given comfort rating, or that there are twice the odds that a route will be rated less comfortably than a given comfort rating". Forest plots[17] illustrating the direction,

---

[17] Forest plots are plots of the odds ratios of logistic regression model coefficients and their corresponding 95% confidence intervals.

magnitude (odds ratio), and variability (confidence interval) of the independent variables'

relationships with route comfort are also presented.

5.2.1 **Trip Attributes**

Three trip attributes were calculated: trip length (miles), trip duration (minutes), and

average trip speed (miles per hour). These attributes had to be calculated from the geo-

matched results, as calculations from the raw GPS coordinates yielded erroneous results

because of the tendency for users to leave ORcycle recording GPS points longer than

they were actually traveling for. These trip attribute variables were tested for significant

relationships with route comfort. The corresponding variable definitions are outlined in

Table 44.

Table 44: Trip attribute variable definitions

Variable Description	Variable Type	Possible Values of Variable (range for Continuous variables)	Median (for Continuous) Mode (for Categorical)	Route Comfort Distribution Plot (in Appendix 0)	z-statistic in single variable cumulative logit	Statistical Significance
Trip length	Continuous	Min: 0.30 miles Max: 29.67 miles	4.75 miles	Figure 65	-2.389	p<0.05
Trip duration	Continuous	Min: 2.51 minutes Max: 166 minutes	29.38 minutes	Figure 66	0.087	Not significant
Average speed	Continuous	Min: 0.63 mph Max: 16.83 mph	9.70 mph	Figure 67	-2.282	p<0.05

After several exploratory models (see section 1.3), a final regression model was selected for the trip attribute variables using a backwards stepwise regression approach. Statistically insignificant variables were dropped from the final model. The final model specification is presented in Table 45, with the odds ratios illustrated in Figure 54. The final model only included average speed, resulting in a statistically significant relationship (β=-0.09, OR=0.91, p<0.05) that decreased route comfort as average speed increased.

Two interpretations could be given to this result: (1) cyclists that travel faster are less comfortable or (2) cyclists that are less comfortable travel faster.

Table 45: Final trip attribute model specification (cumulative logistic)

	Dependent variable:
	Route Comfort Rating
Average Speed	-0.092** (0.037)
Observations	616
Log Likelihood	-778.519
Note:	*p<0.1; **p<0.05; ***p<0.01
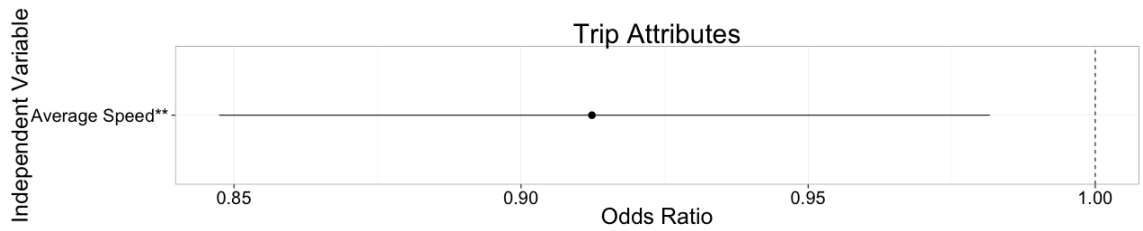
Figure 54: Forest plot of odds ratios of coefficients for final trip attributes model (whiskers correspond to 95% CI)

5.2.2 Temporal Characteristics

Temporal characteristics were tested to explore how route comfort varied over time of day and day of the week. The time a trip started was used to categorize these temporal variables into two groups representing the difference between weekday and weekend travel as well as the difference between peak time travel and off-peak time travel. The corresponding variable definitions are outlined in Table 46.

Table 46: Temporal characteristics variable definitions

Variable Description	Variable Type	Possible Values of Variable (range for Continuous variables)	Median (for Continuous) Mode (for Categorical)	Route Comfort Distribution Plot (in Appendix 0)	Chi-Square, DF	Statistical Significance
Trip day-of-week category	Categorical	• Weekday • Weekend	Weekday	Figure 68	10.57, 8	p<0.05
Trip start time category	Categorical	• Off-Peak Night (6:30 PM to 7:00 AM) • Peak AM (7:00 AM-9:00 AM) • Off-Peak Day (9:00 AM to 4:30 PM) • Peak PM (4:30 PM to 6:30 PM)	Off-Peak Day	Figure 69	8.65, 18	Not significant

After several exploratory models (see section 1.A.3.2), a final regression model was

selected for the temporal characteristics variables using a backwards stepwise regression

approach. Statistically insignificant variables were dropped from the final model. The

final model specification is presented in Table 47, with the odds ratios illustrated in 54.

Trips taking place on a weekday were rated less comfortable than those taken on

weekends (β=-0.43, OR=0.65, p<0.05). The other variables tested were found to be

insignificant and were not included in the final model specification.

Table 47: Final temporal characteristic model specification (cumulative logistic)

	Dependent variable:
	Route Comfort Rating
Trip took place on weekday (reference = weekend)	-0.428** (0.212)
Observations	616
Log Likelihood	-779.480
Note:	*p<0.1; **p<0.05; ***p<0.01
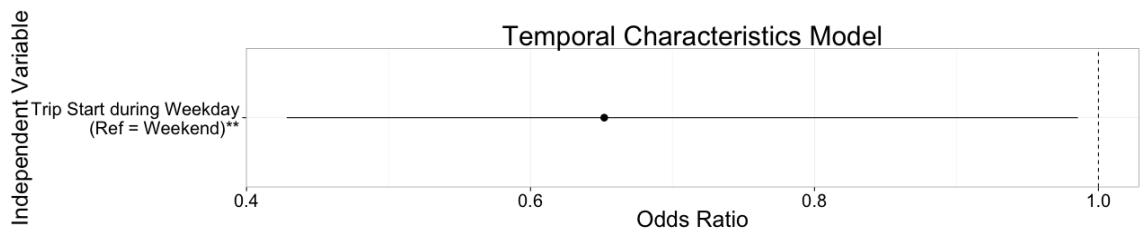
Figure 55: Forest plot of odds ratios of coefficients for final temporal characteristics model (whiskers correspond to 95% CI)

5.2.3 **Self-Reported Trip Characteristics**

The other trip survey question responses (besides route comfort, which was used as the dependent variable) were explored for significant effects on route comfort. The corresponding variable definitions are outlined in Table 48 and Table 49.

Table 48: Trip question response variable definitions

Variable Description	Variable Type	Possible Values of Variable (range for Continuous variables)	Median (for Continuous) Mode (for Categorical)	Route Comfort Distribution Plot (in Appendix 0)	Chi-Square, DF	Statistical Significance
Trip purpose	Categorical	• Commute • School • Work related • Exercise • Social or Entertainment • Shopping or Errands • Transport Access • Other	Commute	Figure 70	58.96, 38	p<0.01
Indication of often the user takes this particular route	Ordinal categorical	1) First time ever 2) Once per year or less 3) Several times per year 4) Several times per month 5) Several times per week	Several times per week	Figure 71	23.91, 23	p<0.10

Table 49: Trip question response variable definitions (continued)

Variable Description	Variable Type	Possible Values of Variable (range for Continuous variables)	Median (for Continuous) Mode (for Categorical)	Route Comfort Distribution Plot (in Appendix 0)	z-statistic in variable group cumulative logit model	Statistical Significance
User indicated they chose this route because it was direct or fast.	Binary categorical	True/False	True	Figure 72	-8.49	p<0.001
User indicated they chose this route because it has good bicycle facilities.	Binary categorical	True/False	False	Figure 72	4.08	p<0.001
User indicated they chose this route because it is enjoyable or has nice scenery.	Binary categorical	True/False	False	Figure 72	1.97	p<0.05
User indicated they chose this route because it is good for a workout.	Binary categorical	True/False	False	Figure 72	-0.54	Not significant
User indicated they chose this route because it has low traffic or low vehicle speeds.	Binary categorical	True/False	False	Figure 72	3.51	p<0.001
User indicated they chose this route because it has few busy intersections.	Binary categorical	True/False	False	Figure 72	2.76	p<0.01
User indicated they chose this route because it has few and/or easy hills.	Binary categorical	True/False	False	Figure 72	0.64	Not significant
User indicated they chose this route because it has other	Binary categorical	True/False	False	Figure 72	1.64	Not significant

riders/people.						
User indicated they chose this route because it is good for families/kids.	Binary categorical	True/False	False	Figure 72	3.71	p<0.001
User indicated they chose this route because they do not know another route.	Binary categorical	True/False	False	Figure 72	-3.24	p<0.01
User indicated they chose this route because they found it online or using their phone.	Binary categorical	True/False	False	Figure 72	1.28	Not significant
User indicated they chose this route because of some other reason.	Binary categorical	True/False	False	Figure 72	-0.82	Not significant
User indicated that on this route they were not concerned with traffic stressors.	Binary categorical	True/False	False	Figure 73	4.23	p<0.001
User indicated that on this route they experienced discomfort as a result of auto traffic.	Binary categorical	True/False	False	Figure 73	-2.81	p<0.01
User indicated that on this route they experienced discomfort as a result of large commercial vehicles/trucks.	Binary categorical	True/False	False	Figure 73	-8.11	p<0.001
User indicated that on this route they experienced discomfort as a result of public transport.	Binary categorical	True/False	False	Figure 73	-1.57	Not significant
User indicated that on this route	Binary categorical	True/False	False	Figure 73	0.92	Not significant

| they experienced discomfort as a result of parked vehicles.						
User indicated that on this route they experienced discomfort as a result of other cyclists.	Binary categorical	True/False	False	Figure 73	2.17	$p < 0.05$
User indicated that on this route they experienced discomfort as a result of pedestrians.	Binary categorical	True/False	False	Figure 73	1.62	Not significant
User indicated that on this route they experienced discomfort as a result of auto traffic.	Binary categorical	True/False	False	Figure 73	0.68	Not significant

After several exploratory models (see section 1.A.3.3), a final regression model was selected for the user question response variables using a backwards stepwise regression approach. Statistically insignificant variables were dropped from the final model. The final model included trip purpose, route frequency, route choice preferences, and route stressors variables. The final model specification is presented in Table 50, with the odds ratios illustrated in Figure 56.

Trip purpose was included in the final model, with the reference case being commute trips. Shopping/errand trips were significantly (β=0.856, OR=2.53, $p<0.01$) more comfortable than commute trips. Work-related trips were significantly (β=-0.909, OR=0.40, $p<0.05$) less comfortable than commute trips. School trips were significantly (β=-1.01, OR=0.36, $p<0.05$) less comfortable than commute trips.

Route frequency was included in the final model, with increased route frequency corresponding to increased route comfort (β=0.598, OR=1.82, $p<0.05$). This result indicates that cyclists riding routes they ride often are more comfortable on those routes.

Several route choice preferences (i.e. self-reported reasons why a user traveled on their particular route) were included in the model. Users who indicated they chose their route because it was direct or fast rated their trips as less comfortable (β=-2.767, OR=0.06, $p<0.01$). Users who indicated choosing their route because it was "good for a workout" also rated their trips as less comfortable (β=-1.66, OR=0.19, $p<0.05$). Users who indicated choosing their routes because it was "good for families/kids" rated their trips as more comfortable (β=2.03, OR=7.61, $p<0.01$). Users who indicated choosing their routes

124

because they did not know another route rated their trips as less comfortable (β=-1.71, OR=0.18, p<0.01).

Route stressors (i.e. self-reported characteristics of the chosen route that made users uncomfortable or stressed) were included in the final model. Users who indicated they were not concerned about traffic stressors on their route rated their routes as more comfortable (β=2.13, OR=8.41, p<0.01). Users who indicated they were concerned about large commercial vehicles on their route rated that route as less comfortable (β=-1.87, OR=0.15, p<0.01). Users who indicated they were concerned about public transit vehicles on their route rated that route as less comfortable (β=-1.93, OR=0.14 p<0.01). Users who indicated they were concerned about parked vehicles along their route also rated their route as less comfortable (β=-1.05, OR=0.35, p<0.01). Finally, users who indicated they were concerned about pedestrians along their route also rated their route as less comfortable (β=-0.62, OR=0.54, p<0.01).

Table 50: Final trip question response model specification (cumulative logistic)

	Route Comfort Rating	
	Coefficient (Standard Error)	Number of Observations
Route Stressors-'Not Concerned'	2.131*** (0.458)	26
Route Stressors-'Large commercial vehicles or trucks'	-1.865*** (0.365)	36
Route Stressors-'Public transport'	-1.932*** (0.487)	21
Route Stressors-'Parked vehicles + being doored'	-1.046*** (0.221)	109
Route Stressors-'Pedestrians'	-0.615*** (0.231)	109
Route Preferences-'It is direct + fast'	-2.767*** (0.277)	77
Route Preferences-'It has good bicycle facilities'	-0.507* (0.295)	51
Route Preferences-'It is good for a workout'	-1.661*** (0.352)	34
Route Preferences-'It is good for families + kids'	2.025*** (0.544)	21
Route Preferences-'I do not know another route'	-1.710*** (0.384)	33
Reference = Commute		
Trip Purpose-'School'	-1.012** (0.432)	21
Trip Purpose-'Shopping + Errands'	0.856*** (0.282)	97
Trip Purpose-'Work-related'	-0.909** (0.376)	29
Route Frequency (Ordinal)	0.598** (0.273)	616
Observations	616	
Log Likelihood	-636.933	
Note:	*p<0.1; **p<0.05; ***p<0.01	
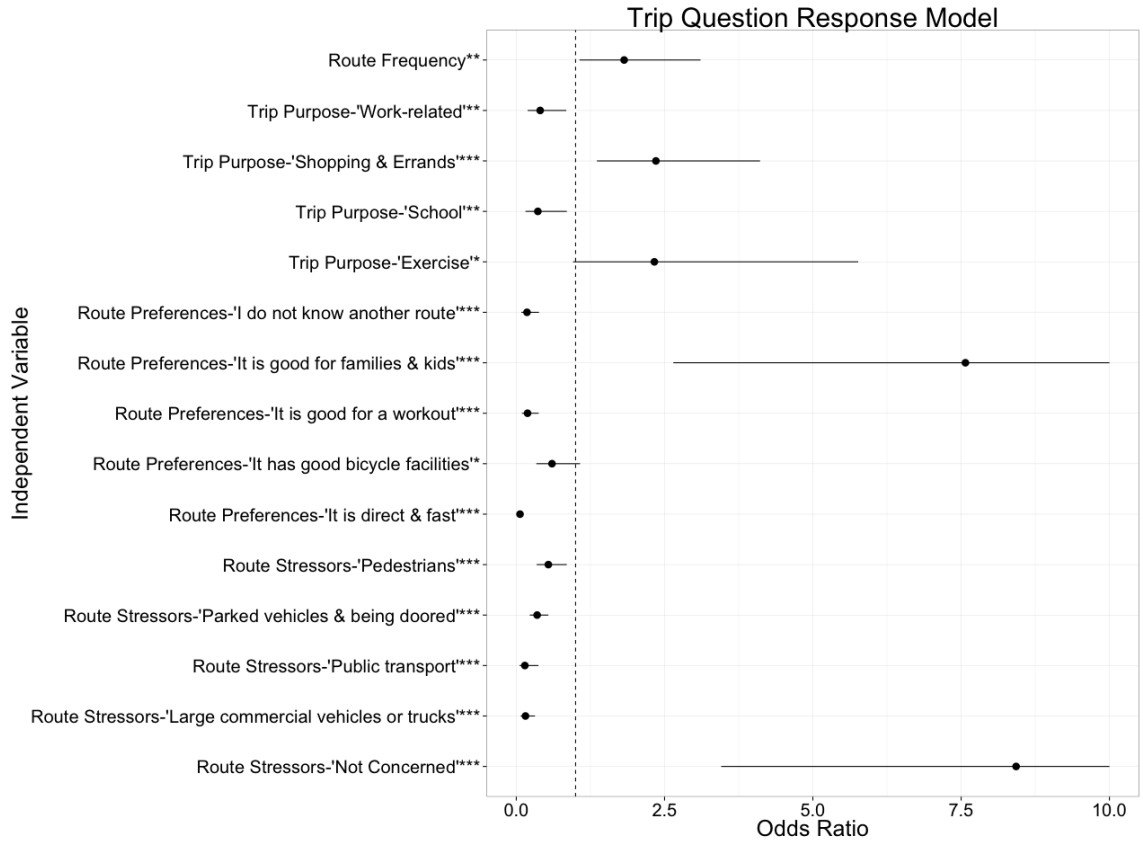
Figure 56: Forest plot of odds ratios of coefficients for final trip question response model (whiskers correspond to 95% CI)

5.2.4 User Attitudes and Socio-Demographic Characteristics

User survey question responses were explored for significant effects on route comfort.

User question variable definitions are outlined in Table 51.

Table 51: User question response variable definitions

Variable Description	Variable Type	Possible Values of Variable (range for Continuous variables)	Median (for Continous) Mode (for Categorical)	Route Comfort Distribution Plot (in Appendix 0)	Chi-Square, DF	Statistical Significance
Age category	Ordered categorical	• Less than 18 • 18-24 • 25-34 • 35-44 • 45-54 • 55-64 • 65+	35-44	Figure 74	122.95, 33	p<0.01
Gender category	Categorical	• Female • Male • Other	Male	Figure 75	81.57, 13	p<0.01
Ethnicity category	Categorical	• African American • Asian American • Hispanic • Native American • White American • Other	White American	Figure 76	95.66, 23	p<0.01
Occupation category	Categorical	• Employed • Student • Retired • Homemaker • Other	Employed	Figure 77	94.37,23	p<0.01
Income category	Ordered categorical	1) Less than $14,999 2) $15,000 to $24,999 3) $25,000 to $34,999 4) $35,000 to $49,999 5) $50,000 to $74,999 6) $75,000 to $99,999 7) $100,000 to $149,999 8) $150,000 or more	$100,000 to $149,999	Figure 79	100, 38	p<0.01

Number of household workers	Ordered categorical	1) 0 Workers 2) 1 Worker 3) 2 Workers 4) 3 Workers or more	2 workers	Figure 78	29.04, 18	$p < 0.05$
Number of household vehicles	Ordered categorical	1) 0 vehicles 2) 1 vehicle 3) 2 vehicles 4) 3 vehicles or more	2 Vehicles	Figure 80	71.21, 18	$p < 0.01$
Number of bicycles owned by user	Ordered categorical	1) 0 bicycles 2) 1 bicycle 3) 2 bicycles 4) 3 bicycles 5) 4 or more bicycles	4 or more bicycles	Figure 81	59.11, 18	$p < 0.01$
Cycling frequency category	Ordered categorical	1) A few times per year 2) A few times per month 3) A few times per week 4) Nearly every day	Nearly every day	Figure 85	10.96, 13	Not significant
Cycling weather category	Categorical	• In any kind of weather • When it does not rain • Usually warm and dry weather • Only with warm and dry weather	In any kind of weather	Figure 82	13.64, 18	$p < 0.1$
Rider ability category	Ordered categorical	1) Very Low 2) Low 3) Average 4) High 5) Very High	Very High	Figure 84	59.3, 23	$p < 0.01$
Rider type category	Categorical	• For nearly all my trips • To and from work • For recreation and/or exercise • For shopping, errands, or	To and from work	Figure 83	61.71, 28	$p < 0.01$

		- visiting friends - Mainly to and from work, but occasionally for exercise, shopping, etc. - Other				

After several exploratory models (see section 1.A.3.4), a final regression model was selected for the temporal characteristics variables using a backwards stepwise regression approach. Statistically insignificant variables were dropped from the final model; many of the user variables were dropped. The final model included gender, ethnicity, and occupation as independent variables. The final model specification is presented in Table 52, with the odds ratios illustrated in Figure 57.

Users identifying as white ethnicities were less comfortable than non-white (β=-1.40, OR=0.61, p<0.01), and users who were employed were more comfortable than non-employed users (β=1.26, OR=3.51, p<0.01).

Table 52: Final user question response model specification (cumulative logistic)

	Route Comfort Rating	
	Coefficient (Standard Error)	Number of Observations
Ethnicity: White	-1.399*** (0.25)	523
Occupation: Employed	1.256*** (0.238)	545
Observations	616	
Log Likelihood	-746.631	

Note: *p<0.1; **p<0.05; ***p<0.01
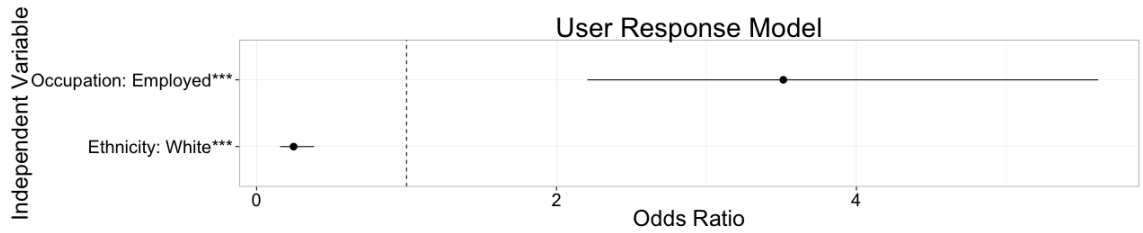
Figure 57: Forest plot of odds ratios of coefficients for final user question response model (whiskers correspond to 95% CI)

5.2.5 Bicycle Facility and Street Type

After geo-matching (geo-matching process explained in section 3.2.2) trips to the Portland metropolitan area street/bike network, we were able to discern the bicycle facility and/or street type of the links used on each trip. Metro's street categorization included over twenty categories, and their bicycle facility categorization included over ten categories. These categories were aggregated to test for contrasts of interest to the researchers. The typology used was adapted from the link typology used in the bicycle GPS study conducted by Dill and Gliebe (Dill and Gliebe 2008). The relevant vocabulary used in the typology is defined below:

- Primary arterials are multi-lane roads that carry high traffic volumes at high speeds

- Minor arterials are multi-lane roads that carry moderate traffic volumes at moderate speeds

- Residential streets are two or one way streets primarily used for residential access

- "Other" streets are those streets that did not fit into the other three categories

132

- Bicycle lanes are dedicated road space for cyclists delineated only by striping, with no lateral separation between bicyclists and motor vehicle traffic

- Buffered bicycle lanes are similar to bicycle lanes, but they have extra buffer space allocated on the roadway using striping to laterally separate bicyclists from motor vehicle traffic.

- Bicycle boulevards are low-traffic streets that have been designated for bicycle travel. They feature bicycle route signage and pavement markings, traffic calming features such as traffic circles or speed humps, and motor vehicle traffic diversion at major intersections.

- Cycletracks (AKA protected bicycle lanes) have lateral separation enforced using some physical buffer, such as planters, plastic posts, parked cars, raised concrete barriers, or other treatments.

- Separated paths are linear transportation facilities where motor vehicle traffic is prohibited but bicycle traffic is allowed and/or encouraged.

- "No Bicycle Facility" means that there was no bicycle facility on the particular link matching any of the above bicycle facility descriptions. In these cases, bicycles share the traffic lane with motor vehicle traffic and no special consideration is given to bicyclists.

Bicycle facility variables are outlined in Table 53.

