UNDERSTANDING BIKE SHARE CYCLIST ROUTE CHOICE BEHAVIOR

UNDERSTANDING BIKE SHARE CYCLIST ROUTE CHOICE BEHAVIOR

BY

WEI LU, B.Env.Sc. (HONS.)

A THESIS

SUBMITTED TO THE SCHOOL OF GEOGRAPHY AND EARTH SCIENCES

AND THE SCHOOL OF GRADUATE STUDIES

OF MCMASTER UNIVERSITY

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF SCIENCE

Master of Science (2017) McMaster University

(School of Geography and Earth Sciences) Hamilton, Ontario, Canada

TITLE:	UNDERSTANDING BIKE SHARE CYCLIST ROUTE CHOICE BEHAVIOR
AUTHOR:	Wei Lu B.Env.Sc. (HONS.) University of Waterloo, Waterloo, Canada
SUPERVISOR:	Dr. Darren M. Scott
NUMBER OF PAGES:	x, 74

Abstract

This thesis examines the existence of a dominant route between a hub pair and factors that influence bike share cyclists route choices. This research collects 132,396 hub-to-hub global positioning system (GPS) trajectories over a 12-month period between April 1, 2015 and March 31, 2016 from 750 bicycles provided by a bike share program (BSP) called SoBi (Social Bicycles) Hamilton. Then, a GIS-based map-matching toolkit is used to convert GPS points to map-matched trips and generate a series of route attributes. In order to create choice sets, unique routes between the same hub pair are extracted from all corresponding repeated trips using a link signature tool. The results from t statistics and Path-size logit models indicate that bike share cyclists are willing to detour for some positive features, such as bicycle facilities and low traffic volumes, but they also try to avoid too circuitous routes, turns, and steep slopes over 4% though detouring may come with a slight increase in turns. This research not only helps us understand BSP cyclist route preferences but also presents a GIS-based approach to determine potential road segments for additional bike facilities on the basis of such preferences.

Acknowledgements

Financial support for this thesis was provided by a grant from the Natural Sciences and Engineering Research Council of Canada, which was awarded to Dr. Darren M. Scott (PI).

The completion of this thesis would not have been possible without the contributions of several individuals. First, I would like to thank Dr. Scott (supervisor), who has provided me with guidance and support, not only in the context of this thesis, but also with many other aspects of my Master career. Thank you to Dr. Ron Dalumpines for his help with Python and a GIS-based map-matching toolkit.

I would also like to acknowledge Peter Topalovic (City of Hamilton) and Chelsea Cox (SoBi Hamilton) for providing all the GPS data for this research. Finally, I would like to thank my TransLAB mates: Charles Burke, Celenna Ciuro, and Justin Hall for supporting and encouraging me during my Master period.

Preface

This thesis is organized as a compendium of related articles. It consists of the following two chapters:

Chapter 2: Understanding bike share cyclist route choice using GPS data: Comparing dominant routes and shortest paths

Chapter 3: Determinants of bike share cyclist route choice behavior using GPS data

Although the two journal articles have been co-authored with the research supervisors and Dr. Ron Dalumpines (the first paper), the thesis author takes responsibility of the content of each chapter, including setting up research objectives, reviewing literatures, processing and analyzing data, coding, specifying and estimating models, and interpreting results. The supervisors contributions include suggesting the topics and methods of this research, discussing all the outcomes, and editing and evaluating the entire papers and prior to journal submission.

Contents

Abstract **iii**

Acknowledgements **iv**

Preface **v**

List of Tables **ix**

List of Figures **x**

1 INTRODUCTION **1**

1.1 THE RESEARCH PROBLEM 1

1.2 RESEARCH OBJECTIVES 3

1.3 THESIS OUTLINE 4

2 UNDERSTANDING BIKE SHARE CYCLIST ROUTE CHOICE USING GPS DATA: COMPARING DOMINANT ROUTES AND SHORTEST PATHS **9**

2.1 INTRODUCTION 9

2.2 BACKGROUND 12

2.2.1 Data Collection Methods 12
2.2.2 Potential Determinants of Cyclist Route Choice Behavior . . . 13
2.3 DATA 16
2.3.1 Study Area 16
2.3.2 Cycling Network 17
2.3.3 GPS Dataset 18
2.4 METHODOLOGY 19
2.4.1 Data Generation 19
2.4.2 Statistical Analysis 21
2.5 RESULTS 23
2.6 CONCLUSIONS 29

3 DETERMINANTS OF BIKE SHARE CYCLIST ROUTE CHOICE BEHAVIOR USING GPS DATA 38

3.1 INTRODUCTION 38
3.2 BACKGROUND 41
3.3 METHODS 43
3.3.1 GPS Data Collection and Post-processing 43
3.3.2 Choice Set Generation 46
3.3.3 Normalized Gini Coefficient 47
3.3.4 Path-Size Logit Model 48
3.4 RESULTS 49
3.4.1 Descriptive Statistics 51
3.4.2 Contribution of Route Attributes 54
3.5 CONCLUSIONS 58