Table 53: Bicycle Facility variable definitions

Variable Description	Variable Type	Possible Values of Variable (range for Continuous variables)	Median (for Continous) Mode (for Categorical)	Route Comfort Distribution Plot (in Appendix 0)	z-statistic in variable group cumulative logit model	Statistical Significance
Number of miles of trip ridden on primary arterials with no bike lanes	Continuous	Min: 0.00 miles Max: 6.54 miles	0.03 miles	Figure 86 and Figure 87	-4.05	p<0.001
Number of miles of trip ridden on minor arterials with no bike lanes	Continuous	Min: 0.00 miles Max: 2.97 miles	0.00 miles	Figure 86 and Figure 87	0.74	Not significant
Number of miles of trip ridden on residential streets with no bike lanes	Continuous	Min: 0.00 miles Max: 12.08 miles	1.38 miles	Figure 86 and Figure 87	-0.07	Not significant
Number of miles of trip ridden on other types of streets with no bike lanes	Continuous	Min: 0.00 miles Max: 5.90 miles	0.01 miles	Figure 86 and Figure 87	-2.26	p<0.05
Number of miles of trip ridden on primary arterials with bike lanes	Continuous	Min: 0.00 miles Max: 9.48 miles	0.20 miles	Figure 86 and Figure 87	-1.93	p<0.1
Number of miles of trip ridden on minor arterials with bike lanes	Continuous	Min: 0.00 miles Max: 4.12 miles	0.09 miles	Figure 86 and Figure 87	-3.59	p<0.001
Number of miles of trip ridden on residential streets with	Continuous	Min: 0.00 miles Max: 6.28 miles	0.34 miles	Figure 86 and Figure 87	-0.58	Not significant

| bike lanes						
Number of miles of trip ridden on other streets with bike lanes	Continuous	Min: 0.00 miles Max: 1.90 miles	0.00 miles	Figure 86 and Figure 87	0.63	Not significant
Number of miles of trip ridden on cycletracks or buffered bicycle lanes	Continuous	Min: 0.00 miles Max: 0.80 miles	0.00 miles	Figure 86 and Figure 87	0.34	Not significant
Number of miles of trip ridden on bicycle boulevards	Continuous	Min: 0.00 miles Max: 5.03 miles	0.34 miles	Figure 86 and Figure 87	1.34	Not significant
Number of miles of trip ridden on separated paths	Continuous	Min: 0.00 miles Max: 12.54 miles	0.27 miles	Figure 86 and Figure 87	4.89	p<0.001

135

After several exploratory models (see section 1.A.3.5), a final regression model was selected for the bicycle facility/street type variables using a backwards stepwise regression approach. Statistically insignificant variables were dropped from the final model. The final model specification is presented in Table 54, with the odds ratios illustrated in Figure 58.

Trip miles on links typed as "no bike facility, primary arterial" detracted from route comfort (β=-0.49, OR=0.61, p<0.01). Trip miles on links typed as "no bike facility, other" also detracted from route comfort (β=-0.41, OR=0.66, p<0.05). Trip miles on links typed as "bike lane, primary arterial" also detracted from route comfort (β=-0.167, OR=0.85, p<0.05). Trip miles on links typed as "bike lane, minor arterial" also detracted from route comfort (β=-0.54, OR=0.58, p<0.01). Finally, trip miles on links typed as separated paths increased route comfort (β=0.33, OR=1.41, p<0.01).

Table 54: Final bike facility/street type model specification (cumulative logistic)

	Route Comfort Rating	
	Coefficient (Standard Error)	Mileage within Model Sample
Trip Miles on 'No Bike Facility, Primary Arterial'	-0.487*** (0.114)	168
Trip Miles on 'No Bike Facility, Other'	-0.410** (0.173)	83
Trip Miles on 'Bike Lane, Primary Arterial'	-0.167** (0.077)	335
Trip Miles on 'Bike Lane, Minor Arterial'	-0.544*** (0.153)	212
Trip Miles on 'Separated Path'	0.331*** (0.068)	417
		Total = 3,200
Observations	616	
Log Likelihood	-743.114	
Note:	*p<0.1; **p<0.05; ***p<0.01	
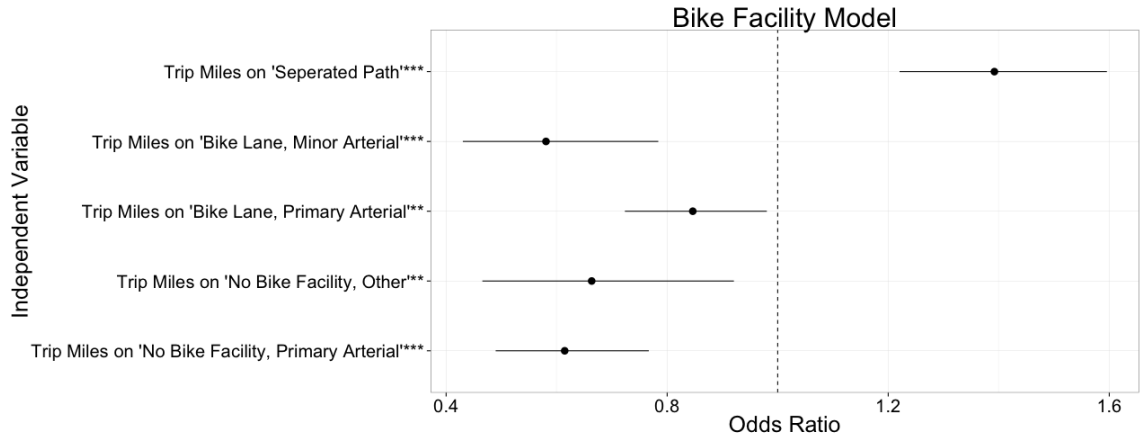
Figure 58: Forest plot of odds ratios of coefficients for final bike facility/street type model (whiskers correspond to 95% CI)

These results align with literature indicating that more separation from traffic has a

positive effect on cyclist comfort (see literature review).

5.2.6 **Topography**

The average slopes of network segments were calculated using a digital elevation model

provided by Metro's Regional Land Information System (RLIS). The slope variable

definitions are outlined in Table 55.

Table 55: Topography variable definitions

Variable Description	Variable Type	Possible Values of Variable (range for Continuous variables)	Median (for Continuous) Mode (for Categorical)	Route Comfort Distribution Plot (in Appendix 0)	z-statistic in variable group cumulative logit model	Statistical Significance
Number of miles of trip ridden on slope category "less than -2%	Continuous	Min: 0.00 miles Max: 4.49 miles	0.67 miles	Figure 88 and Figure 89	-2.76	p<0.01
Number of miles of trip ridden on slope category "between -2% and +2%"	Continuous	Min: 0.10 miles Max: 24.59 miles	3.21 miles	Figure 88 and Figure 89	1.20	Not significant
Number of miles of trip ridden on slope category "between +2% and +4%"	Continuous	Min: 0.00 miles Max: 2.59 miles	0.42 miles	Figure 88 and Figure 89	1.74	p<0.1
Number of miles of trip ridden on slope category "between +4% and +6%"	Continuous	Min: 0.00 miles Max: 2.46 miles	0.10 miles	Figure 88 and Figure 89	0.20	Not significant
Number of miles of trip ridden on slope category "greater than +6%"	Continuous	Min: 0.00 miles Max: 1.27 miles	0.05 miles	Figure 88 and Figure 89	-4.37	p<0.001

After several exploratory models (see section 1.A.3.6), a final regression model was

selected for the segment grade variables using a backwards stepwise regression approach.

Statistically insignificant variables were dropped from the final model. The final model

specification is presented in Table 56, with the odds ratios illustrated in Figure 59.

Trip miles on network segments with grades less than -2% detracted from route comfort

(β=-0.38, OR=0.68, p<0.01). Trip miles on network segments with grades between +2%

and +4% increased route comfort (β=0.53, OR=1.69, p<0.05). Trip miles on network

segments with grades greater than +6% detracted from route comfort (β=-2.76, OR=0.06,

p<0.01).

Table 56: Final segment grade model specification (cumulative logistic)

	Route Comfort Rating	
	Coefficient (Standard Error)	Mileage Within Model Sample
Trip miles on grades <-2%	-0.384[**] (0.150)	490
Trip miles on grades +2% to 4%	0.525[**] (0.243)	302
Trip miles on grades >+6%	-2.760[***] (0.607)	78
		Total=3,200
Observations	616	
Log Likelihood	-755.227	
Note:	[*]p<0.1; [**]p<0.05; [***]p<0.01	
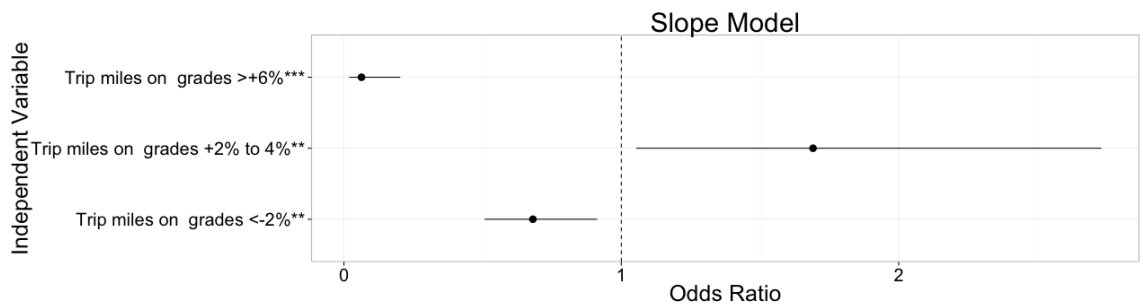
Figure 59: Forest plot of odds ratios of coefficients for final segment grade model (whiskers correspond to 95% CI)

5.2.7 **Daily Traffic Volume**

Interpolated daily motor vehicle traffic volumes estimated for links in the bike/street network were provided courtesy of Joseph Broach; the network volume model was developed as part of Dill and Gliebe's bicycle GPS study (Dill and Gliebe 2008) based on City of Portland traffic volumes. Where traffic volumes for links were unavailable, a linear regression based on a link's functional classification (for links where volume was estimated) was used to predict the missing traffic volumes. Route comfort was modeled over the estimated daily vehicle volumes to discern if estimated daily traffic volumes were significantly related. Traffic volume variable definitions are presented in Table 57.

Table 57: Traffic Volume variable definitions

Variable Description	Variable Type	Possible Values of Variable (range for Continuous variables)	Median (for Continuous) Mode (for Categorical)	Route Comfort Distribution Plot (in Appendix 0)	z-statistic in variable group cumulative logit model	Statistical Significance
Number of miles of trip ridden on traffic category "less than 5,000 vehicles per day"	Continuous	Min: 0.00 miles Max: 18.17 miles	2.59 miles	Figure 90 and Figure 91	4.24	p<0.001
Number of miles of trip ridden on traffic category "between 5,000 and 10,000 vehicles per day"	Continuous	Min: 0.00 miles Max: 6.58 miles	0.45 miles	Figure 90 and Figure 91	-1.18	Not significant
Number of miles of trip ridden on traffic category "between 10,000 and 20,000 vehicles per day"	Continuous	Min: 0.00 miles Max: 5.96 miles	0.60 miles	Figure 90 and Figure 91	-1.70	p<0.1
Number of miles of trip ridden on traffic category "between 20,000 and 30,000 vehicles per day"	Continuous	Min: 0.00 miles Max: 9.51 miles	0.13 miles	Figure 90 and Figure 91	-2.53	p<0.05

| Number of miles of trip ridden on traffic category "greater than 30,000 vehicles per day"	Continuous	Min: 0.00 miles Max: 4.04 miles	0.00 miles	Figure 90 and Figure 91	-4.35	p<0.001

After several exploratory models (see section 1.A.3.7), a final regression model was

selected for the traffic volume variables using a backwards stepwise regression approach.

Statistically insignificant variables were dropped from the final model. The final model

specification is presented in Table 58, with the odds ratios illustrated in Figure 59.

Trip miles on network segments with traffic volumes less than 5,000 vehicles per day

increased route comfort (β=0.32, OR=1.38, p<0.01). Trip miles on network segments

with traffic volumes greater than 30,000 vehicles per day decreased route comfort (β=-

0.78, OR=0.46, p<0.01). Controlling for trip length resulted in a significant relationship

in this model, with route comfort decreasing as trip length increased (β=-0.16, OR=0.85,

p<0.01)

Table 58: Final traffic volume model specification (cumulative logistic)

	Route Comfort Rating	
	Coefficient (Standard Error)	Mileage Within Model Sample
Trip miles on links with 'Less than 5k veh/day'	0.323*** (0.055)	1,740
Trip miles on links with 'Greater than 30k veh/day'	-0.777*** (0.203)	100
Trip length (miles)	-0.164*** (0.037)	3,200
		Total = 3,200
Observations	613	
Log Likelihood	-743.328	
Note:	*p<0.1; **p<0.05; ***p<0.01	
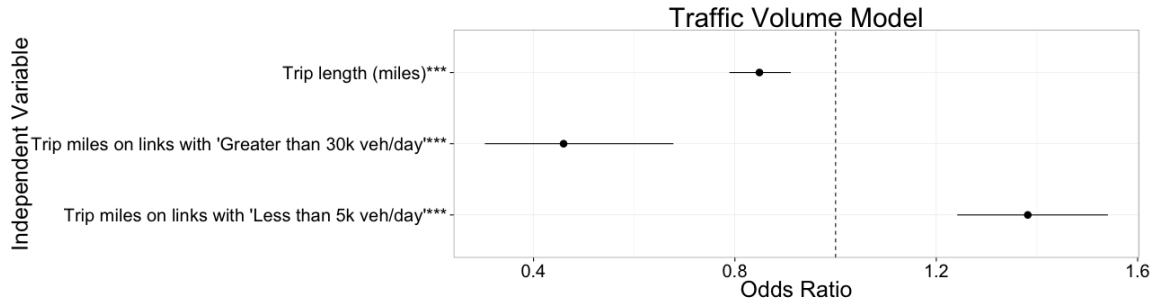
Figure 60: Forest plot of odds ratios of coefficients for final traffic volume model (whiskers correspond to 95% CI)

5.2.8 Traffic Speed

Posted speed limits were provided by Metro for links in the bike/street network. Where posted speeds were missing, a linear regression based on link functional class was used to estimate the missing speeds. Route comfort was modeled over the posted traffic speed to determine if route comfort was related to traffic speed. Traffic speed variable definitions are presented in Table 78.

Table 59: Speed Volume variable definitions

| Variable Description	Variable Type	Possible Values of Variable (range for Continuous variables)	Median (for Continuous) Mode (for Categorical)	Route Comfort Distribution Plot (in Appendix 0)	z-statistic in variable group cumulative logit model	Statistical Significance
Number of miles of trip ridden on speed category "less than or equal to 20 mph"	Continuous	Min: 0.00 miles Max: 12.84 miles	0.40 miles	Figure 92 and Figure 93	4.10	p<0.001
Number of miles of trip ridden on speed category "between 20 and 35 mph"	Continuous	Min: 0.00 miles Max: 13.34miles	3.05 miles	Figure 92 and Figure 93	-0.64	Not significant
Number of miles of trip ridden on speed category "greater than or equal to 35 mph"	Continuous	Min: 0.00 miles Max: 13.07 miles	0.51 miles	Figure 92 and Figure 93	-5.64	p<0.001

After several exploratory models (see section 1.A.3.8), a final regression model was selected for the posted traffic speed variables using a backwards stepwise regression approach. Statistically insignificant variables were dropped from the final model. The final model specification is presented in Table 60, with the odds ratios illustrated in Figure 61.

Trip miles on network segments with posted speeds less than or equal to 20 mph increased route comfort (β=0.25, OR=1.28, p<0.01). Trip miles on network segments with posted speeds greater than 35 mph decreased route comfort (β=-0.36, OR=0.70, p<0.01).

Table 60: Final posted traffic speed model specification (cumulative logistic)

	Route Comfort Rating	
	Coefficient (Standard Error)	Mileage Within Model Sample
Trip miles on links with posted speed <= 20 MPH	0.246*** (0.060)	514
Trip miles on links with posted speed >35 MPH	-0.361*** (0.061)	577
		Total = 3,200
Observations	616	
Log Likelihood	-755.374	
Note:	*p<0.1; **p<0.05; ***p<0.01	

146
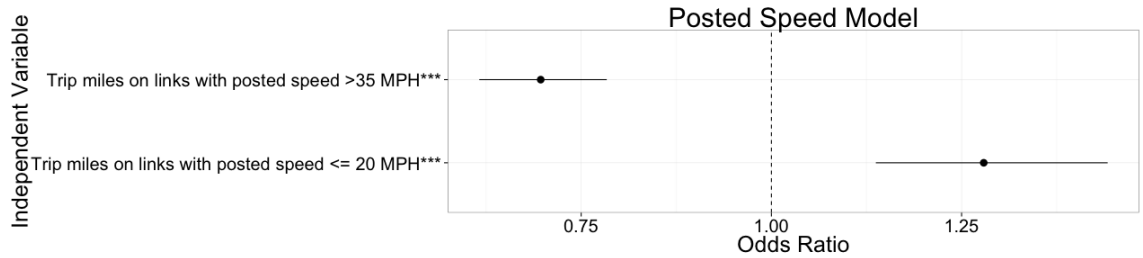
Figure 61: Forest plot of odds ratios of coefficients for final posted traffic speed model (whiskers correspond to 95% CI)

5.2.9 Weather

Weather data was pulled from airport weather stations in the Portland metropolitan region. The weather data for each trip was pulled from the airport weather station nearest to the start location of the trip, and the readings attached to each trip pertained to the temporally nearest weather record. Temperature, wind speed, wind gust speed, precipitation volume, and weather conditions (category provided by weather stations) were tested for significant effects on route comfort. The weather variable definitions are outlined in Table 61.

Table 61: Weather variable definitions

Variable Description	Variable Type	Possible Values of Variable (range for Continuous variables)	Median (for Continuous) Mode (for Categorical)	Route Comfort Distribution Plot (in Appendix 0)	z-statistic in single variable cumulative logit	Statistical Significance
Temperature (degrees Farenheit)	Continuous	Min: 26.10 °F Max: 66.90 °F	45.00 °F	Figure 94	1.04	Not significant
Wind speed (miles per hour)	Continuous	Min: 0.00 mph Max: 38.00 mph	8.10 mph	Figure 96	-1.45	Not significant
Wind gust speed (miles per hour)	Continuous	Min: 0.00 mph Max: 48.30 mph	0.00 mph	Figure 97	-0.87	Not significant
Hourly precipitation	Continuous	Min: 0.00 inches Max: 0.10 inches	0.00 inches	Figure 95	0.73	Not significant
Variable Description	Variable Type	Possible Values of Variable (range for Continuous variables)	Median (for Continuous) Mode (for Categorical)	Route Comfort Distribution Plot (in Appendix 0)	Chi-Square, DF	Statistical Significance
Weather conditions category	Categorical	• Clear • Fog • Light Clouds • Heavy Clouds • Light Rain • Heavy Rain	Heavy Clouds	Figure 98	28.79, 28	Not significant

Several exploratory models (see section 1.A.3.9) were constructed to test for significant relationships between the weather variables and route comfort. Using a stepwise regression approach, no final models were selected because most relationships were insignificant. It is likely this is due to the low variation in weather conditions during the study period.

5.3 POOLED MODEL

There are numerous combinations of the above explanatory variables that could be combined to form pooled regression models (i.e. models containing more than one variable group), but the following model (specified in Table 62) was selected using a backwards stepwise regression approach. A forest plot illustrating the odds ratios corresponding to each model coefficient is presented in Figure 62.

Table 62: Pooled regression model specification (cumulative logistic)

	Route Comfort Rating	
	Coefficient (Standard Error)	Number of Observations or Trip Mileage
Route Frequency (Ordinal)	0.677** (0.270)	613
Trip Purpose: Exercise	1.723*** (0.512)	27
Trip Purpose: Shopping/Errands	1.043*** (0.271)	97
Occupation: Employed	1.496*** (0.277)	545
Ethnicity: White	-1.369*** (0.265)	523
Household Vehicles (Ordinal)	2.116*** (0.500)	613
Route Preferences: It is direct or fast	-2.178*** (0.272)	77
Route preferences: It is good for a workout	-1.345*** (0.350)	34
Route Preferences: It has few busy intersections	0.739** (0.293)	62
Route Preferences: It is good for families/kids	2.347*** (0.568)	21
Route Stressors: Not concerned	2.111*** (0.484)	26
Route Stressors: Large commercial vehicles/trucks	-1.751*** (0.366)	36
Route Stressors: Parked vehicles/being doored	-0.637*** (0.215)	109
Trip miles on grades >+6%	-2.351*** (0.548)	78 miles
Trip miles on links with <5k veh/day	0.250*** (0.056)	1,740 miles
Trip miles on links with posted speed >35 mph	-0.242*** (0.066)	577 miles
Trip miles on links with posted speed 20-35 mph	-0.148*** (0.054)	2,056 miles
Observations	613	
Log Likelihood	-587.259	
Note:	*p<0.1; **p<0.05; ***p<0.01	
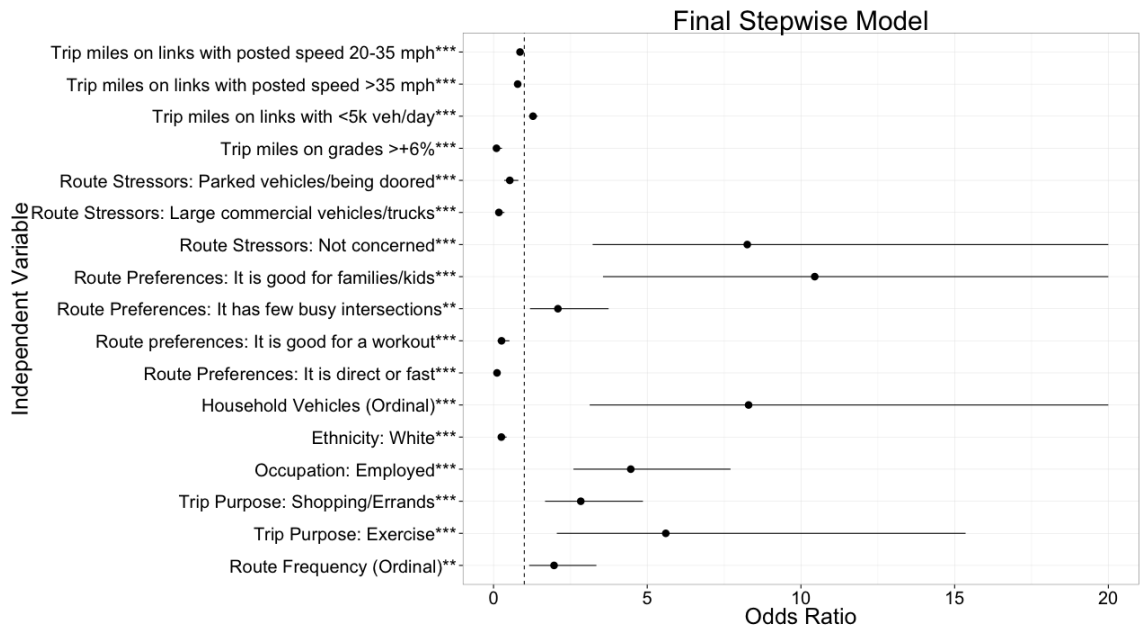
Figure 62: Forest plot of odds ratios of coefficients for final pooled model (whiskers correspond to 95% CI)

Due to non-intuitive sign changes in the bicycle facility and street type variables when incorporated into the pooled model, these variables were not incorporated in the final pooled model. These sign changes were likely due to other variables controlled for within the pooled model, though narrowing down which particular variables caused the sign change would require many more model runs. These models are exploratory in nature and developing a robust predictive model was not the goal of this thesis.

In order to gauge the relative contribution of each predictor in the pooled regression model, the difference in Log Likelihood when each variable was removed one at a time *ceteris parabus* (all other variables remaining in the model) was calculated. A percentage contribution to the model was then calculated by dividing the individual log-likelihood

change by the difference between the model log-likelihood and the null log-likelihood (i.e. the variance in the dependent variable accounted for by the predictors in the model). These results are presented in Table 63.

Conducting these post-hoc analyses yields several insights. It illustrates the relative importance of the predictive variables included in the model. For example, the variable describing a route choice preference of "It is direct or fast" accounted for 16% of the accounted for model variance in route comfort, whereas the number of trip miles on links with low traffic volumes (<5k veh/day) accounted for 5.9% of the variance in route comfort. If one were to develop a predictive model of route comfort, the variables with the highest contribution (top portion of the table) would be the most important to include.

Another important note is the number of variables that were dropped from the final model even though they proved significant in the smaller models. This is likely due to correlation between predictors as well as predictors that were confounding the effect of some other variable. With more model testing, confounding relationships could likely be discovered. To deal with the correlation between predictors, indices could be developed to incorporate multiple correlated factors, such as using a variable that incorporated both traffic volume and posted speed (since these variables are correlated).

Table 63: Pooled regression model results summary

Independent Variable	Statistics from full pooled model			Log Likelioods from *Ceteris parabus* removed variable model	
	Coefficient	Odds Ratio	z-statistic	Log-Likelihood Difference	Contribution
Route Preferences: It is direct or fast	-2.37	0.09	-8.41	57.18	26.5%
Occupation: Employed	1.60	4.94	5.69	39.72	18.4%
Ethnicity: White	-1.40	0.25	-5.25	38.97	18.1%
Route Stressors: Large commercial vehicles/trucks	-1.42	0.24	-3.54	36.49	16.9%
Household Vehicles (Ordinal)	2.10	8.13	4.15	35.56	16.5%
Trip miles on links with <5k veh/day	0.25	1.29	4.43	34.70	16.1%
Trip miles on grades >+6%	-2.24	0.11	-4.03	34.67	16.1%
Route Stressors: Not concerned	2.18	8.85	4.49	34.46	16.0%
Route Preferences: It is good for families/kids	2.35	10.46	4.10	34.43	16.0%
Trip Purpose: Shopping/Errands	1.08	2.95	3.92	32.45	15.0%
Route preferences: It is good for a workout	-1.44	0.24	-4.05	32.37	15.0%
Trip miles on links with posted speed >35 mph	-0.21	0.81	-3.19	31.96	14.8%
Trip Purpose: Exercise	1.71	5.55	3.35	30.50	14.1%
Route Frequency (Ordinal)	0.71	2.04	2.62	30.24	14.0%
Route Stressors: Parked vehicles/being doored	-0.63	0.53	-2.92	29.22	13.5%
Trip miles on links with posted speed 20-35 mph	-0.17	0.85	-3.03	28.49	13.2%
Route Preferences: It has few busy intersections	0.66	1.94	2.23	28.28	13.1%

5.4   MODEL INTERPRETATION

The models reviewed above became increasingly complex as additional predictive variables were added, and the practical application of these models beyond the analysis of this data is limited due to the very unique nature of the dataset. However, the models do add further evidence to the growing body of research surrounding factors contributing to or detracting from cyclist comfort. In general, the signs of the predictive variables reviewed herein align with those presented in the relevant literature. In Table 64, the primary trends from each model group are summarized, and it is noted whether these trends held in the pooled model. Several variables at the bottom of the table were only significant within the pooled model. Those trends that held within the single variable group models and the pooled model can be considered most robust.

Table 64: Model Interpretation Summary

Variable Group	Significant Variables		Influence on Route Comfort	Relationship held significant in same direction in pooled model
Trip Statistics	Average speed		Decrease	No
Temporal Characteristics	Weekday Trip		Decrease	No
Self-reported characteristics	Route Stressors	Not concerned	Increase	Yes
		Large commercial vehicles	Decrease	Yes
		Public transportation	Decrease	No
		Parked vehicles + being doored	Decrease	Yes
		Pedestrians	Decrease	No
	Route Preferences	It is direct + fast	Decrease	Yes
		It has good bicycle facilities	Decrease	No
		It is good for a workout	Decrease	Yes
		It is good for families + kids	Increase	Yes
		I do not know another route	Decrease	No
	Route Frequency (Ordinal)		Increase	Yes
	Trip Purpose (Reference = Commute)	School	Decrease	No
		Shopping + Errands	Increase	Yes
		Work-related	Decrease	No
Self-reported user characteristics	Ethnicity: White (Dummy)		Decrease	Yes
	Occupation: Employed (Dummy)		Increase	Yes
Bicycle facility and street type (Trip miles on links with…)	No bike facility, primary arterial		Decrease	No
	No bike facility, other		Decrease	No
	Bike lane, primary arterial		Decrease	No
	Bike lane, minor arterial		Decrease	No
	Separated path		Increase	No
Network segment grade (Trip miles on links with…)	Grades <-2%		Decrease	No
	Grades +2% to +4%		Increase	No
	Grades > +6%		Decrease	Yes
Traffic volume (Trip miles on links with…)	< 5,000 vehicles per day		Increase	Yes
	> 30,000 vehicles per day		Decrease	No
Traffic speed (Trip miles on links with posted speeds…)	<= 20 mph		Increase	No
	> 35 mph		Decrease	Yes
Weather conditions	No significant variables		N/A	N/A

Variables only in pooled model	Trip miles on links with posted speeds 20-35 mph	Decrease	-
	Route preferences: it has few busy intersections	Decrease	-
	Trip Purpose: Exercise	Increase	-

6.   DISCUSSION

Following the presentation of data in chapters 4 and 5, discussion about the applicability of this research is merited. Limitations regarding the conclusions presented herein are then briefly discussed. Lessons learned from this research that may prove valuable to future researchers are then outlined. Finally, thoughts about the future of this research are presented.

6.1   APPLICATIONS

For transportation agencies interested in inventorying areas in their transportation networks where bicycle facilities require improvement, smartphone applications like ORcycle provide a cost-effective and high-resolution crowdsourcing solution. Transportation agencies are increasingly turning to smartphone technology to efficiently manage transportation assets and communicate with transportation users through smartphone applications managing parking supplies[18], detecting potholes[19], routing transit users[20], distributing transit tickets[21], and many more uses. This section proposes several applications for ORcycle and its resultant data sets. First, prior applications of

---

[18] https://itunes.apple.com/us/app/sfpark/id426208076?mt=8

[19] https://itunes.apple.com/us/app/street-bump/id528964742?mt=8

[20] https://itunes.apple.com/us/app/transit-app-real-time-bus/id498151501?mt=8

[21] https://itunes.apple.com/us/app/trimet-tickets/id687943985?mt=8

similar datasets are reviewed, and ideas of how ORcycle could reproduce or improve these applications considering the unique capabilities of ORcycle are discussed. Then, several new applications specific to ORcycle and its unique datasets are proposed.

6.1.1 Prior Applications

Three applications of the trip data are highlighted in an M.S. project report from Joel Meyer at the University of Texas, Austin (Meyer 2013). More information about the Austin deployment of the Cycletracks smartphone application and its results are reviewed in section 2.5.2. After those applications, another interesting application proposed by Strauss et al. (Strauss, Miranda-Moreno, and Morency 2015) is discussed. ORcycle data could be used for all of these applications and more.

6.1.1.1   *Bicycle Network Planning*

Meyer uses GPS traces in Austin to map both the observed network volumes (collected by Cycletracks) and the hypothetical network volumes that would occur if each cyclist were to use the shortest path from their origin to destination. In comparing these two flow maps, mismatches between actual use and shortest paths can help to identify where adding links to the bicycle network would have the most benefit in terms of cyclist volumes. Bicycle network planning relies to some extent on the knowing the desired paths of bicyclists. By identifying these desired paths in a comprehensive way, transportation planners can be more empirically informed about where future bicycle infrastructure improvements are necessary. ORcycle could allow this analysis to be carried out in various regions around the State of Oregon.

6.1.1.2   *Barrier Identification*

Building upon the "bicycle network planning" application, Meyer calculates the

difference between the actual bicycle volume and shortest path volume for each link in

the network to quantify the degree to which particular links are being avoided by cyclists.

This measure of avoidance can help to identify where barriers exist in the street and

bicycle network. Mitigating these barriers through infrastructure modification or

provision may be a viable and cost effective method for improving bicycle networks.

Critical connections made by mitigating barriers can have network-wide benefits for

bicycle travelers. ORcycle could help planners in Oregon to identify critical barriers in

their bicycle networks.

6.1.1.3   *Before/After Analysis*

ORcycle and Cycletracks users can record multiple trips, and can record these trips at

different points in time over the same or different geographies. This type of data presents

panel applications, of which Meyer highlights the potential for before/after analysis of

bicycle infrastructure improvements. Planners can quantify the difference in volumes

using a particular facility after it is improved; though causality of changes in bicycle

volumes from the facility installation is difficult to determine because smartphone GPS

collection is passive in nature. Perhaps more importantly, planners can use the

demographic questions associated with cyclists to see if different *types* of cyclists are

using a new facility. Further, in the interest of panel applications, planners could even see

if the comfort experienced by a single cyclist changed with the provision of new

infrastructure. This application could be crucial in validating assumptions about what

159

facility types different types of cyclists prefer. ORcycle's deep investigation of cyclist types lends itself well to this particular application.

6.1.1.4 *Crash and/or Injury Risk Models*

As reviewed in section 2.5.4.3, trip data sourced from Mon RésoVélo were used in combination with point bicycle counts and geocoded crash data to develop an injury risk model (Strauss, Miranda-Moreno, and Morency 2015). The GPS traces from the application were combined with point bicycle counts to form bicyclist exposure rates for each link in Montreal's network. The crash/injury data is then modeled over the exposure rates to model the risk of injury in the network. The data from ORcycle in combination with bicycle counts and geocoded crash data could be used to reproduce this model in the Portland area and other areas in Oregon. ORcycle has the potential to make a crash model more effective since it also collects crash information from its users.

6.1.2 Newly Proposed Applications

There are several other applications for ORcycle's unique dataset besides those reviewed above. ORcycle supplements the trip recording functionality of CycleTracks (which was used for Meyer's analysis in Austin) with many more demographic and cyclist type questions, more details about riders' trips, and adds safety problem reporting functionality. Similarly, it also offers many more survey questions than Mon RésoVélo which could be exploited for valuable insights. Several additional applications of the unique ORcycle dataset are proposed below:

6.1.3.1   *User Typology Analysis*

Factor and cluster analyses could be used on the wealth of survey data available from ORcycle to distinguish specific cyclist types empirically. While Geller's "Four types of cyclists" methodology (Geller 2006) is widely cited as a satisfactory cyclist typology for cyclist planning (in Portland and elsewhere), the methodology rests on limited analytical rigor. Geller's original categorization in 2006 made educated guesses at the proportions of the Portland population falling with the four categories (see section 2.2.1), and his proportions were approximately validated by a randomized phone survey (n=908) conducted by Dill and McNeil (Dill and McNeil 2012). However, this typology was not validated using revealed preference data, and has not been validated outside of Portland. By using statistical techniques to group the data sourced from ORcycle, Portland's cyclist typology (and other regions) can begin to approach a more accurate and empirically verified cyclist typology.

6.1.3.2   *LTS Analysis Calibration*

Many transportation agencies in Oregon (including the Oregon Department of Transportation and the Portland Bureau of Transportation) are invested in using Level of Traffic Stress (LTS) methods (see section 2.1.2) to analyze their bicycle networks and determine areas of critical need through a comprehensive, research-based methodology. ORcycle could help to calibrate this analysis to the unique conditions of Oregon's bicycle networks and user populations. For example, ORcycle could provide relative comfort differences for different bicycle facilities, which could then be modeled in relation to

measurable characteristics about those bicycle facilities. In this way, the LTS

categorization could be better attuned to Oregon cyclists needs and desires.

6.1.3.3 *Route Choice Model Improvement*

Oregon Metro's bicycle route choice model was developed based on empirical data

collected in Dill and Gliebe's landmark bicycle GPS study (Dill and Gliebe 2008). While

this model is a positive step towards an accurate bicycle route choice model, it is based

on a relatively small sample of cyclists (164 cyclists) and trips (1,449 trips). The Portland

metro area's bicycle network has also changed considerably since 2007, and is planned to

change even more drastically in the future. By collecting more data about cyclists in the

Portland metro area, ORcycle could help to improve Metro's bicycle route choice model.

In addition, other areas interested in modeling bicycle travel (such as Lane County, OR)

could use the data from ORcycle to calibrate such models. As ODOT and Oregon's other

transportation agencies make cycling an increasingly central focus of their transportation

planning efforts, it will be important to develop bicycle travel models to effectively

analyze and predict the needs of growing cycling populations.

6.1.3.4 *Safety concern identification*

Though not analyzed in depth within this thesis, ORcycle's safety report functionality

presents critical applications for identifying areas of high safety concern for cyclists. The

report data presents applications for pinpointing critical areas for bicycle infrastructure

improvement. The report data also gives cyclists another channel for reporting crash

events and conflicts, which often go unreported in cases where the crash did not seriously

harm anyone ("Pedestrian & Bicycle Information Center" 2015). Using geospatial analysis and factor-cluster methods, planners and researchers could prioritize areas in critical need of safety improvement based on the safety report data from ORcycle.

6.2   LIMITATIONS

There are several limitations to the conclusions presented from the analysis conducted herein. First, sample biases will be discussed, then the biases due to the multiple imputation algorithm used will be mentioned. Finally, biases due to small category frequencies will be briefly discussed.

6.2.1   User Sample Biases

The user, trip, and report samples were all collected between the beginning of November 2014 and the end of March 2015. Though this time period was a relatively mild winter[22] in Oregon, winter cyclists are typically different demographically than their fair-weather counterparts (Damant-Sirois, Grimsrud, and El-Geneidy 2014; Ahmed et al. 2012). Within the user sample, there are potentially biases resulting from the method of data collection; namely that it was necessary to have access to an iOS or Android smartphone to participate in the data collection. Among potential users that did own smartphones, there were also likely differences among those who would be willing to participate in the ORcycle data collection. There were also likely differences among those who uploaded

---

[22] http://www.ktvz.com/news/as-oregons-warm-winter-ends-snowpack-worries-rise/31718584

many trips and/or reports when compared to users who only uploaded minimal data. User sample biases are quantified in section 4.2 through comparisons with a travel survey dataset from the Oregon Household Activity Survey. All of these discussed biases could be potentially mitigated to some degree through a longer and more rigorous sampling period.

6.2.2 **User Participation Biases**

As with the other smartphone GPS studies reviewed (Hood, Sall, and Charlton 2011; Hudson et al. 2012) some users participated more than others by uploading more information. The user level of participation was not considered in the models presented, and is consequently biased by over-participation from some users. As illustrated in Figure 63, a few users had many more trips in the model than others, with one particular user having 53 trips included in the model. While some of these trips included by the same users were different among that user's trip set, there were also trips that were very similar to other trips they had taken. In future research, it will be necessary to test for the similarity of trips and remove these similar trips form the model set so that a certain kind of trip (or user) is not over-represented in models.

Histogram of number of trips per user in model set
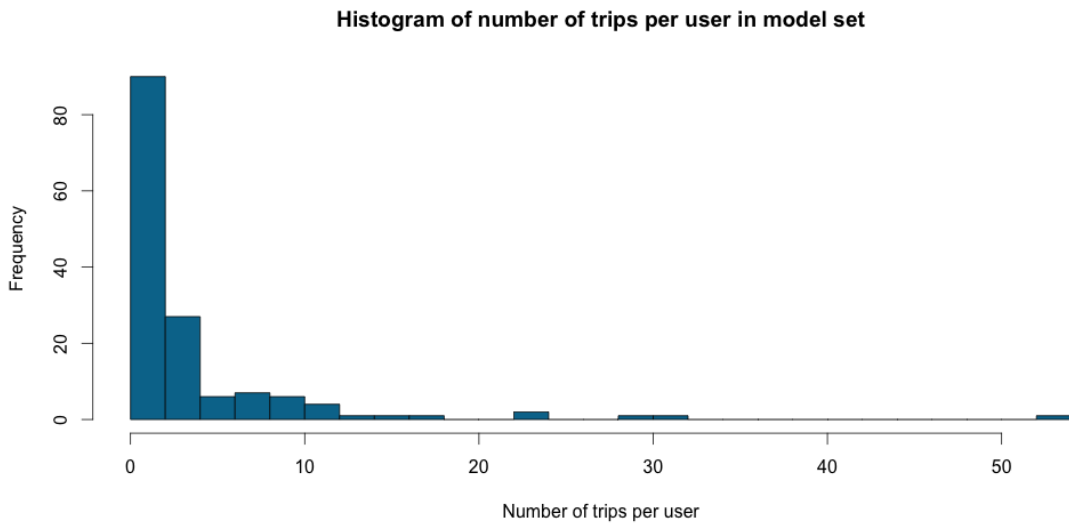
Figure 63: Histogram of number of trips per user in regression model analysis set

6.2.3 **Multiple Imputation Algorithm**

Where survey responses were missing, a multiple imputation algorithm was used to
generate likely survey responses (see section 3.2 for more information). Missing survey
responses were generated based on the other responses that had been made for a trip and
the responses that had been made for similar trips. This is especially important to
consider for the dependent variable used ("route comfort"), which was missing 30% of
the responses among the model set, as illustrated in Figure 64. These missing responses
were imputed based on the rest of the data in the model set presented in section 5. Model
specifications where route comfort was not imputed are presented in Appendix 9.4.
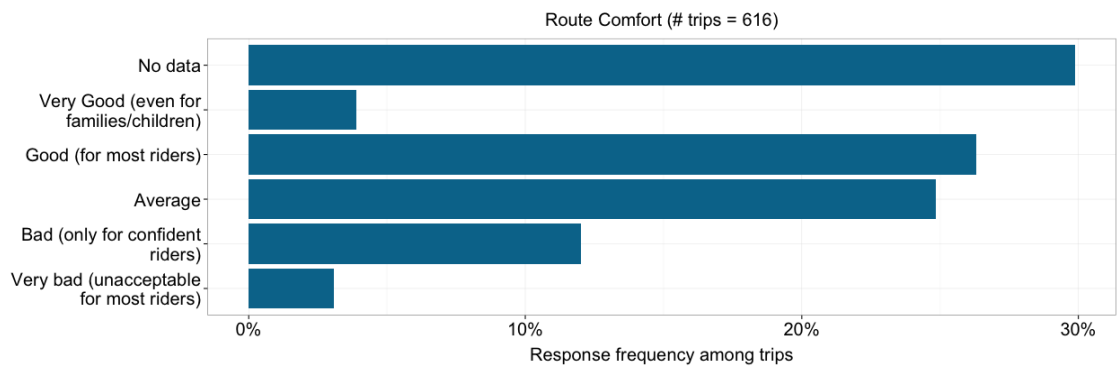
Figure 64: Route comfort distribution among model set

The multiple imputation algorithm could have systematically biased the data by
reinforcing trends because trip variables were no longer independent of one another. A
larger sample with more variation could help to mitigate the bias due to multiple
imputation by broadening the range of responses received. In future research, the bias due
to multiple imputation of missing responses should be quantified.

6.2.4 **Small Category Frequencies**

Several of the categories (especially among user variables such as occupation or ethnicity) had small frequencies. This resulted in statistically significant coefficients that were the result of small-sample bias in the maximum log-likelihood estimation used to generate the coefficients for logistic regression models[23]. While these coefficients were statistically significant, they may not be practically significant, and the inclusion of them in the models may bias the values of other model coefficients. To address this issue, categories in gender, occupation, ethnicity, and trip purpose were pooled to make larger groups. However, this did not enable interpretations of the effects of these smaller categories on route comfort. In future research, it may be valuable to increase the frequency of the small-sample categories through a larger or broader overall sample, or specifically targeted sampling efforts. There are also statistical methods available for accounting for these biases that could be employed. The coefficients could also be removed from the model if not deemed valuable, or categories could be pooled for larger frequencies.

6.3   **LESSONS LEARNED**

A project of this size and complexity required a great deal of teamwork and the development of new technical skills among all members of the research team.

---

[23] More info on small-smaple bias in logistic regression:
http://www3.nd.edu/~rwilliam/stats3/RareEvents.pdf

Developing and distributing a smartphone application capable of crowdsourcing the type of data analyzed herein is likely a task most transportation agencies are not currently equipped to handle without outside consultants; potentially making a project like this considerably more expensive. The open source code used to build the smartphone application and the web server was crucial; and this factor can help lead to the use of this type of application in other regions.

6.4   FUTURE RESEARCH

In future research on this topic, several of the limitations discussed above can be addressed or mitigated to improve the robustness of the trip comfort models proposed herein. While this thesis delves deep into analyzing the results of the trip data obtained from ORcycle, the report data obtained was not critically analyzed herein. In addition to the analysis of the report data, building a larger sample through well-executed outreach could make the conclusions presented much more robust, which opens up more possibilities for investigating the applications discussed in section 6.1. More outreach for the ORcycle smartphone application is planned, and this will likely result in a larger and more diverse sample of users, trips, and reports. This thesis provides guidance on how to utilize future samples from ORcycle and applications like it.

7.   CONCLUSIONS

The results of this research present two main sources of value for transportation researchers and planners: (1) this research provides further evidence that bicyclists are more comfortable on bicycle facilities separated from high traffic/high speed motor vehicle centric roadways; and (2) this research outlines innovative methodologies for critically analyzing the data obtained from smartphone applications crowdsourcing information from cyclists.

This research utilized statistical methods to model the reported comfort of cyclists along GPS trip trajectories. Rich geographic data were joined to these trajectories, and through these data combinations we were able to observe statistically significant differences in user-reported cyclist comfort as a function of geographic, temporal, and user-reported attributes. Bicycle trips taken on weekends were more comfortable than those taken during the week. Routes that were indicated to have stressful amounts of parked vehicles or commercial vehicle traffic were less comfortable. Separated bicycle facilities, low traffic volumes, and low traffic speeds were found to positively affect route comfort within the sample. These conclusions are in line with the body of literature surrounding bicyclist comfort evaluation, lending further evidence to the efficacy of separated, low stress bicycle facilities and/or bicycle facilities on low traffic/low speed roadways.

The innovative methods used to analyze the unique dataset provided by ORcycle provide guidance for regions interested in using bicycle data crowdsourcing applications. By outlining the methods by which this data may be analyzed, we have removed a considerable technical hurdle from the use of these types of applications elsewhere in the world.

8. WORKS CITED

Ahmed, F, Geoffrey Rose, M Figliozzi, and C Jakob. 2012. "Commuter Cyclist's Sensitivity to Changes in Weather: Insight from Two Cities with Different Climatic Conditions." In *Proceeedings of the Transportation Research Board 92nd Annual Meeting*, 1–20. Washington, DC: Transportation Research Board. http://web.cecs.pdx.edu/~maf/Conference_Proceedings/2012_Commuter_Cyclist%E2%80%99s_Sensitivity_to_Changes_in_Weather.pdf.

Ali, Asma, Cerasela Cristei, and Aimee Flannery. 2012. "Using Cumulative Logistic Regression Model for Evaluating Bicycle Faiclities on Urban Arterials." In *Proceedings of the Transportation Research Board 91st Annual Meeting*. Washington, DC: Transportation Research Board.

Bicycle Transportation Alliance. 2014. "Oregon Bicycle Transportation Alliance." Accessed February 4. http://btaoregon.org/about/#publications.

Birk, Mia, Kim Voros, Michael Rose, Roger Geller, Denver Igarta, and Brian Patterson. 2010. "Cycle Zone Analysis : An Innovative Approach to Bicycle Planning." In *Proceedings of the Transportation Research Board 89th Annual Meeting*.

Botma, H. 1995. "Method to Determine Level of Service for Bicycle Paths and Pedestrian-Bicycle Paths." *Transportation Research Record: Journal of the Transportation Research Board* (1502): 38–44. http://cat.inist.fr/?aModele=afficheN&cpsidt=3002078.

Broach, Joseph, Jennifer Dill, and John Gliebe. 2012. "Where Do Cyclists Ride? A Route Choice Model Developed with Revealed Preference GPS Data." *Transportation Research Part A: Policy and Practice* 46 (10): 1730–1740.

Carter, Daniel, William Hunter, Charles Zegger, and J. Richard Stewart. 2007. "Pedestrian and Bicyclist Intersection Safety Indices". McLean, VA.

Casello, Jeffrey M, and Vladimir Usyukov. 2014. "Modeling Cyclists' Route Choice Based on GPS Data." *Transportation Research Record: Journal of the Transportation Research Board* 2430 (1): 155–161.

Casello, Jeffrey, Akram Nour, Ks Rewa, and John Hill. 2011. "An Analysis of Stated Preference and GPS Data for Bicycle Travel Forecasting." In *90th Annual Meeting of the Transportation Research Board*. Washington, D.C.

Damant-Sirois, Gabriel, Michael Grimsrud, and Ahmed M El-Geneidy. 2014. "What's Your Type: A Multidimensional Cyclist Typology." *Transportation*: 1–17.

Davis, W. Jeffrey. 1987. "Bicycle Safety Evaluation". Chattanooga, TN.

———. 1995. "Bicycle Test Route Evaluation for Urban Road Conditions." *Transportation Congress, Volumes 1 and 2: Civil Engineers—Key to the World's Infrastructure*: 1063–1076. http://cedb.asce.org/cgi/WWWdisplay.cgi?9505730.

Dill, Jennifer, and John Gliebe. 2008. "Understanding And Measuring Bicycling Behavior." *Bicycling*. Vol. 29. Portland, OR. doi:10.1177/1091581810385956. http://www.lulu.com/items/volume_64/5687000/5687029/1/print/OTREC-RR-08-03_Dill_BicyclingBehavior_FinalReport.pdf.

Dill, Jennifer, and Nathan McNeil. 2012. "Four Types of Cyclists? Testing a Typology to Better Understand Bicycling Behavior and Potential". Portland, OR. http://web.pdx.edu/~jdill/Types_of_Cyclists_PSUWorkingPaper.pdf.

Dill, Jennifer, and Kim Voros. 2007. "Factors Affecting Bicycling Demand: Initial Survey Findings from the Portland, Oregon, Region." *Transportation Research Record: Journal of the Transportation Research Board* 2031 (1): 9–17.

Dixon, Linda. 1996. "Bicycle and Pedestrian Level-of-Service Performance Measures and Standards for Congestion Management Systems." *Transportation Research Record: Journal of the Transportation Research Board* (1538): 1–9.

Emery, J., and C. Crump. 2003. "The WABSA Project: Assessing and Improving Your Community's Walkability & Bikeability". Chapel Hill, NC.

Epperson, B. 1994. "Evaluating Suitability of Roadways for Bicycle Use: Toward a Cycling Level-of-Service Standard." *Transportation Research Record: Journal of the Transportation Research Board* 1438: 9–16.

Fernández-Heredia, Álvaro, Andres Monzon, and Sergio Jara-Díaz. 2014. "Understanding Cyclists' Perceptions, Keys for a Successful Bicycle Promotion." *Transportation Research Part A: Policy and Practice* 63: 1–11.

Figliozzi, Miguel, Bryan Blanc, and Pamela Johnson. 2014. "Evaluating the Use of Crowdsourcing as a Data Collection Method for Bicycle Performance Measures and Identification of Facility Improvement Needs, Tasks 1 to 3 Report". Salem, OR.

Foster, Nick, Christopher M Monsere, Jennifer Dill, and Kelly Clifton. 2015. "A Level-of-Service Model for Protected Bike Lanes 3." In *Transportation Research Board 94th Annual Meeting*. Washington, D.C.

Geller, Roger. 2006. "Four Types of Cyclists". Portland, OR. https://www.portlandoregon.gov/transportation/article/237507.

Harkey, David, Donald Reinfurt, and Matthew Knuiman. 1998. "Development of the Bicycle Compatibility Index." *Transportation Research Record*. doi:10.3141/1636-03.

Harvey, Francis, Kevin J Krizek, and Reuben Collins. 2008. "Using GPS Data to Assess Bicycle Commuter Route Choice." In *Transportation Research Board 87th Annual Meeting*.

Haworth, Narelle, and Amy Schramm. 2011. "How Do Level Of Experience , Purpose For Riding And Preference For Facilities Affect Location Of Riding ? A Study Of Adult Bicycle Riders In Queensland , Australia." In *Proceedings of the Transportation Research Board 90th Annual Meeting*. Washington, DC: Transportation Research Board.

Hood, Jeffrey, Elizabeth Sall, and Billy Charlton. 2011. "A GPS-Based Bicycle Route Choice Model for San Francisco, California." *Transportation Letters: The International Journal of Transportation Research* 3 (1) (January 1): 63–75. doi:10.3328/TL.2011.03.01.63-75. http://jrosspub.metapress.com/openurl.asp?genre=article&id=doi:10.3328/TL.2011.03.01.63-75.

Hudson, Joan G, Jennifer C Duthie, Yatinkumar K Rathod, Katie A Larsen, and Joel L Meyer. 2012. "Using Smartphones to Collect Bicycle Travel Data in Texas". Austin, TX.

Jackson, Stewart, Luis F. Miranda-Moreno, Colin Rothfels, and Yannick Roy. 2014. "Adaption and Implementation of a System for Collecting and Analyzing Cyclist Route Data Using Smartphones." In *Proceedings of the Transportation Research Board 93rd Annual Meeting*, 1–18. Washington, DC: Transportation Research Board.

Jensen, Soren Underlien. 2007. "Pedestrian and Bicyclist Level of Service on Roadway Segments." In *Proceedings of the Transportation Research Board 87th Annual Meeting*, 1–18. http://trb.metapress.com/index/T041T470008UW836.pdf.

———. 2013. "Pedestrian and Bicycle Level of Service at Intersections, Roundabouts, and Other Crossings." *Transportation Research Board 92nd Annual Meeting* (July 2012): 1–19. http://trafitec.dk/sites/default/files/publications/ped bike los at intersec.pdf.

Jones, Elizabeth, and Timothy Carlson. 2003. "Development of Bicycle Compatibility Index for Rural Roads in Nebraska." *Transportation Research Record: Journal of the Transportation Research Board* 1828: 124–132. doi:10.3141/1828-15.

Landis, Bruce W. 1994. "Bicycle Interaction Hazard Score: A Theoretical Model." *Transportation Research Record: Journal of the Transportation Research Board* 1438: 3–8.

Lowry, M, and D Callister. 2012. "Using Bicycle Level of Service to Assess Community-Wide Bikeability." In *Proceedings of the Transportation Research Board 91st Annual Meeting*, 1–15. Washington, DC: Transportation Research Board. http://docs.trb.org/prp/12-4154.pdf.

McNeil, N. 2011. "Bikeability and the 20-Min Neighborhood: How Infrastructure and Destinations Influence Bicycle Accessibility." *Transportation Research Record* (2247): 53–63. doi:10.3141/2247-07. http://www.scopus.com/inward/record.url?eid=2-s2.0-84855236221&partnerID=40&md5=68d2042ab7c0f9b65f63aa323ca4af74.

Mekuria, MC, PG Furth, and H Nixon. 2012. "Low-Stress Bicycling and Network Connectivity". San Jose, CA. http://ntl.bts.gov/lib/45000/45100/45148/1005-low-stress-bicycling-network-connectivity-brief.pdf.

Menghini, Gianluca, Nelson Carrasco, Nadine Schüssler, and Kay W Axhausen. 2010. "Route Choice of Cyclists in Zurich." *Transportation Research Part A: Policy and Practice* 44 (9): 754–765.

Meyer, Joel Loren. 2013. "Big Bike Data: How GPS Route Data Collected from Smartphones Can Benefit Bicycle Planning". Austin, TX.

Misra, Aditi, Aaron Gooze, Kari Watkins, Miriam Asad, and Chrisptopher A. Le Dantec. 2014. "Crowdsourcing and Its Application to Transportation Data Collection and Management." In *Proceedings of the Transportation Research Board 93rd Annual Meeting*, 1–16. Washington, DC: Transportation Research Board.

Noël, N., C. Leclerc, and M. Lee-Gosselin. 2003. "CRC INDEX : Compatibility of Roads for Cyclists in Rural and Urban Fringe Areas." In *Proceedings of the Transportation*

Research Board 82nd Annual Meeting. Washington, DC: Transportation Research Board.

"Pedestrian & Bicycle Information Center." 2015. Accessed April 15. http://www.pedbikeinfo.org/data/factsheet_crash.cfm.

Pedestrian and Bicyle Information Center. 2002. "Bikeability Checklist: How Bikeable Is Your Community?" http://www.pedbikeinfo.org/pdf/bikeability_checklist.pdf.

Petritsch, Theodore A., Bruce W. Landis, Herman F. Huang, Peyton S. McLeod, Daniel Lamb, Waddah Farah, and Martin Guttenplan. 2008. "Bicycle Level of Service for Arterials." *Transportation Research Record*. doi:10.3141/2031-05.

Pucher, John, Ralph Buehler, and Mark Seinen. 2011. "Bicycling Renaissance in North America? An Update and Re-Appraisal of Cycling Trends and Policies." *Transportation Research Part A: Policy and Practice* 45: 451–475. http://policy.rutgers.edu/faculty/pucher/TRA960_01April2011.pdf.

Pucher, John, Jennifer Dill, and Susan Handy. 2010. "Infrastructure, Programs, and Policies to Increase Bicycling: An International Review." *Preventive Medicine* 50: S106–S125.

Reid, J. 2011. "Market Segmentation of Cyclists–understanding Attitudes toward Safety." In *Australasian College of Road Safety Conference*. Melbourne, Victoria, Australia.

San Francisco Department of Public Health. 2009. "Bicycle Environmental Quality Index (BEQI) Draft Report". San Francisco, CA.

Schuessler, Nadine, and Kay W Axhausen. 2009. "Processing Raw Data from Global Positioning Systems without Additional Information." *Transportation Research Record: Journal of the Transportation Research Board* 2105 (1): 28–36.

Schuessler, Nadine, Kay W Axhausen, and E T H Zurich. 2009. "Map-Matching of GPS Traces on High-Resolution Navigation Networks Using the Multiple Hypothesis Technique (MHT)."

Schwartz, Michael, and Jeff Hood. 2011. "Bicycle Route Choice Data Collection Using GPS-Enabled Smartphones": 1–10.

Sener, Ipek N., Naveen Eluru, and Chandra R. Bhat. 2009. "An Analysis of Bicycle Route Choice Preferences in Texas, US." *Transportation* 36 (5) (April 16): 511–

539. doi:10.1007/s11116-009-9201-4. http://link.springer.com/10.1007/s11116-009-9201-4.

Sorton, A, and T Walsh. 1994. "Bicycle Stress Level as a Tool to Evaluate Urban and Suburban Bicycle Compatibility." *Transportation Research Record: Journal of the Transportation Research Board* 1438: 17–24.

Stein, Bill. 2011. "Regional Bicycle Demand Model: In Use Today in Portland." In *Transportation Research Board Applications Conference*. Reno, NV.

Strauss, Jillian, Luis F Miranda-Moreno, and Patrick Morency. 2015. "Mapping Cyclist Activity and Injury Risk in a Network Combining Smartphone GPS Data and Bicycle Counts." In *Transportation Research Board 94th Annual Meeting*.

Teschke, Kay, and Meghan Winters. 2013. "Motivating Cycling." *Cycling in Cities*. http://cyclingincities.spph.ubc.ca/opinion-survey/.

Transportation Research Board. 2010. *Highway Capacity Manual 2010*. Washington, DC: Transportation Research Board.

Turner, S, CS Shafer, and WP Stewart. 1997. "Bicycle Suitability Criteria for State Roadways in Texas". College Station, TX. http://safety.fhwa.dot.gov/ped_bike/docs/txdot_3988s.pdf.

"Walk Score." 2014. Accessed April 2. http://www.walkscore.com/.

Winters, Meghan, Gavin Davidson, Diana Kao, and Kay Teschke. 2011. "Motivators and Deterrents of Bicycling: Comparing Influences on Decisions to Ride." *Transportation* 38 (1): 153–168.

Zolnik, Edmund J., and Ellen K. Cromley. 2008. "Poisson Multilevel Methodology of Bicycle Levels of Service for Road Networks." *Transportation Research Record*. doi:10.3141/2031-01.

Zorn, Lisa, Elizabeth Sall, and Matthew Bomber. 2011. "Completing The Cycle: Incorporating Cycletracks Into Sf-Champ." In *Proceedings of the 4th Transportation Research Board Conference on Innovations in Travel Modeling*. Washington, DC: Transportation Research Board. http://www.sfcta.org/sites/default/files/content/IT/CycleTracks/CompletingTheCycle_DRAFTNOV2011.pdf.

APPENDIX A

A.1 USER SAMPLE COMPARISON

Table 65: Age sample comparison

Sample	OHAS	OHAS Bicycle Commuters	ORcycle
N	45,695	802	339
< 18	20.0 %	1.6 %	0.3 %
18-24	4.0 %	5.9 %	4.7 %
25-34	6.4 %	12.5 %	32.1 %
35-44	10.8 %	23.8 %	30.2 %
45-54	16.5 %	29.8 %	19.5 %
55-64	21.1 %	23.6 %	10.1 %
65+	21.2 %	2.9 %	3.1 %

Table 66: Gender sample comparison

Sample	OHAS	OHAS Bicycle Commuters	ORcycle
N	46,368	818	335
Female	52.2 %	33.7 %	17.8 %
Male	47.8 %	66.3 %	82.2 %

Table 67: Ethnicity sample comparison

Sample	OHAS	OHAS Bicycle Commuters	ORcycle
N	19,526	711	332
African American	0.5%	0.0%	0.3%
Asian American	1.0%	2.0%	0.3%
Hispanic	2.8%	2.4%	5.9%
Native American	0.9%	0.3%	1.6%
White American	93.8%	94.4%	87.6%
Other	1.0%	1.0%	4.2%

Table 68: Household Income sample comparison

Sample	OHAS	OHAS Bicycle Commuters	ORcycle
N	18,637	690	316
$0-$14,999	6.7%	2.9%	6.3%
$15,000-$24,999	10.4%	5.5%	4.5%
$25,000-$34,999	9.8%	7.4%	7.7%
$35,000-$49,999	14.2%	10.3%	6.6%
$50,000-$74,999	23.1%	25.9%	20.6%
$75,000-$99,999	17.2%	23.3%	24.0%
$100,000-$149,999	12.8%	17.0%	22.0%
$150,000 or more	5.8%	7.7%	8.4%

Table 69: Household number of vehicles comparison

Sample	OHAS	OHAS Bicycle Commuters	ORcycle
N	19,932	736	339
0 Vehicles	4.3 %	4.9 %	13.8 %
1 Vehicles	27.8 %	34.6 %	39.2 %
2 Vehicles	40.5 %	41.4 %	37.6 %
3 or more Vehicles	27.3 %	19 %	9.4 %

Table 70: Household number of workers comparison

Sample	OHAS	OHAS Bicycle Commuters	ORcycle
N	19,932	736	334
0 Workers	23.9 %	0.4 %	7.4 %
1 Worker	36.9 %	27.3 %	35.6 %
2 Workers	34.5 %	61.3 %	52.2 %
3 or more Workers	4.7 %	11 %	4.8 %

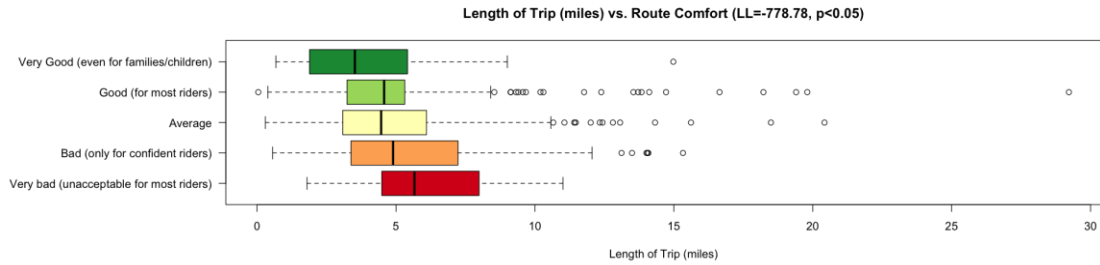## A.2   ROUTE COMFORT DISTRIBUTION PLOTS

A.2.1 Trip Attributes

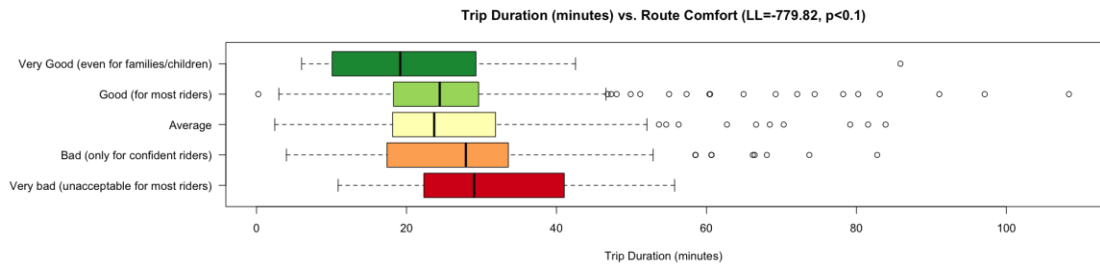**Figure 65: Trip Length distribution over Route Comfort**

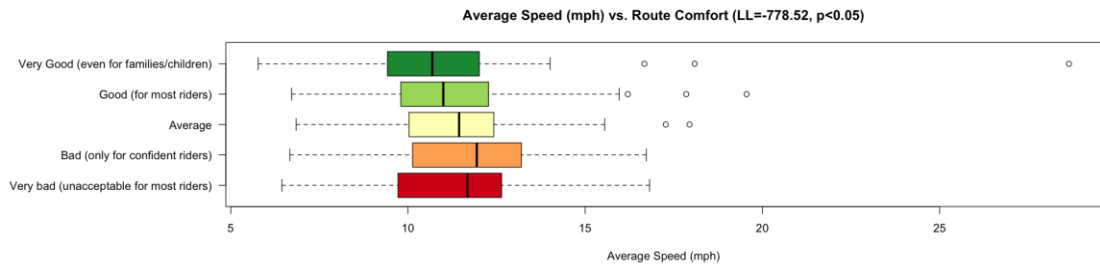**Figure 66: Trip Duration distribution over Route Comfort**
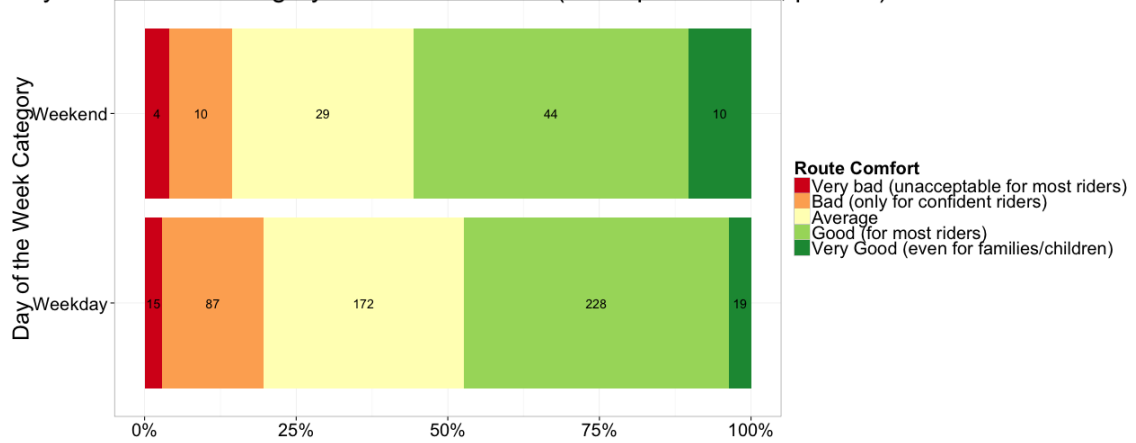

Figure 67: Average Speed distribution over Route Comfort

A.2.2 Temporal Characteristics

Day of the Week Category vs. Route Comfort (Chi-square=10.57, p<0.05)

Figure 68: Route Comfort distribution among Day of the Week Category

Start Time Category vs. Route Comfort (Chi-square=8.65, NS)
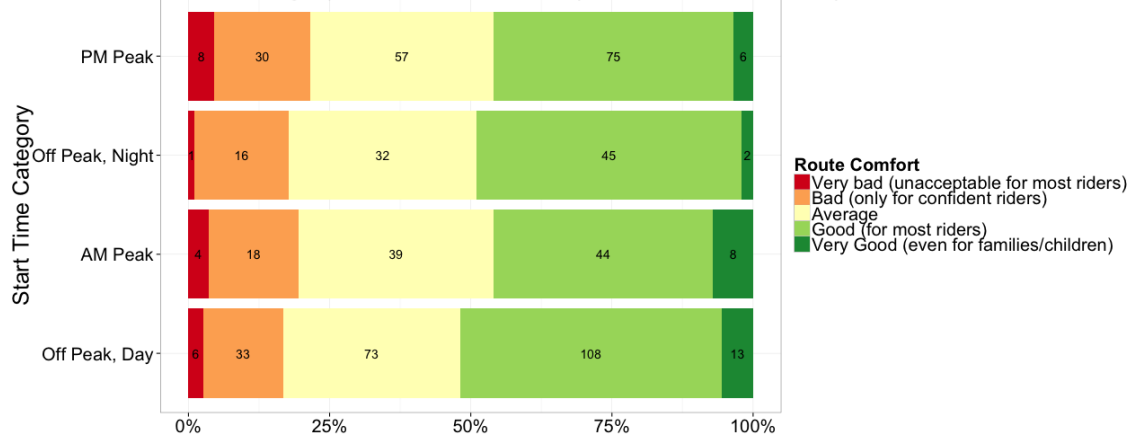
Figure 69: Route Comfort distribution among Start Time categories

A.2.3 Trip Question Responses
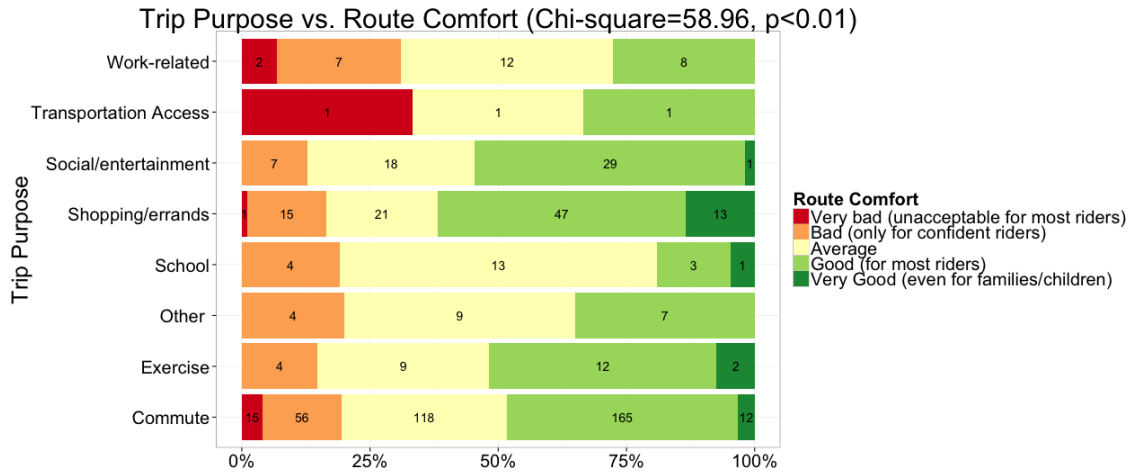
Figure 70: Route Comfort distribution among Trip Purpose
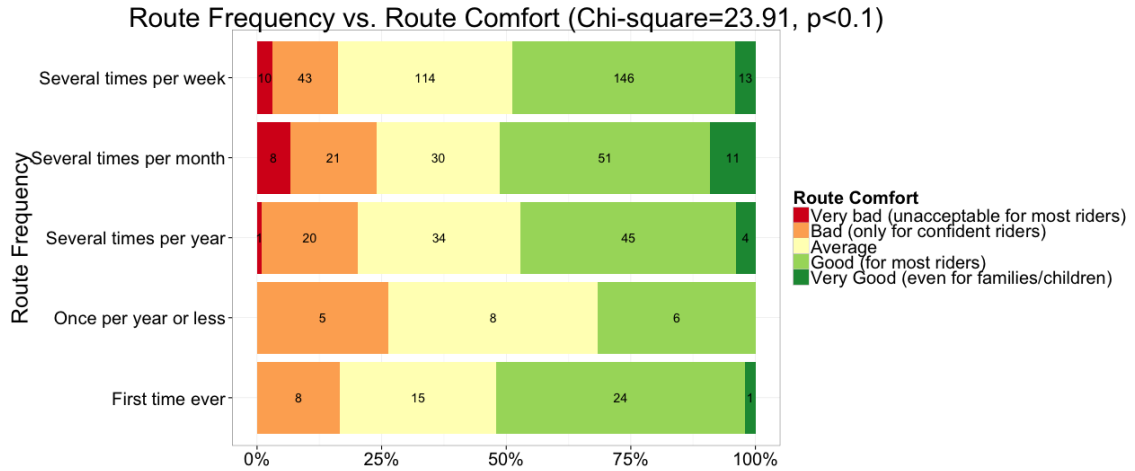
Figure 71: Route Comfort distribution among Route Frequency
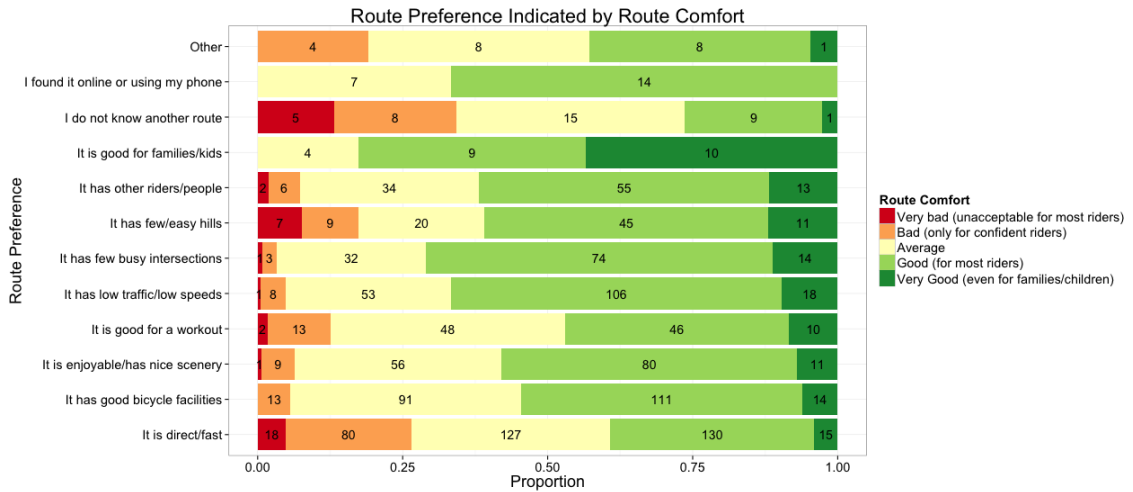
Figure 72: Route Comfort distribution among Route Choice Preferences
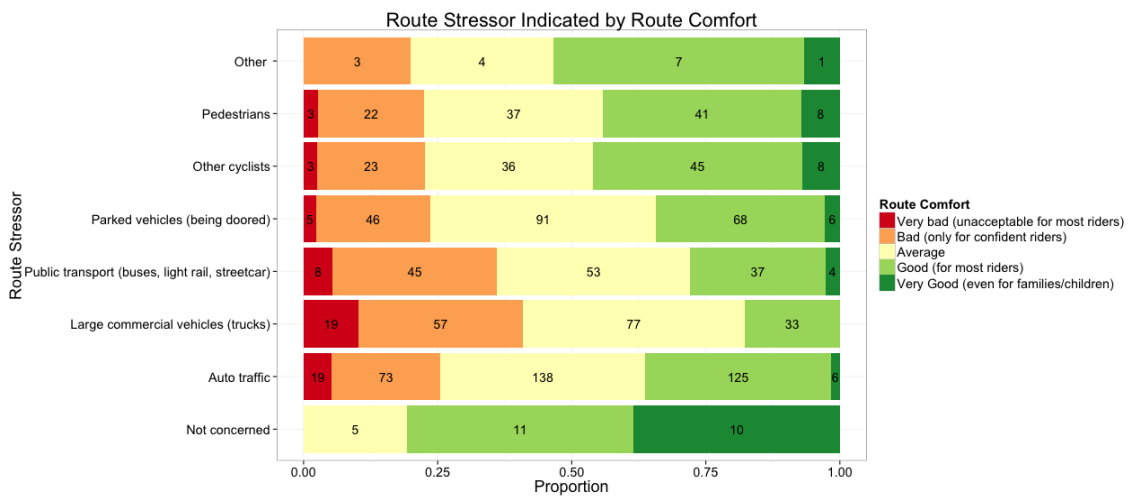
Figure 73: Route Comfort distribution among Route Stressors

A.2.4 User Question Responses
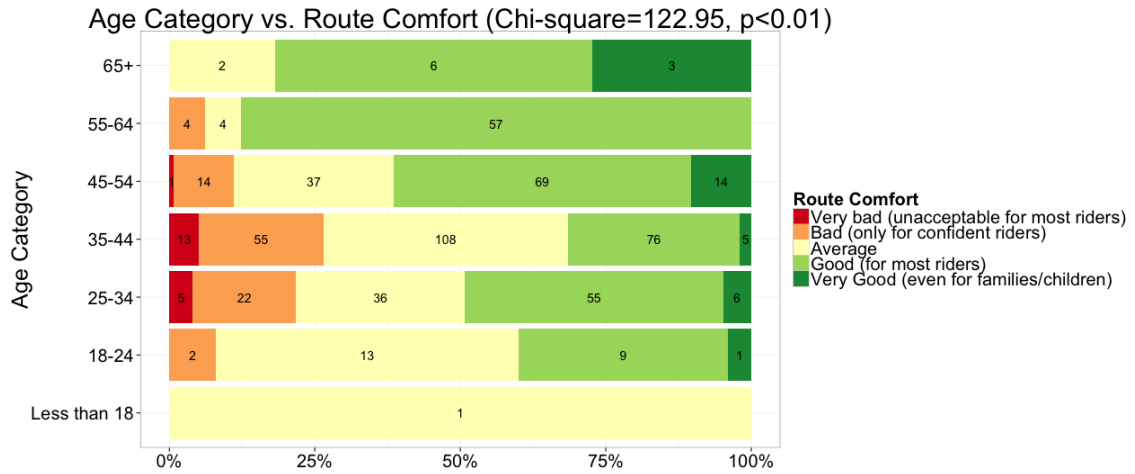
Figure 74: Route Comfort distribution among Age categories
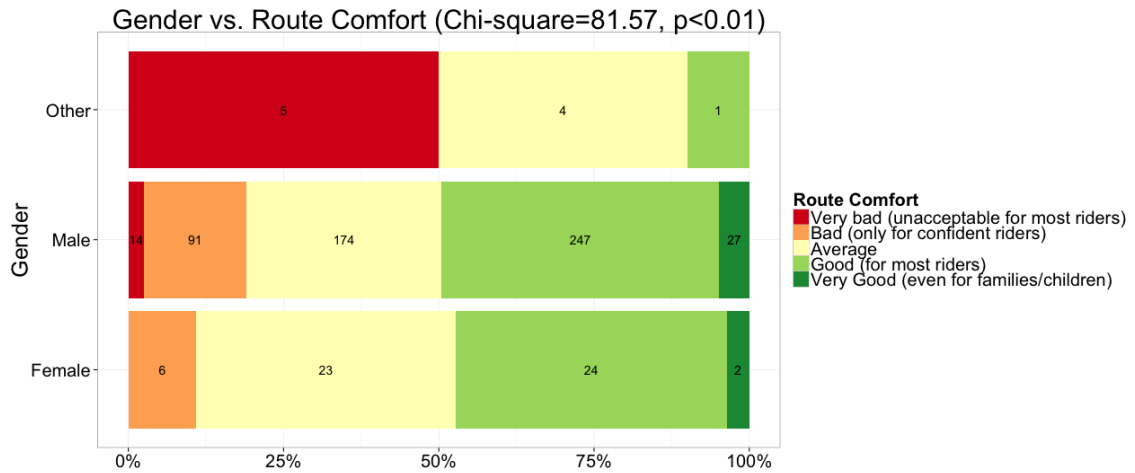
Figure 75: Route Comfort distribution among Gender categories
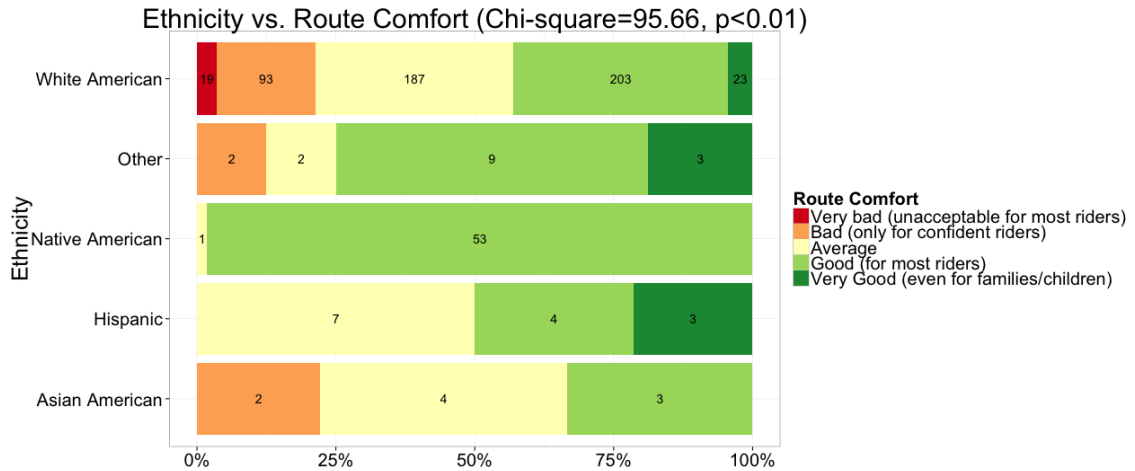
Figure 76: Route Comfort distribution among Ethnicity
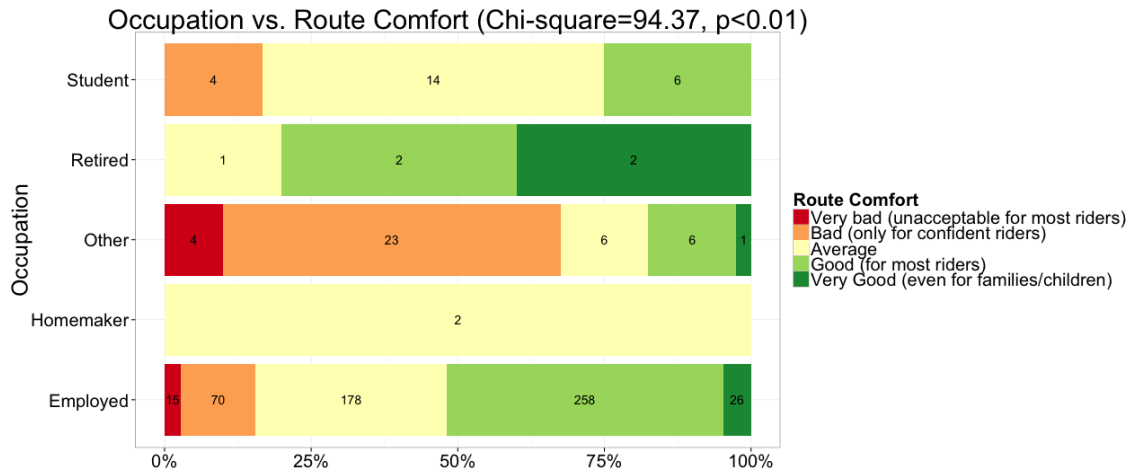
Figure 77: Route Comfort distribution among Occupation

Household Workers Category vs. Route Comfort (Chi-square=29.04, p<0.05)
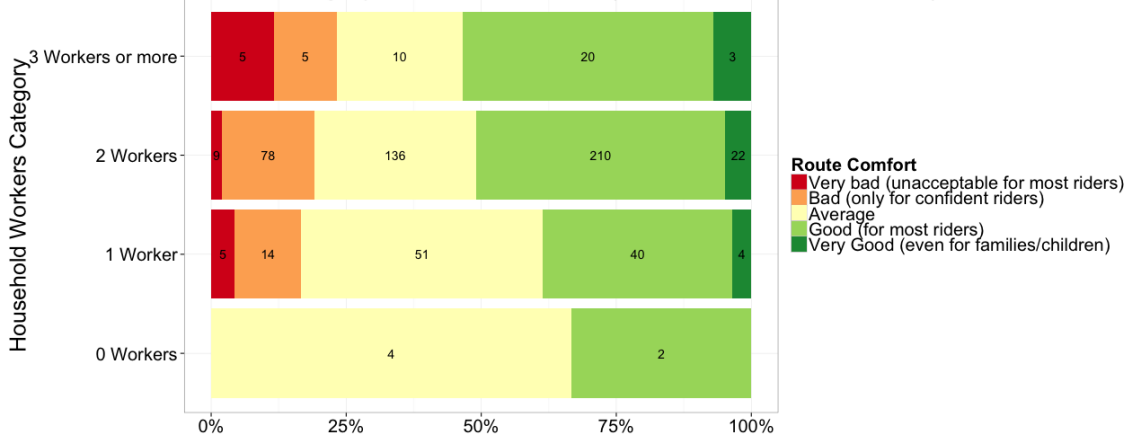
Figure 78: Route Comfort distribution among Household Workers

Household Income Category vs. Route Comfort (Chi-square=100, p<0.01)
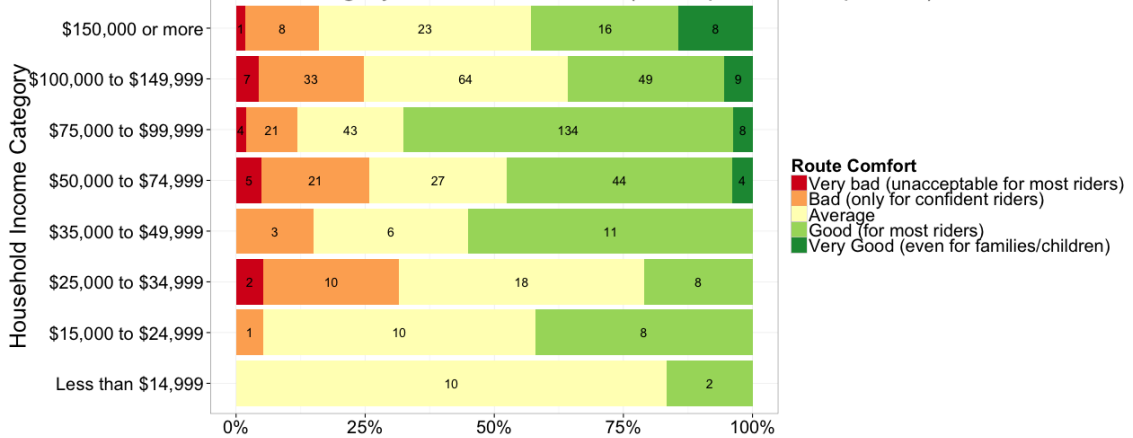
Figure 79: Route Comfort distribution among Household Income Category

Household Vehicles Category vs. Route Comfort (Chi-square=71.21, p<0.01)
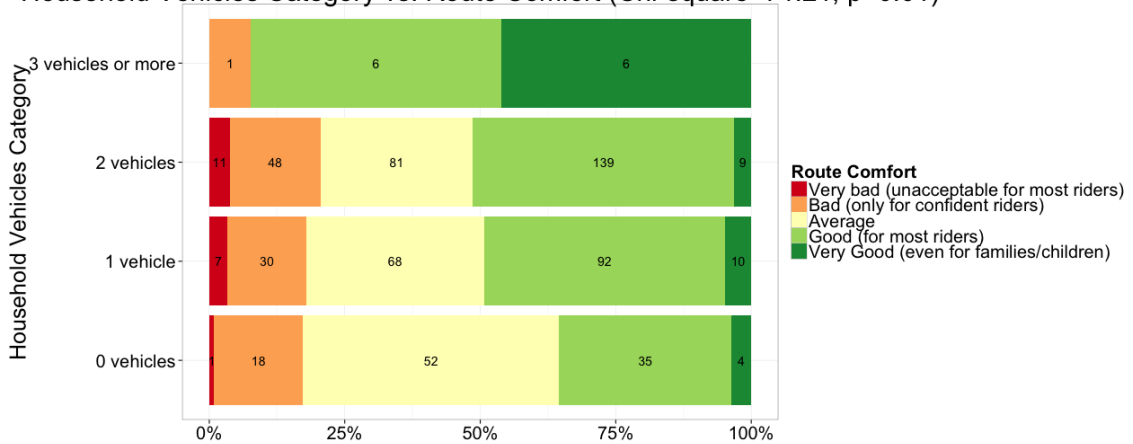
Figure 80: Route Comfort distribution among Household Vehicles

Number of Bicycles Owned vs. Route Comfort (Chi-square=59.11, p<0.01)
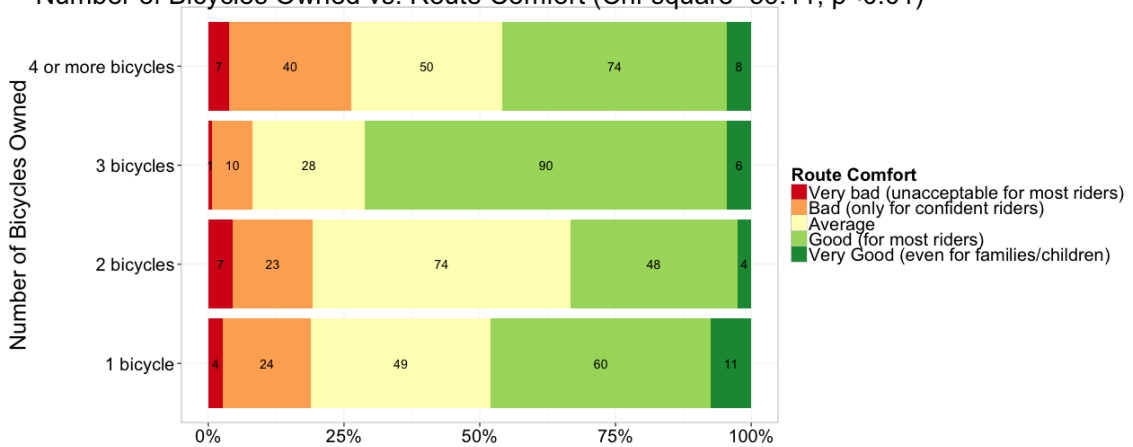
Figure 81: Route Comfort distribution among Number of Bicycles Owned
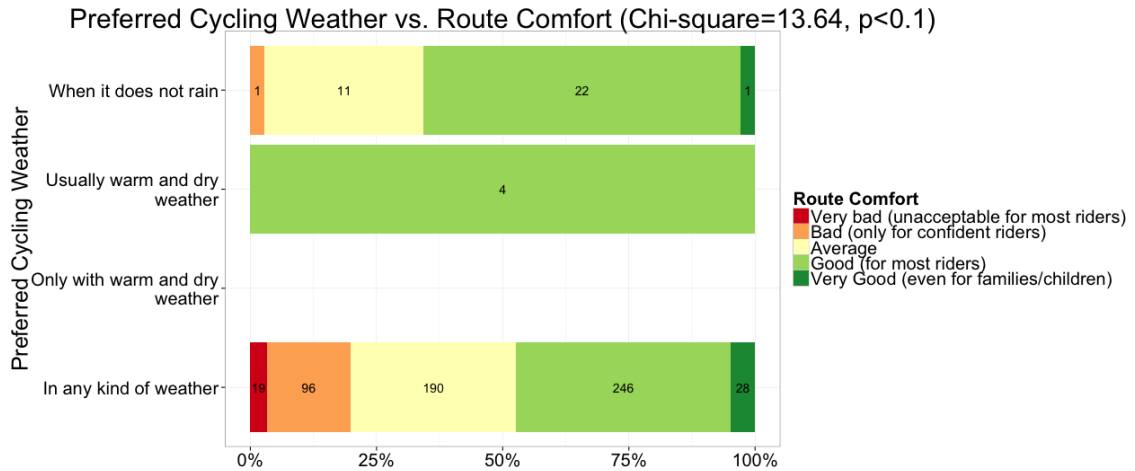
Figure 82: Route Comfort distribution among Preferred Cycling Weather
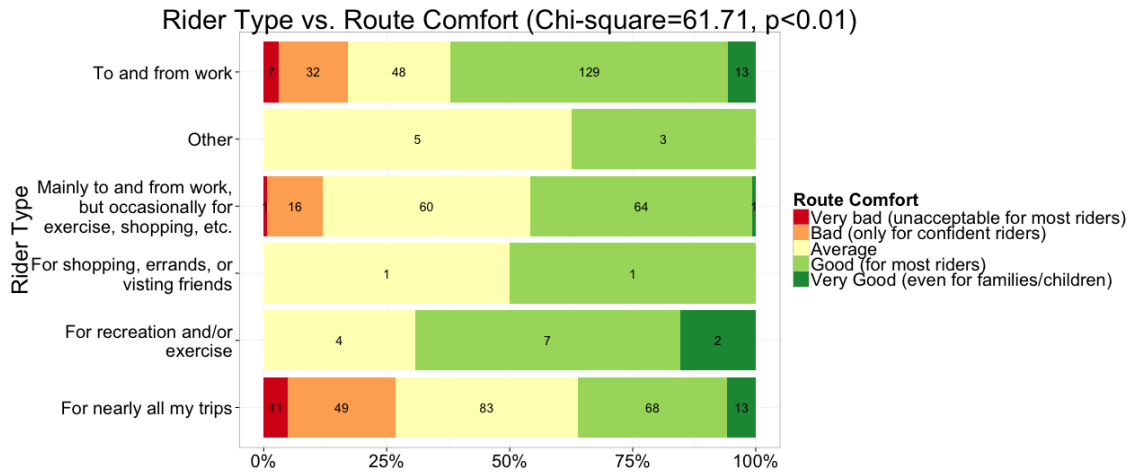
Figure 83: Route Comfort distribution among Rider Type

187
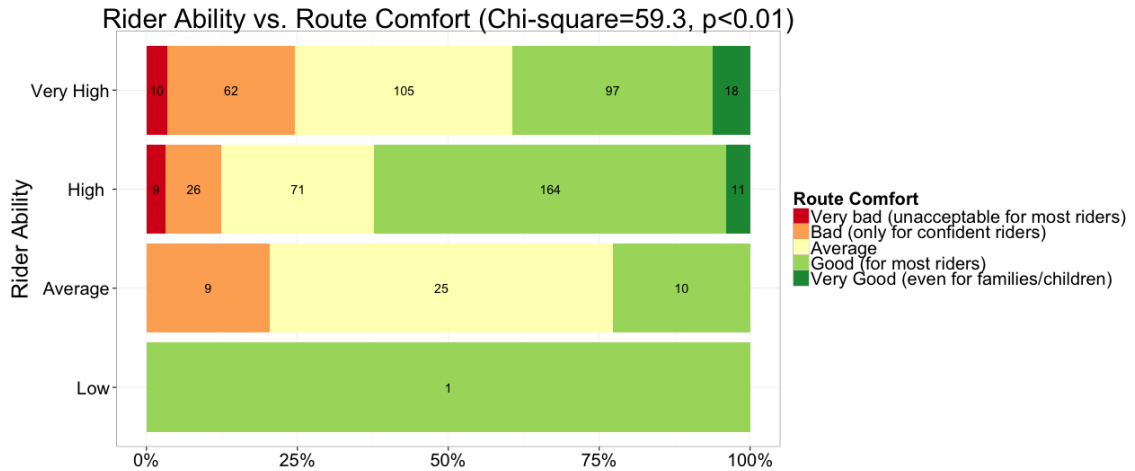
Figure 84: Route Comfort distribution among Rider Ability
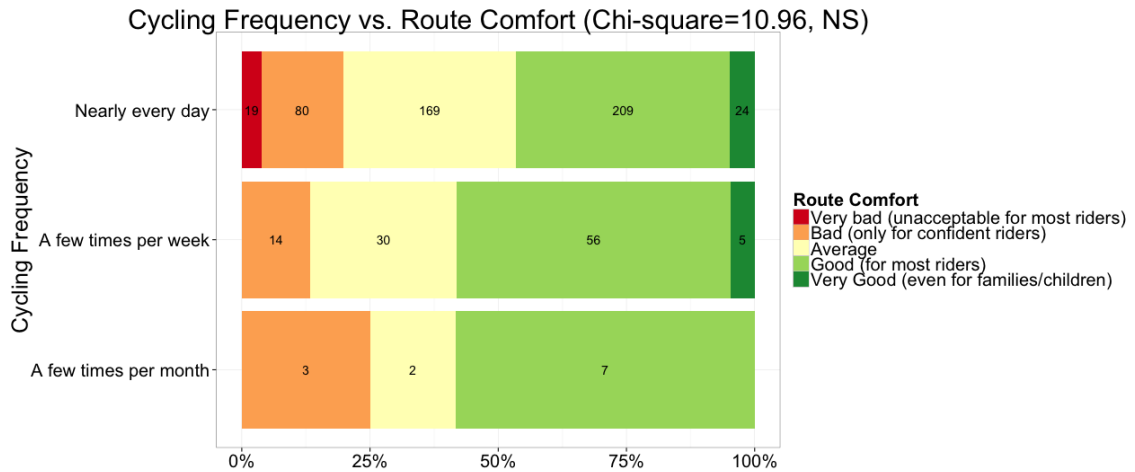
Figure 85: Route Comfort distribution Cycling Frequency

A.2.5 Bicycle Facility and Street Type
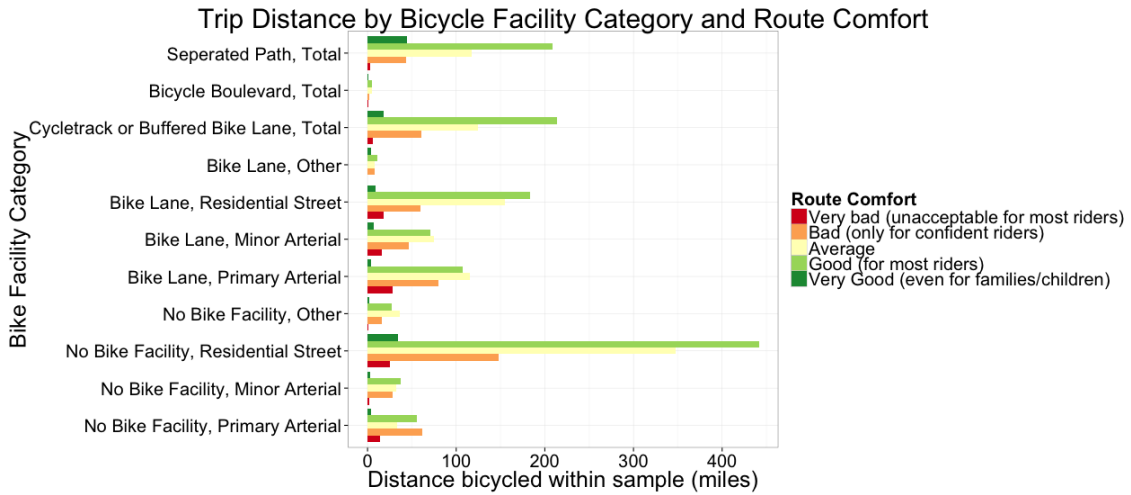
Figure 86: Route Comfort vs. distance bicycled on links with different bicycle facility types
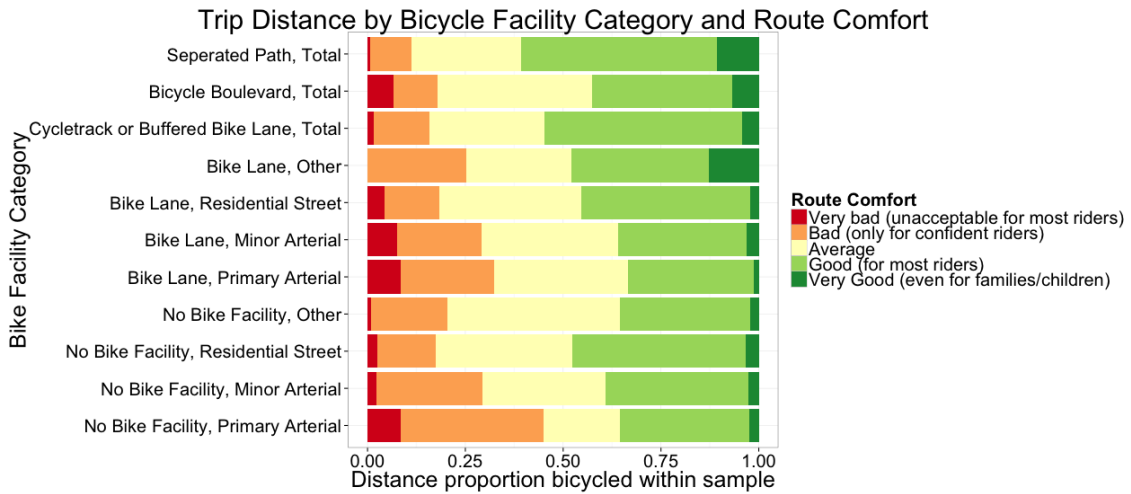
Figure 87: Route Comfort vs. distance proportion bicycled on links with different bicycle facility types

A.2.6 Topography
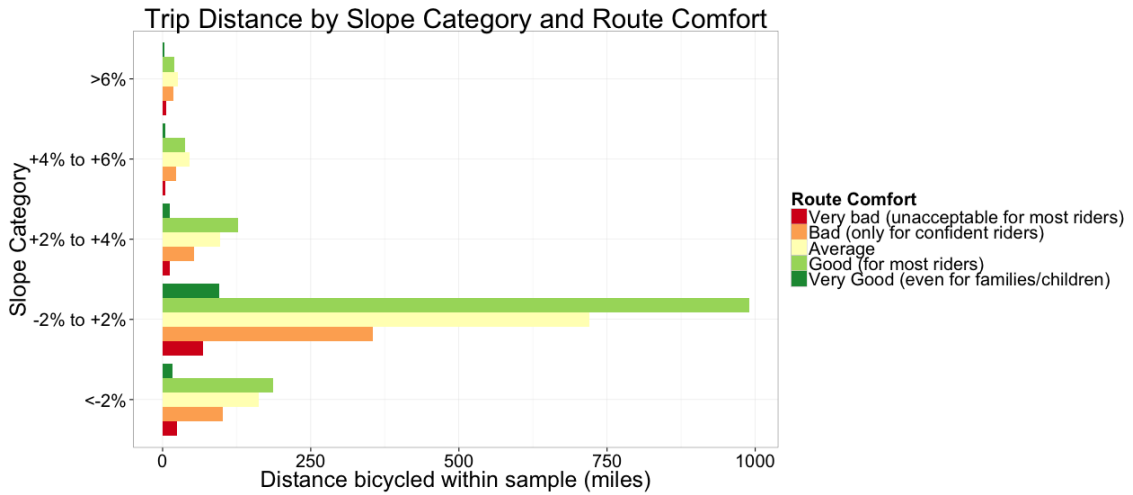
Figure 88: Route Comfort vs. distance bicycled on links with different average slopes
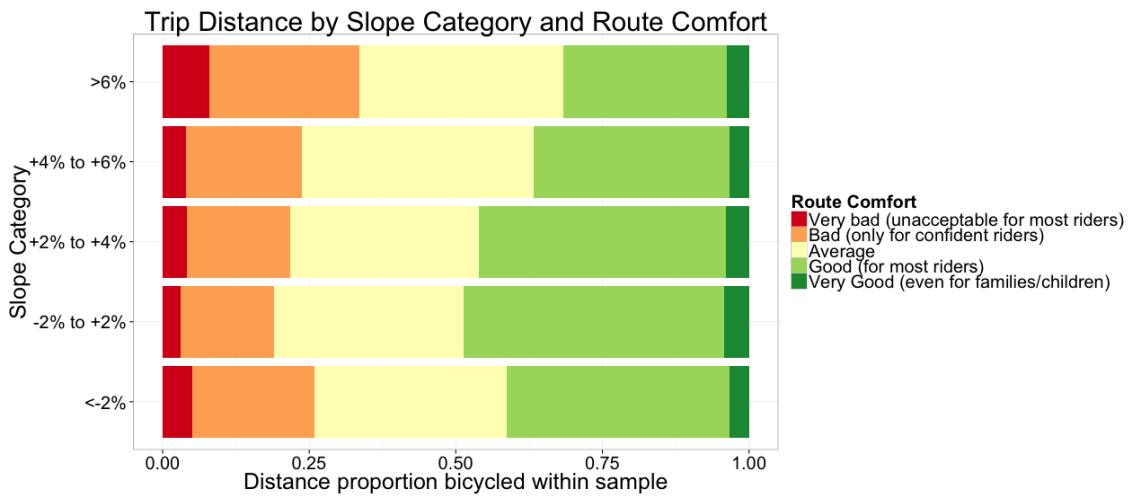
Figure 89: Route Comfort vs. distance proportion bicycled on links with different average slopes

A.2.7 Traffic Volume
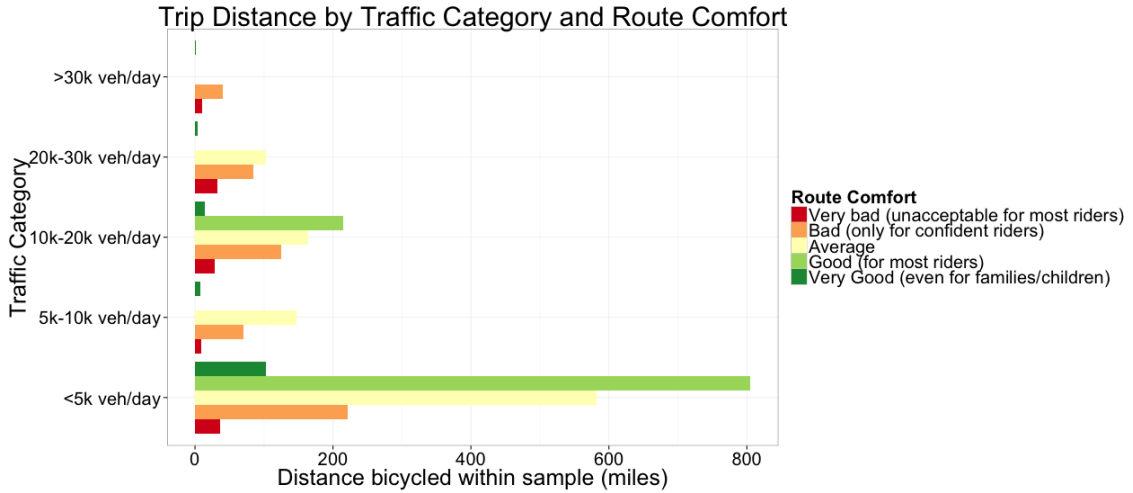
Figure 90: Route Comfort vs. distance bicycled on links with different traffic volumes
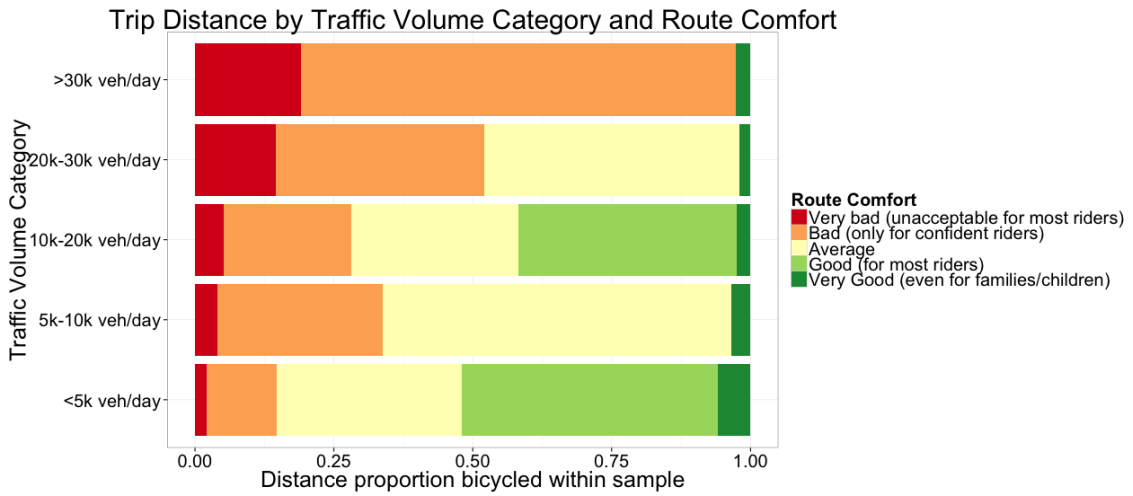
Figure 91: Route Comfort vs. distance proportion bicycled on links with different traffic volumes

A.2.8 Traffic Speed
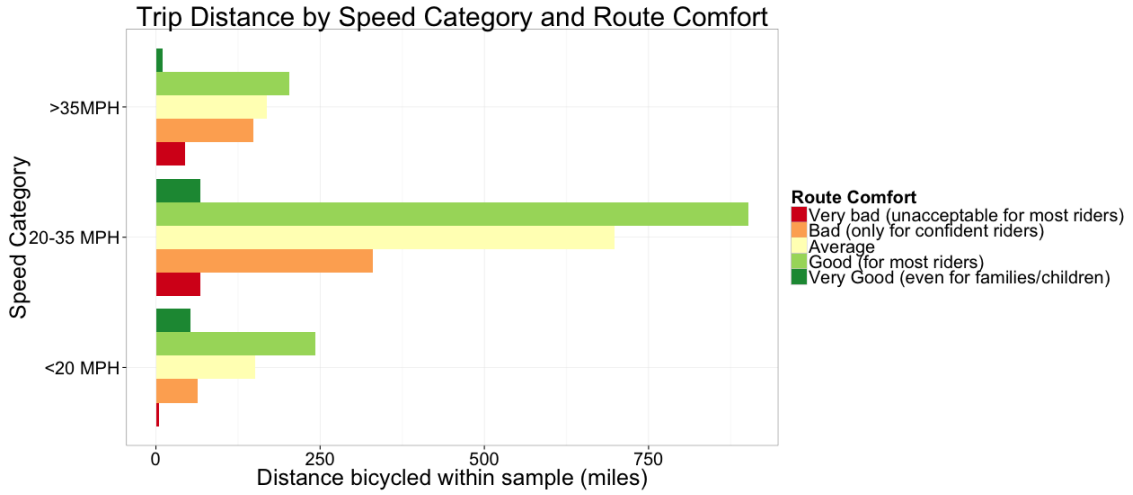
Figure 92: Route Comfort vs. distance bicycled on links with different traffic speeds
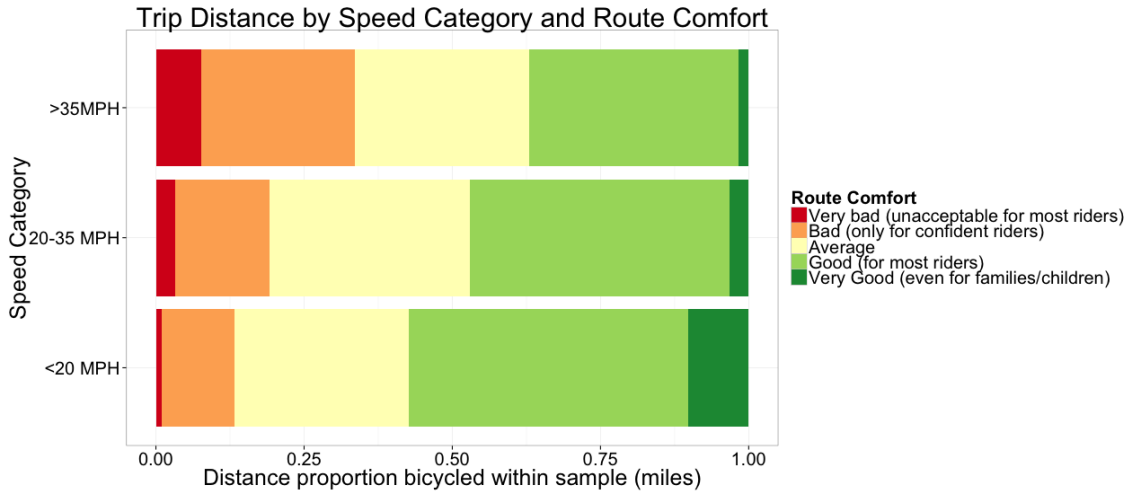
Figure 93: Route Comfort vs. distance proportion bicycled on links with different traffic speeds

A.2.9 Weather Variables

Temperature (Degrees F) vs. Route Comfort (LL=-782.93, NS)
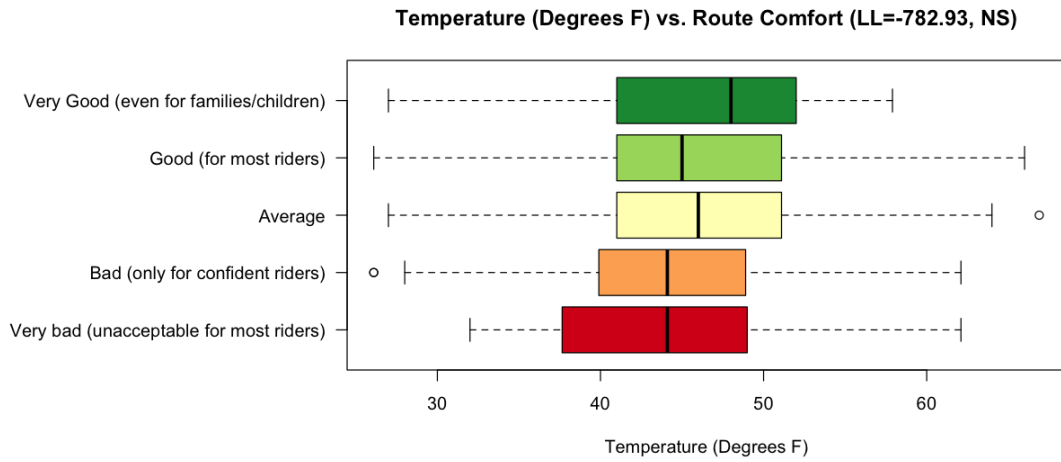
Figure 94: Temperature distribution over Route Comfort

Hourly Precipitation (inches) vs. Route Comfort (LL=-783.22, NS)
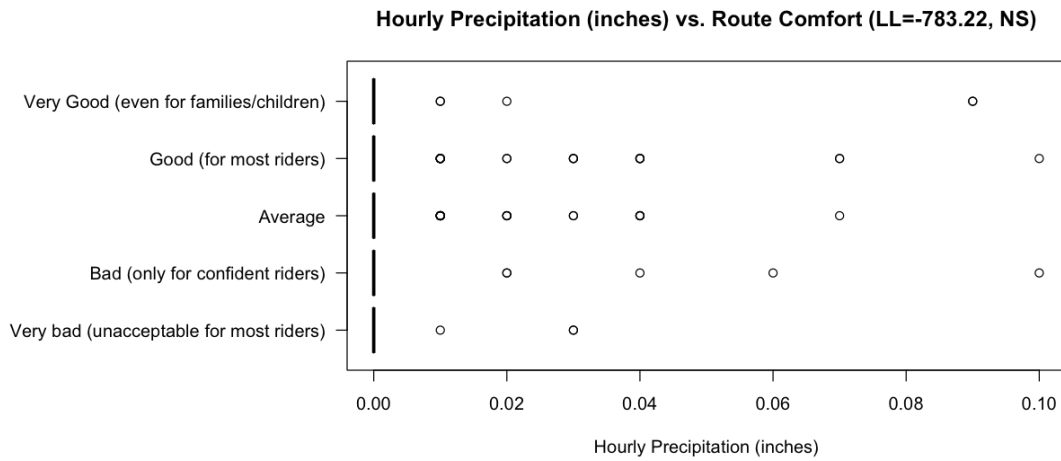
Figure 95: Hourly Precipitation distribution over Route Comfort
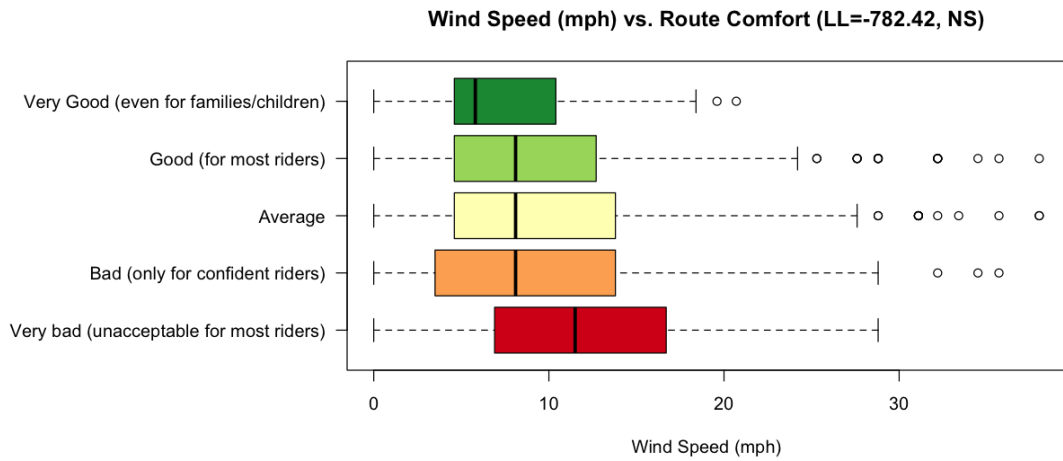
Figure 96: Wind Speed distribution over Route Comfort
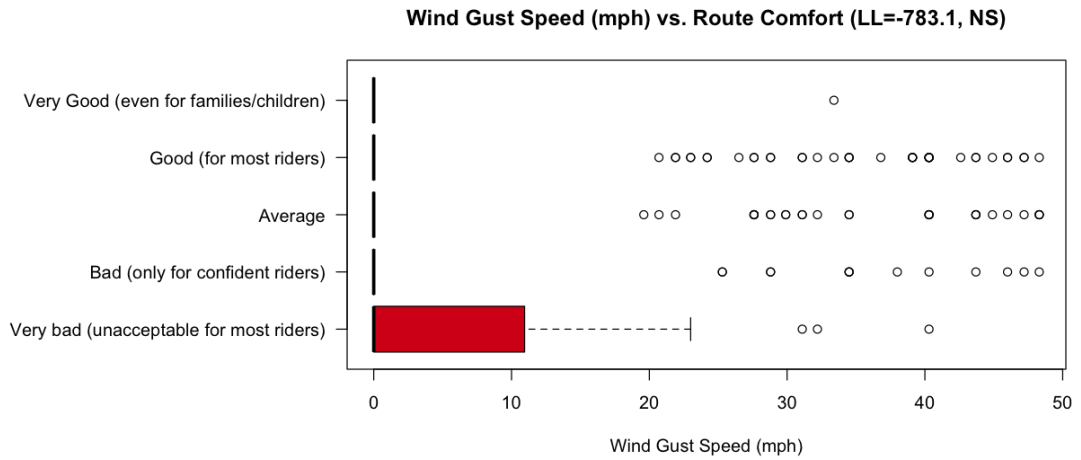
Figure 97: Wind Gust Speed distribution over Route Comfort

194
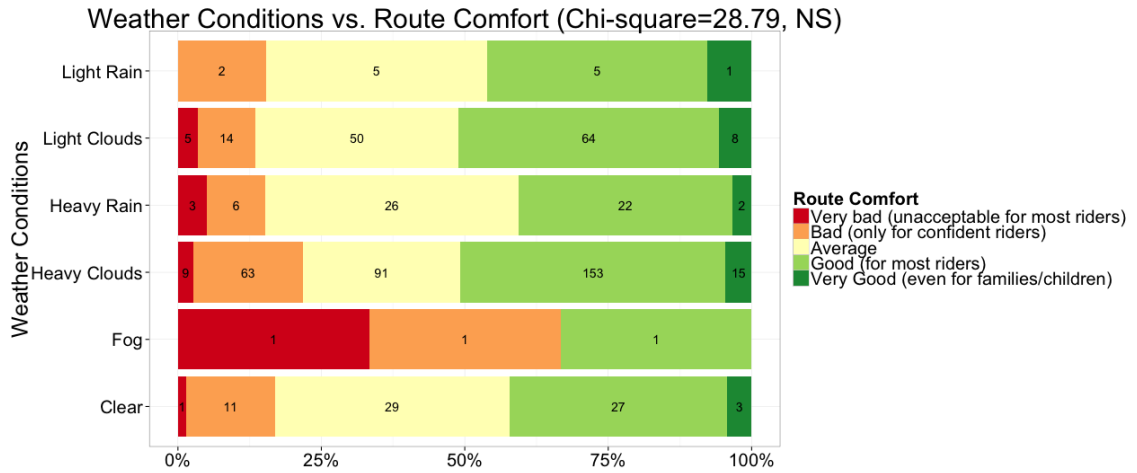
Figure 98: Route Comfort distribution among Weather Conditions

A.3   EXPLORATORY MODELS

A.3.1 Trip Attributes

Several cumulative logistic regression models were tested, with the results presented in
Table 71. In the first model, trip length is tested on its own, resulting in a statistically
significant relationship (β=-0.055, OR=0.95, $p<0.05$) that lowered the route comfort
rating as trip length increased. The second model tested trip duration alone, resulting in a
statistically significant relationship (β=-0.009, OR=0.99, $p<0.1$) that lowered the route
comfort rating as trip duration increased. In the third model, average speed was tested
alone, resulting in a statistically significant relationship (β=-0.092, OR=0.91, $p<0.05$) that
decreased route comfort as average speed increased. In the fourth model, all three
variables were tested, which resulted in no statistically significant relationships.

**Table 71: Cumulative Logit model specification (DV = Route Comfort, IV = Trip
Attribute Variables)**

	Dependent variable:			
	Route Comfort Rating			
	(1)	(2)	(3)	(4)
Trip Length (miles)	-0.055**			-0.014
	(0.024)			(0.128)
Trip Duration (minutes)		-0.009*		-0.005
		(0.005)		(0.025)
Average Speed (mph)			-0.092**	-0.076
			(0.037)	(0.067)

Observations	616	616	616	616
Log Likelihood	-778.782	-779.825	-778.519	-777.455

Note: *p<0.1; **p<0.05; ***p<0.01

A.3.2 Temporal Characteristics

Several cumulative logistic regression models were tested for the temporal

characteristics; the results are presented in

Table 72. The influence of trips taking place on a weekday (as opposed to a weekend) was tested in the first model, resulting a statistically significant (β=-0.428, OR=0.65, p<0.05) negative influence. This indicates that trips taken within this sample during the week were, on average, less comfortable than trips taken on weekends. The influence of trip start time was tested in the second model, with no statistically significant results.

In the third model, day-of-week category and trip start time were tested simultaneously, with the two variable sets interacted. This formulation resulted in mostly statistically significant relationships. When controlling for start time, trips taking place on a weekday had a larger negative influence (β=-0.963, OR=0.38, p<0.01) on route comfort than in the model where day-of-week was considered alone. After controlling for day-of-week, trip start time also became statistically significant in two of three categories. Trips starting during the AM peak were less comfortable (β=-2.000, OR=0.14, p<0.05) than trips starting during the daytime off-peak period. Trips starting during the nighttime off-peak were less comfortable (β=-1.500, OR=0.22, p<0.01) than trips starting during the daytime off-peak period. The odds ratios for each statistically significant coefficient in the third model are presented graphically in Figure 99.
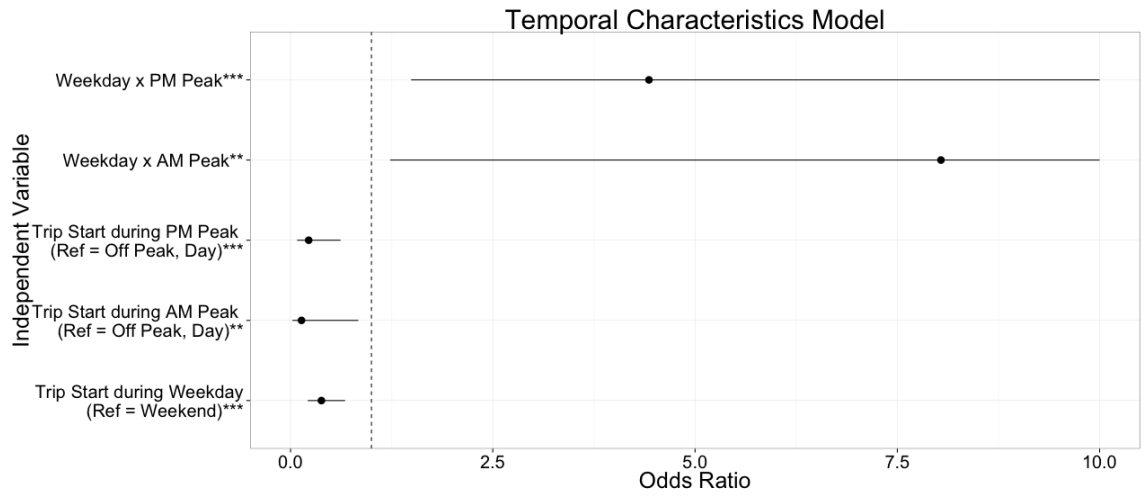
Figure 99: Forest plot of Route Comfort Odds Ratios of coefficients for temporal characteristics model 3 (whiskers correspond to 95% CI)

Table 72: Cumulative Logit model specification (DV = Route Comfort, IV = Temporal Characteristics)

	Dependent variable:		
	Route Comfort Rating		
	(1)	(2)	(3)
Trip took place on weekday	-0.428**		-0.963***
	(0.212)		(0.294)
Trip started between 7:00 AM-9:00 AM (AM Peak)		-0.189	-2.000**
		(0.215)	(0.900)
Trip started between 6:30 PM-7:00 AM (Off-Peak Night)		-0.285	-1.500***
		(0.185)	(0.517)
Trip started between 4:30 PM – 6:30 PM (PM Peak)		-0.143	-0.825
		(0.222)	(0.543)
(Weekday * AM Peak)			2.080**
			(0.929)
(Weekday * PM Peak)			1.490***
			(0.556)
(Weekday * Off-Peak Night)			0.901
			(0.596)
Observations	616	616	616
Log Likelihood	-779.000	-780.000	-773.000

Note: *p<0.1; **p<0.05; ***p<0.01

A.3.3 Trip Question Responses

Several cumulative logistic regression models were tested using the trip question responses as independent variables, with the results presented in

Figure 100: Forest plot of Route Comfort Odds Ratios of coefficients for trip question responses model 6 (whiskers correspond to 95% CI)

Table 73. Trip purpose was tested in the first model, with the reference case being commute trips. Shopping/errand trips were significantly (β=0.674, OR=1.96, p<0.01) more comfortable than commute trips. Work-related trips were significantly (β=-0.772, OR=0.46, p<0.05) less comfortable than commute trips.

Route frequency was tested in the second model, with no statistically significant results. Trip purpose and route frequency were tested simultaneously in the third model, making some of the coefficients more statistically significant. Exercise trips were more comfortable than commute trips, shopping/errand trips were more comfortable than commute trips, and route comfort increased as route frequency increased.

Route choice preferences (i.e. self-reported reasons why a user traveled on their particular route) were tested in the fourth model. Users who indicated they chose their route because it was direct or fast rated their trips as less comfortable (β=-2.940, OR=0.05, p<0.01). Users who indicated that they chose their route because it had good bicycle facilities also rated their trips as less comfortable (β=-0.994, OR=0.37, p<0.01). Users who indicated choosing their route because of enjoyable or nice scenery also rated their trips as less comfortable (β=-0.924, OR=0.40, p<0.05). Users who indicated choosing their route because it was "good for a workout" also rated their trips as less comfortable (β=-1.920, OR=0.15, p<0.05). Users who indicated choosing their route because it had few hills also rated their trips as less comfortable (β=-0.929, OR=0.40, p<0.01). Users who indicated choosing their routes because it was "good for families/kids" rated their trips as more comfortable (β=2.01, OR=7.46, p<0.01). Users who indicated choosing their routes because they did not know another route rated their trips as less comfortable (β=-2.180, OR=0.11, p<0.01). Users who indicated choosing their routes for some other reason not available also rated their trips as less comfortable (β=-0.990, OR=0.37, p<0.05).

Route stressors (i.e. self-reported characteristics of the chosen route that made users uncomfortable or stressed) were tested in the fifth model. Users who indicated they were not concerned about traffic stressors on their route rated their routes as more comfortable (β=2.000, OR=7.39, p<0.01). Users who indicated they were concerned about large commercial vehicles on their route rated that route as less comfortable (β=-1.930, OR=0.15, p<0.01). Users who indicated they were concerned about public transit on their

route rated that route as less comfortable (β=-2.610, OR=0.07, p<0.01). Users who indicated they were concerned about parked vehicles along their route also rated their route as less comfortable (β=-1.040, OR=0.35, p<0.01). Finally, users who indicated they were concerned about pedestrians along their route also rated their route as less comfortable (β=-0.638, OR=0.52, p<0.01).

The final model tested all the previously tested variables together. Overall trends were the same, though some coefficients and statistical significances changed by small amounts. The odds ratios for each statistically significant coefficient in the third model are presented graphically in Figure 100.
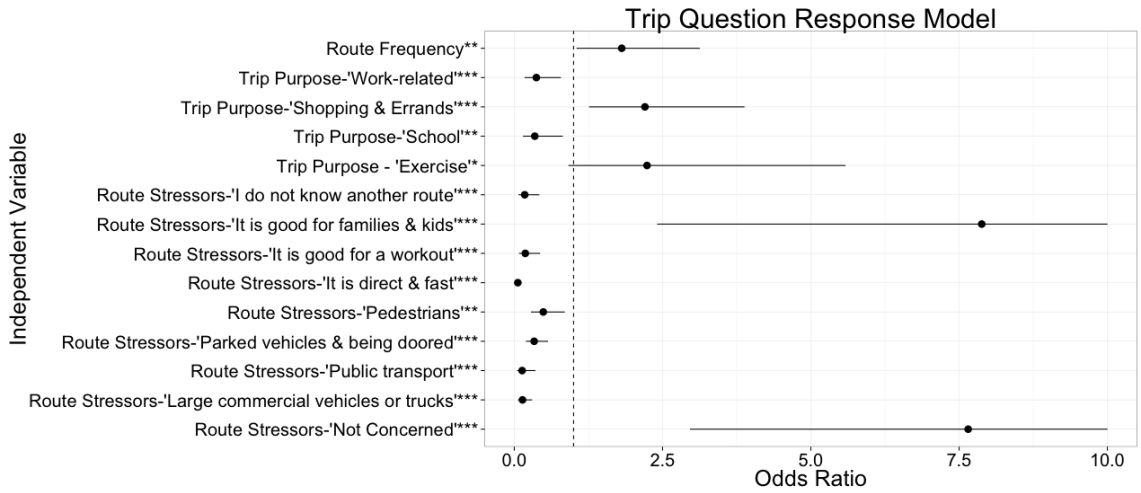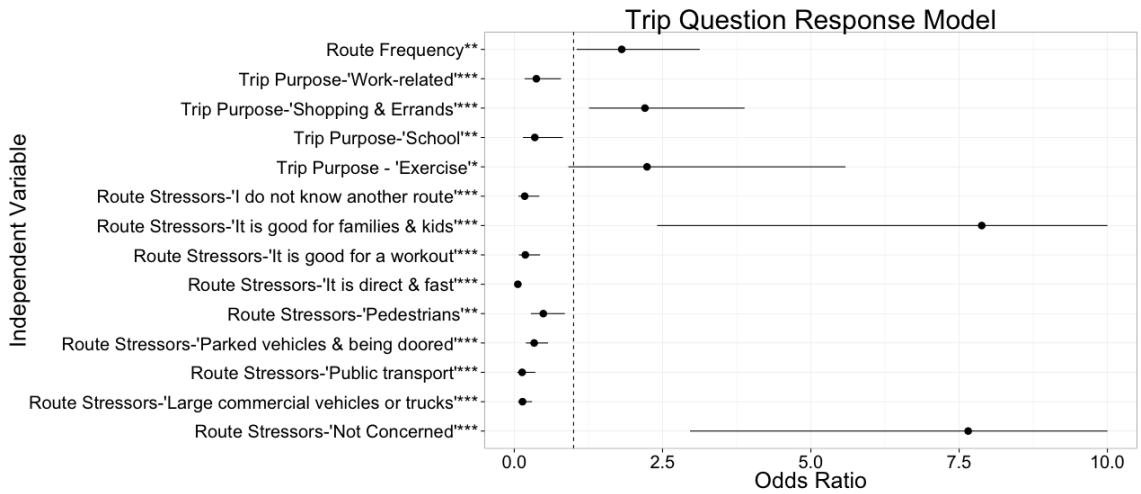

Figure 100: Forest plot of Route Comfort Odds Ratios of coefficients for trip question responses model 6 (whiskers correspond to 95% CI)

Table 73: Cumulative Logit model specification (DV = Route Comfort, IV = Trip Question Responses)

	Dependent variable:					
	routeComfort					
	(1)	(2)	(3)	(4)	(5)	(6)
Trip Purpose - Exercise	0.267		0.688*			0.805*
	(0.373)		(0.413)			(0.461)
Trip Purpose - Other	-0.345		-0.036			-0.519
	(0.401)		(0.425)			(0.488)
Trip Purpose - School	-0.604		-0.528			-1.060**
	(0.381)		(0.382)			(0.436)
Trip Purpose - Shopping/errands	0.674***		0.995***			0.790***
	(0.228)		(0.270)			(0.287)
Trip Purpose - Social/entertainment	0.268		0.523*			0.241
	(0.264)		(0.294)			(0.318)
Trip Purpose - Transportation Access	-1.060		-1.160			1.360
	(1.200)		(1.210)			(1.350)
Trip Purpose -Work-related	-0.772**		-0.587			-0.989***
	(0.346)		(0.363)			(0.382)
Route Frequency		0.153	0.471*			0.594**
		(0.226)	(0.253)			(0.278)
Route Preferences – "It is direct/fast"				-2.940***		-2.820***
				(0.308)		(0.360)
Route Preferences "It has good bicycle facilities"				-0.994***		-0.510
				(0.321)		(0.380)
Route Preferences – "It is enjoyable/has nice scenery"				-0.924**		-0.596
				(0.456)		(0.496)
Route Preferences- "It is good for a workout"				-1.920***		-1.680***
				(0.362)		(0.437)

Route Preferences – "It has low traffic/low speeds"	-0.464	-0.216
	(0.339)	(0.388)
Route Preferences – "It has few busy intersections"	-0.033	0.145
	(0.310)	(0.357)
Route Preferences – "It has few/easy hills"	-0.929***	-0.549
	(0.355)	(0.409)
Route Preferences "It has other riders/people"	-0.283	0.075
	(0.291)	(0.358)
Route Preferences – "It is good for families/kids"	2.010***	2.060***
	(0.527)	(0.611)
Route Preferences – "I do not know another route"	-2.180***	-1.740***
	(0.404)	(0.450)
Route Preferences – "I found online or using my phone"	0.148	0.788
	(0.478)	(0.566)
Route Preferences – "Other"	-0.990**	-0.490
	(0.458)	(0.512)
Route Stressors – "Not Concerned"	2.000***	2.040***
	(0.461)	(0.487)
Route Stressors – "Auto Traffic"	-0.264	-0.081
	(0.268)	(0.318)
Route Stressors – "Large commercial vehicles/trucks"	-1.930***	-1.980***
	(0.349)	(0.398)
Route Stressors – "Public transports (buses, light rail, streetcar, etc.)	-2.610***	-2.020***
	(0.440)	(0.508)

Route Stressors – "Parked vehicles/being doored"					-1.040*** (0.220)	-1.090*** (0.272)
Route Stressors – "Other cyclists"					-0.407 (0.350)	-0.490 (0.401)
Route Stressors – "Pedestrians"					-0.638*** (0.229)	-0.715** (0.284)
Route Stressors – "Other"					-0.330 (0.519)	-0.201 (0.570)
Observations	616	616	616	616	616	616
Log Likelihood	-771.000	-780.000	-767.000	-692.000	-731.000	-634.000

Note: *p<0.1; **p<0.05; ***p<0.01

A.3.4 User Question Responses

Several cumulative logistic regression models were tested using the user question responses, with the results presented in Table 74. The first model tested the variables corresponding to user demographics; namely age, gender, and ethnicity. Route comfort increase as age increased (β=1.840, OR=6.29, p<0.1). Users identifying as "Other" genders were more likely to indicate routes were uncomfortable than males (β=-2.060, OR=0.13, p<0.01). Hispanic users (β=0.896, OR=2.45, p<0.1) and Native American users (β=2.170, OR=8.76, p<0.01) were more likely to rate routes comfortably than White users.

The second model tested variables corresponding to socioeconomic status; namely income, vehicle per worker ratio, and occupation. Neither income or vehicles/workers ratio were statistically significant. However, users reporting "Other" (β=-2.160, OR=0.12, p<0.01) or "Student" (β=-1.110, OR=0.33, p<0.01) occupations were less likely to rate routes comfortably than employed users. Users reporting "Retired" occupations were more likely to rate routes comfortably (β=2.020, OR=7.54, p<0.1).

The third model tested variables associated with bicycling attitudes; namely "preferred cycling weather", "rider type", "rider ability", and the number of bicycles owned. Preferred cycling weather did not produce any statistically significant coefficients. Users identifying with rider types "for recreation and exercise" (β=1.350, OR=3.86, p<0.05) and "to and from work" (β=0.788, OR=2.20, p<0.01) were more likely to rate routes as more comfortable than users who indicated they bicycled for nearly all their trips. Rider

ability and number of bicycles owned did not result in statistically significant coefficients.

In the fourth model, the most significant variable from each of the above models was included; namely age, occupation, and rider type. When controlling for occupation and rider type, age was no longer significant. Users with "Other" occupations were less likely to rate routes as more comfortable than employed users. When controlling for age and occupation only riders identifying as "to and from work" rider types were significantly different from those riding for all trips, with those riders rating routes more comfortably.

Table 74: Cumulative Logit model specification (DV = Route Comfort, IV = User Question Responses)

	Dependent variable:			
	Route Comfort Rating			
	(1)	(2)	(3)	(4)
Age (Ordinal)	1.840[*]			1.350
	(1.010)			(1.120)
Gender (Reference = Male)				
Female	-0.061			
	(0.280)			
Other	-2.060[***]			
	(0.721)			
Ethnicity (Reference = White)				
Asian American	-0.683			
	(0.753)			
Hispanic	0.896[*]			
	(0.530)			
Native American	2.170[***]			
	(0.590)			
Other	0.845			
	(0.593)			
Income (Ordinal)		-0.408		
		(0.396)		
Vehicles/Workers Ratio		0.318		
		(0.275)		
Occupation (Reference = Employed)				
Homemaker		-1.130		-0.187
		(1.150)		(1.380)
Other		-2.160[***]		-1.630[***]
		(0.355)		(0.345)
Retired		2.020[*]		0.711

209

			(1.050)	(1.220)
Student			-1.110***	-0.457
			(0.428)	(0.399)
Preferred Cycling Weather (Reference = In any kind of weather)				
Usually warm and dry weather			1.780	
			(1.200)	
When it does not rain			0.544	
			(0.381)	
Rider Type (Reference = For nearly all my trips)				
For recreation and/or exercise			1.350**	0.543
			(0.592)	(0.650)
For shopping, errands, or visting friends			0.662	0.693
			(1.240)	(1.250)
Mainly to and from work, but occasionally for exercise, shopping, etc.			0.318	0.141
			(0.215)	(0.215)
Other			0.087	0.028
			(0.674)	(0.836)
To and from work			0.788***	0.359*
			(0.204)	(0.213)
Rider Ability (Ordinal)			0.535	
			(1.510)	
Number of Bicycles (Ordinal)			0.080	
			(0.169)	
Observations	616	616	616	616
Log Likelihood	-727.000	-734.000	-742.000	-723.000

Note: *p<0.1; **p<0.05; ***p<0.01

A.3.5 Bicycle Facility and Street Type

Cumulative logistic regression models were tested for the relationship between route comfort and the miles of a trip ridden on different bicycle facility types. The results are presented in

Table 75. In the first model, only the mileages on the different bicycle facility types were tested. Trip miles on links typed as "no bike facility, primary arterial" detracted from route comfort (β=-0.548, OR=0.58, p<0.01). Trip miles on links typed as "no bike facility, other" also detracted from route comfort (β=-0.402, OR=0.67, p<0.05). Trip miles on links typed as "bike lane, primary arterial"" also detracted from route comfort (β=-0.152, OR=0.98, p<0.1). Trip miles on links typed as "bike lane, minor arterial" also detracted from route comfort (β=-0.555, OR=0.57, p<0.01). Finally, trip miles on links typed as separated paths increased route comfort (β=0.341, OR=1.41, p<0.01).

In the second model, trip length is controlled for in addition to all the variables included in the first model. Trip miles ridden on links typed as "bike lane, primary arterial" no longer had a statistically significant contribution to route comfort. The other coefficients changed slightly, but the overall trends remained the same. The odds ratios for each statistically significant coefficient in the second model are presented graphically in Figure 101.
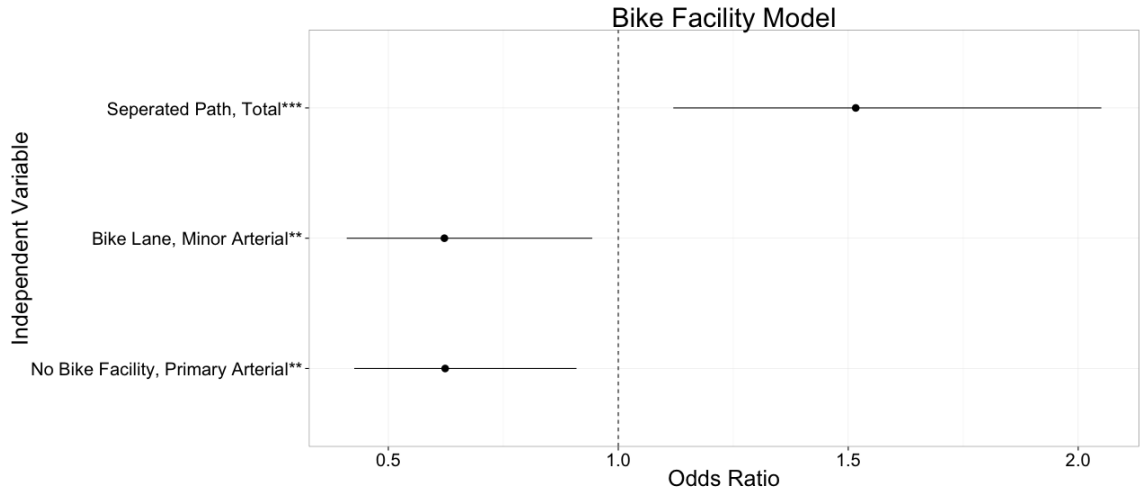
Figure 101: Forest plot of Route Comfort Odds Ratios of statistically significant coefficients for bike facility model 2 (whiskers correspond to 95% CI)

Table 75: Cumulative Logit model specification (DV = Route Comfort, IV =Trip miles on different bike facility types)

	Dependent variable:	
	Route Comfort Rating	
	(1)	(2)
Trip miles on link type "No Bike Facility, Primary Arterial"	-0.548[***]	-0.472[**]
	(0.135)	(0.193)
Trip miles on link type "No Bike Facility, Minor Arterial"	0.183	0.250
	(0.246)	(0.275)
Trip miles on link type "No Bike Facility, Residential Street"	-0.004	0.071
	(0.059)	(0.150)
Trip miles on link type "No Bike Facility, Other"	-0.402[**]	-0.330
	(0.178)	(0.221)
Trip miles on link type "Bike Lane, Primary Arterial"	-0.152[*]	-0.078
	(0.079)	(0.157)
Trip miles on link type "Bike Lane, Minor Arterial"	-0.555[***]	-0.475[**]
	(0.154)	(0.212)

212

Trip miles on link type "Bike Lane, Residential Street"	-0.050	0.028
	(0.086)	(0.167)
Trip miles on link type "Bike Lane, Other"	0.273	0.339
	(0.433)	(0.450)
Trip miles on link type "Cycletrack or Buffered Bike Lane, Total"	0.308	0.408
	(0.895)	(0.914)
Trip miles on link type "Bicycle Boulevard, Total"	0.119	0.200
	(0.089)	(0.172)
Trip miles on link type "Separated Path, Total"	0.341***	0.416***
	(0.070)	(0.154)
Trip Length		-0.076
		(0.139)
Observations	616	616
Log Likelihood	-741.000	-741.000

Note: *p<0.1; **p<0.05; ***p<0.01

A.3.6 Topography

Two cumulative logistic regression models were tested for the comparison of route comfort and the slope of trip segments; the results are presented in

Table 76. In both models, the number of miles of a trip ridden on a particular slope category is the independent variable over which route comfort is regressed. In the first model, the number of miles ridden on each slope category is tested. As the number of miles ridden on grades less than -2% increased, a trip was rated less comfortably (β=-0.433, OR=0.65,p<0.01). As the number of miles ridden on grades greater than +6% increased, a trip was rated less comfortably with a higher effect size (β=-2.730, OR=0.07,p<0.01). As the number of miles ridden on grades between +2% and +4% increased, the comfort rating increased (β=0.442, OR=1.56, p<0.1). The second model contains all the variables of the first, but controls for overall trip length. Coefficients changed marginally, but the same overall trends are observed. The odds ratios for each statistically significant coefficient in the second model are presented graphically in Figure 102.
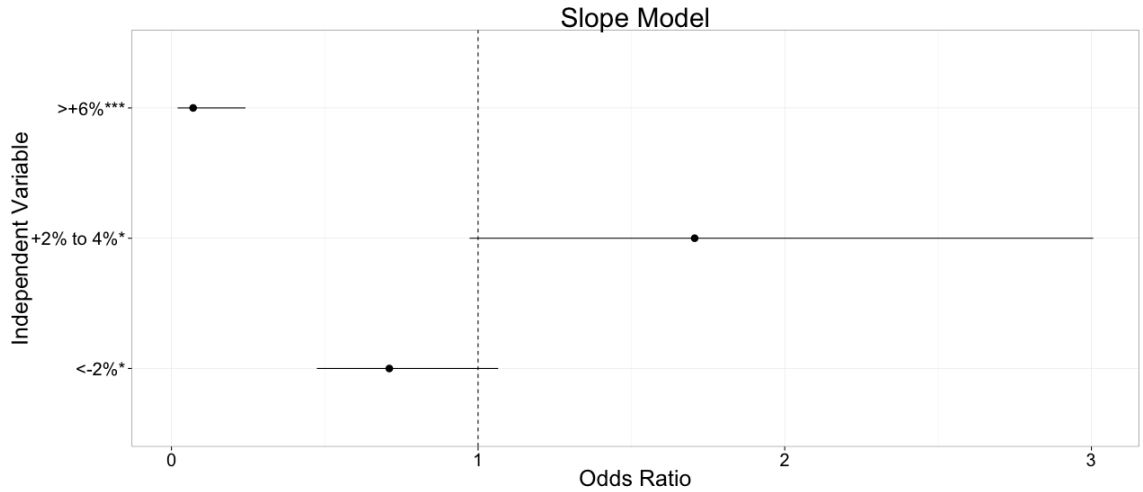
Figure 102: Forest plot of Route Comfort Odds Ratios of statistically significant coefficients for slope model 2 (whiskers correspond to 95% CI)

Table 76: Cumulative Logit model specification (DV = Route Comfort, IV = Slope categories)

	Dependent variable:	
	Route Comfort Rating	
	(1)	(2)
Trip miles on <-2% grade	-0.433***	-0.342*
	(0.157)	(0.206)
Trip miles on -2% to +2% grade	0.040	0.132
	(0.033)	(0.138)
Trip miles on +2% to +4% grade	0.442*	0.534*
	(0.255)	(0.288)
Trip miles on +4% to +6% grade	0.055	0.130
	(0.271)	(0.292)
Trip miles on >+6% grade	-2.730***	-2.650***
	(0.625)	(0.636)

215

Trip length (miles)	-0.091	
	(0.132)	
Observations	616	616
Log Likelihood	-754.000	-754.000

A.3.7 Daily Traffic Volume

Cumulative logistic regression models were tested for the relationship between route comfort and the number of miles ridden on network links with various daily traffic volume categories. Results of the models are presented in

Table 77 In the first model, the mileage per trip on the different traffic categories was tested. Comfort increased as more miles of a trip were ridden on links with traffic volume less than 5,000 vehicles per day (β=0.167, OR=1.18, p<0.01). Comfort decreased with greater effect size sequentially as mileage on higher traffic volume links increased, with comfort decreasing the most per mile on links with more than 30,000 vehicles/day (β=-0.864, OR=0.42, p<0.01). The second model included all the independent variables of the first but also controls for overall trip length. Controlling for trip length changed coefficients and significances slightly, but the same overall trend can be observed. The odds ratios for each statistically significant coefficient in the second model are presented graphically in Figure 103.
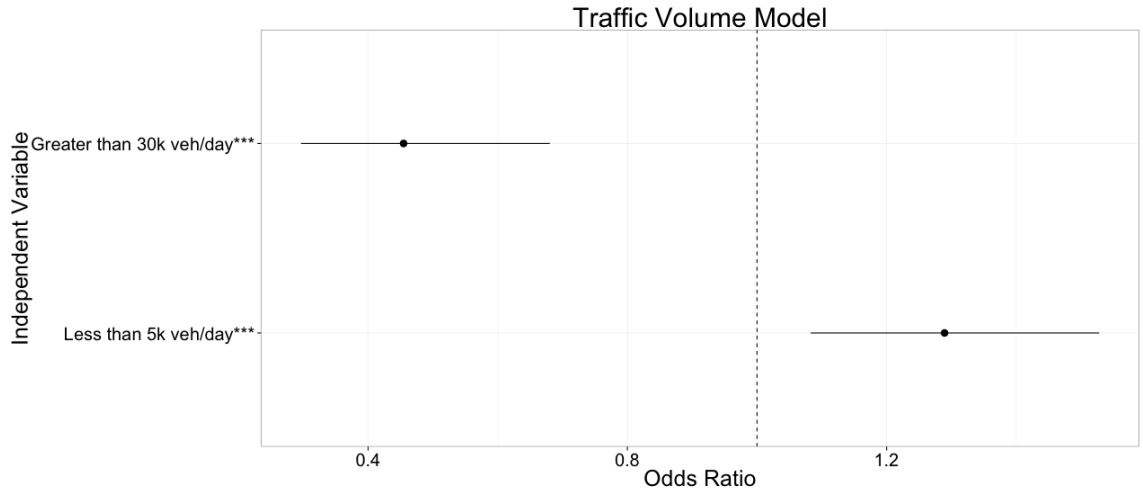
Figure 103: Forest plot of Route Comfort Odds Ratios of statistically significant coefficients for traffic model 2 (whiskers correspond to 95% CI)

Table 77: Cumulative Logit model specification (DV = Route Comfort, IV = Traffic volume categories)

	Dependent variable:	
	Route Comfort Rating	
	(1)	(2)
Miles ridden on links with traffic volume less than 5,000 vehicles/day	0.167[***]	0.254[***]
	(0.039)	(0.086)
Miles ridden on links with traffic volume between 5,000 and 10,000 vehicles/day	-0.112	-0.017
	(0.095)	(0.127)
Miles ridden on links with traffic volume between 10,000 and 20,000 vehicles/day	-0.177[*]	-0.087
	(0.105)	(0.131)
Miles ridden on links with traffic volume between 20,000 and 30,000 vehicles/day	-0.231[**]	-0.138
	(0.091)	(0.123)
Miles ridden on links with traffic volume greater	-0.864[***]	-0.788[***]

218

than 30,000 vehicles/day

	(0.199)	(0.210)
Trip Length (miles)		-0.092
		(0.082)
Observations	613	613
Log Likelihood	-743.000	-742.000

Note: *p<0.1; **p<0.05; ***p<0.01

A.3.8 Traffic Speed

Cumulative logistic regression models were tested for the relationship between route

comfort and the number of miles ridden on network links with various posted traffic

speed categories. In the first model, trip mileage on three different categories of traffic

speeds was tested. As trip mileage on links with speeds less than or equal to 20 mph

increased, users were more likely to rate routes as comfortable (β=-0.248, OR=1.28,

p<0.01). As trip mileage on links with speeds greater than or equal to 35 mph increased,

users were more likely to rate routes as less comfortable (β=-0.353, OR=0.70, p<0.01). In

the second model, trip length is controlled for. The coefficients change slightly as a

result, but the statistical significances and interpretations remain the same. The odds

ratios for each coefficient in the second model are presented graphically in Figure 104.

Figure 104: Forest plot of Route Comfort Odds Ratios of coefficients for speed model 2 (whiskers correspond to 95% CI)

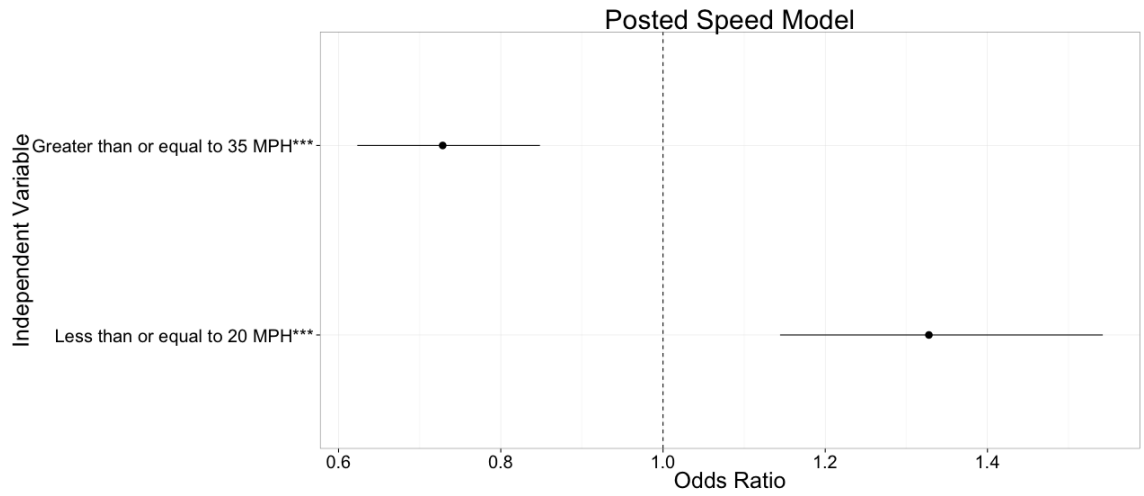**Table 78: Cumulative Logit model specification (DV = Route Comfort, IV = Posted traffic speed categories)**

	Dependent variable:	
	Route Comfort	
	(1)	(2)
Trip miles on links with posted traffic speeds less than or equal to 20 mph	0.248***	0.284***
	(0.061)	(0.076)
Trip miles on links with posted traffic speeds between 20 mph and 35 mph	-0.023	0.013
	(0.036)	(0.058)
Trip miles on links with posted traffic speed greater than or equal to 35 mph	-0.353***	-0.317***
	(0.063)	(0.078)
Trip Length (miles)		-0.039
		(0.050)
Observations	616	616
Log Likelihood	-755.000	-755.000
Note:	*p<0.1; **p<0.05; ***p<0.01	

A.3.9 Weather Variables

Cumulative logistic regression models were tested for several weather variables. None of the models had statistically significant variables except for the last one tested, which contained all the variables. In the last model, Fog had a statistically significant negative influence on route comfort (β=-2.12, OR=0.12, p<0.1).

Table 79: Cumulative Logit model specification (DV = Route Comfort, IV = Weather variables)

	Dependent variable:					
	routeComfort					
	(1)	(2)	(3)	(4)	(5)	(6)
Temperature (Deg F)	0.010					0.006
	(0.010)					(0.010)
Wind Speed (mph)		-0.013				-0.021
		(0.009)				(0.016)
Wind Gust Speed (mph)			-0.005			0.005
			(0.006)			(0.010)
Precipitation				5.510		11.900
				(7.540)		(8.790)
Weather Conditions (Reference Category = Clear)						
Fog					-2.040	-2.120[*]
					(1.260)	(1.270)
Heavy Clouds					0.144	0.081
					(0.240)	(0.248)
Heavy Rain					-0.035	-0.256
					(0.318)	(0.363)
Light Clouds					0.301	0.311
					(0.267)	(0.270)
Light Rain					0.198	0.095
					(0.553)	(0.562)
Observations	616	616	616	616	616	616
Log Likelihood	-781.000	-780.000	-781.000	-781.000	-779.000	-776.000

Note: [*]$p<0.1$; [**]$p<0.05$; [***]$p<0.01$

223

A.4 NON-IMPUTED MODEL SPECIFICATIONS

In the following model specifications, route comfort (the dependent variable) was not imputed. The independent variables were imputed. Backwards stepwise specifications were used for each single model group, and the same model specification as used in section 5.3 was used for the pooled model. Less observations were used in each model because of the loss of imputed data. In general, signs remained the same and were intuitive.

A.4.1 Trip Attributes

Table 80: Cumulative logistic regression model specification for trip attributes

	Dependent variable:
	Route Comfort Rating
avgLinkSpeed	-0.088** (0.043)
Observations	431
Log Likelihood	-573.882
Note:	*p<0.1; **p<0.05; ***p<0.01

A.4.2 Temporal Characteristics

Table 81: Cumulative logistic regression model specification for temporal characteristics

	Dependent variable:
	Route Comfort Rating
Trip Start during Weekday (Ref = Weekend)	-0.473[*] (0.246)
Observations	431
Log Likelihood	-574.085
Note:	[*]p<0.1; [**]p<0.05; [***]p<0.01

A.4.3 Trip Question Responses

Table 82: Cumulative logistic regression model specification for trip question responses

	Dependent variable:
	Route Comfort Rating
routeStressors_Not.concerned	2.142*** (0.474)
routeStressors_Large.commercial.vehicles..trucks.	-1.695*** (0.375)
routeStressors_Public.transport..buses..light.rail..streetcar.	-1.641*** (0.469)
routeStressors_Parked.vehicles..being.doored.	-0.807*** (0.246)
routeStressors_Pedestrians	-0.511** (0.258)
routePrefs_It.is.direct.fast	-0.824** (0.341)
routePrefs_It.has.good.bicycle.facilities	1.232*** (0.389)
routePrefs_It.is.enjoyable.has.nice.scenery	0.800 (0.528)
routePrefs_It.has.low.traffic.low.speeds	1.240*** (0.422)
routePrefs_It.has.few.busy.intersections	1.512*** (0.374)
routePrefs_It.has.few.easy.hills	1.038*** (0.397)
routePrefs_It.has.other.riders.people	1.657*** (0.344)
routePrefs_It.is.good.for.families.kids	3.423*** (0.583)
routePrefs_I.found.it.online.or.using.my.phone	2.130*** (0.532)
routePrefs_Other	1.034** (0.513)
Observations	431
Log Likelihood	-486.412
Note:	*p<0.1; **p<0.05; ***p<0.01

A.4.4 User Question Responses

Table 83: Cumulative logistic regression model specification for user question responses

	Dependent variable:
	Route Comfort Rating
Ethnicity: White	-0.760** (0.316)
Occupation: Employed	0.864*** (0.266)
Observations	431
Log Likelihood	-567.882
Note:	*p<0.1; **p<0.05; ***p<0.01

A.4.5 Bicycle Facility and Street Type

Table 84: Cumulative logistic regression model specification for biycle facility and street typology

	Dependent variable:
	Route Comfort Rating
Trip Miles on 'No Bike Facility, Primary Arterial'	-0.628*** (0.127)
Trip Miles on 'No Bike Facility, Other'	-0.469** (0.198)
Trip Miles on 'Bike Lane, Primary Arterial'	-0.871*** (0.161)
Trip Miles on 'Bike Lane, Minor Arterial'	0.820* (0.483)
Trip Miles on 'Bicycle Boulevard'	0.188* (0.104)
Trip Miles on 'Seperated Path'	0.445*** (0.084)
Observations	431
Log Likelihood	-529.612
Note:	*p<0.1; **p<0.05; ***p<0.01

A.4.6 Topography

Table 85: Cumulative logistic regression model specification for topography

	Dependent variable:
	Route Comfort Rating
Trip miles on grades <-2%	-0.573*** (0.164)
Trip miles on grades +2% to 4%	0.086** (0.035)
Trip miles on grades >+6%	-2.656*** (0.686)
Observations	431
Log Likelihood	-545.798
Note:	*p<0.1; **p<0.05; ***p<0.01

A.4.7 Daily Traffic Volume

Table 86: Cumulative logistic regression model specification for traffic volume

	Dependent variable:
	Route Comfort Rating
Trip miles on links with 'Less than 5k veh/day'	0.325^{***} (0.076)
Trip miles on links with '5k - 10k veh/day'	-0.373^{***} (0.137)
Trip miles on links with 'Greater than 30k veh/day'	-0.504^{**} (0.216)
Trip length (miles)	-0.114^{*} (0.058)
Observations	430
Log Likelihood	-537.029
Note:	*p<0.1; **p<0.05; ***p<0.01

A.4.8 Traffic Speed

Table 87: Cumulative logistic regression model specification for traffic speed

	Dependent variable:
	routeComfort
Trip miles on links with posted speed <= 20 MPH	0.324*** (0.071)
Trip miles on links with posted speed >35 MPH	-0.362*** (0.066)
Observations	431
Log Likelihood	-550.207
Note:	*p<0.1; **p<0.05; ***p<0.01

A.4.9 Pooled Model

Table 88: Cumulative logistic regression model specification for pooled model

	Dependent variable:
	Route Comfort Rating
Route Frequency (Ordinal)	0.669** (0.327)
Trip Purpose: Exercise	2.377*** (0.582)
Trip Purpose: Shopping/Errands	0.841*** (0.317)
Occupation: Employed	1.208*** (0.312)
Ethnicity: White	-0.472 (0.353)
Household Vehicles (Ordinal)	2.217*** (0.639)
Route Preferences: It is direct or fast	-1.915*** (0.307)
Route preferences: It is good for a workout	-1.297*** (0.377)
Route Preferences: It has few busy intersections	0.487 (0.332)
Route Preferences: It is good for families/kids	2.103*** (0.571)
Route Stressors: Not concerned	1.902*** (0.499)
Route Stressors: Large commercial vehicles/trucks	-1.618*** (0.414)
Route Stressors: Parked vehicles/being doored	-0.636*** (0.233)
Trip miles on grades >+6%	-3.302*** (0.670)
Trip miles on links with <5k veh/day	0.296*** (0.068)
Trip miles on links with posted speed >35 mph	-0.230*** (0.070)
Trip miles on links with posted speed 20-35 mph	-0.197*** (0.064)
Observations	420
Log Likelihood	-420.735
Note:	*p<0.1; **p<0.05; ***p<0.01