4 CONCLUSIONS **67**

4.1 INTRODUCTION 67

4.2 SUMMARY OF FINDINGS AND RESEARCH CONTRIBUTIONS . 68

4.3 RESEARCH LIMITATIONS 70

4.4 CONCLUSIONS AND FUTURE RESEARCH 71

List of Tables

2.1 GPS data processing summary. 23

2.2 Route attributes of dominant routes compared to corresponding shortest paths (distance). 27

3.1 Descriptive statistics of route attributes for global, medium G_n, and high G_n datasets. 53

3.2 PSL model estimation results for global, medium G_n, and high G_n datasets. 56

List of Figures

2.1 Study area of SoBi Hamilton. 17

2.2 Conversion from GPS trajectories to an observed route using the map-matching toolkit. 20

2.3 Generation of core link signatures for each hub-hub trip. 21

2.4 Cyclist route choices. 24

2.5 Spatial distribution of dominant routes. 26

3.1 Study area and road network of SoBi Hamilton. 44

3.2 Bikeway distribution in study area. 46

3.3 (a) Demonstration of a hub pair with a high G_n. (b) Demonstration of a hub pair with a low G_n. 50

3.4 Route distance pattern for routes from three datasets: global, medium G_n, and high G_n. 51

3.5 Demand pattern of SoBi hubs. 57

3.6 Demonstration of routes between a hub pair mainly along Cannon Street. 58

Chapter 1

INTRODUCTION

1.1 THE RESEARCH PROBLEM

Physical inactivity is a significant element accounting for death, chronic morbidity, and disability; statistically, it contributed to 21.5% of ischemic heart disease, 16% of colon cancer, 14% of diabetes, 11% of ischemic stroke, and 10% of breast cancer all over the world (Bull et al., 2004). Over the past 20 years, adult and childhood obesity rates have doubled, while adolescent obesity has tripled (Ford et al., 2014). According to WHO (2016), the global overweight rate of adults over 18 years old is around 39%, and the overall obesity rate is 13% in 2014. Over the past decade, the obesity rate of 35 countries participating in the Organization for Economic Co-operation and Development (OECD) has increased to 19.5% in 2015, and 25.8% of adults in Canada aged 15 years and over are obese (OECD, 2017). It is essential to engage in physical activity that can benefit not only physical health by reducing the potential of obesity (Flynn et al., 2006), osteoporosis (Biddle et al., 2004), and cardiovascular disease (Andersen et al., 2006), but also mental health to some extent (Biddle et al., 2004; Panter et al., 2008).

Transport accounted for 23% of global energy-related greenhouse gas emissions in 2004, where around three quarters were from road vehicles, and total transport-related emissions were predicted to increase by about 80% from 2007 to 2030 (Kahn et al., 2007). According to Kunzli et al. (2000), air pollution was associated with morbidity, such as respiratory and cardiovascular diseases, bronchitis, and asthma attacks, and accounted for 6% of total mortality every year, where half was caused by automobile emissions. In this case, active travel, primarily including walking and cycling, has been encouraged given its benefits for public health by increasing physical activity (Pucher et al., 2010), reducing the probability of obesity (Saunders et al., 2013) and chronic diseases (Giles-Corti et al., 2010), and for the environment by mitigating air and noise pollution (De Nazelle et al., 2011).

Dill (2009) claimed that it is more likely for cycling than walking to replace motorized traffic because cycling is faster and can cover much longer distances. However, switching from car to bicycle also means a higher risk of traffic accidents and inhalation of air pollution on the road (De Hartog et al., 2010). It has been found that positive effects of cycling, including decreased air pollution emissions and increased physical fitness, were generally larger than its aforementioned risks, especially considering its influences for the whole society instead of individuals (De Hartog et al., 2010). Rojas-Rueda et al. (2011) further demonstrated that benefits of bike share programs (BSPs), such as Bicing in Barcelona, were larger than risks in terms of public health and carbon emissions. According to Fishman (2015), BSPs have been advocated in the past few years as the number of participating cities has increased from just a few in the late 1990s to over 800 in 2014. No matter for cycling or driving, route choice analysis is essential and important because it can reveal commuters perceptions about route attributes and reactions or adaptations to the source of information, understand individual travel behavior, and predict traffic volumes (Prato, 2009).

In order to motivate the use of BSPs, it is necessary to explore bike share cyclists route choices to learn about corresponding spatial distributions and factors that can contribute to their travel behavior, which will help transport planners determine which road segments can be equipped with what types of bicycle facilities.

1.2 RESEARCH OBJECTIVES

The goal of this research is to determine what and how route attributes can affect bike share cyclist route choice behavior using GPS dataset derived from SoBi (Social Bicycle) in Hamilton. To achieve this goal, the following step-by-step objectives are required:

- Gnerate cyclists actual trips from hub to hub using GPS trajectories and extract unique routes to constitute a choice set for each hub pair.
- Determine whether a dominant route exists between a hub pair according to normalized Gini coefficients and explore the spatial distribution of bike share cyclists dominant routes.
- Compare dominant routes with shortest paths based on distance and examine differences.

- Identify determinants that contribute to SoBi cyclist route choice decision making processes in terms of route attributes based on path-size logit models.

Realizing these objectives will contribute to the existing body of literature about cyclist route choice behavior. Firstly, it will fill a major research gap regarding what and how route attributes influence bike share cyclists route choice decisions. These findings will help urban planners and policy makers to better understand the preferences of BSP cyclists and find solutions to promote the use of BSP, especially within the area of Hamilton. Secondly, it will fill a research gap left by Lima et al. (2016) who introduces the use of Gn and finds a tendency towards a preferred route among all the routes for a drivers routine journey. To the authors knowledge, this is the first study to generate attributes for each trip taken by BSP users, to use Gn to determine whether a dominant route exists between a BSP hub pair, and to evaluate the efficiency of dominant routes by comparing with shortest distance routes. In addition, this research proposes a new method to generate choice sets using BSP cyclists actual routes instead of hypothetical routes created by conventional choice set generation techniques. In this case, findings in this study will reveal the actual preferences of cyclists for route attributes because all the evaluated routes are cyclists hub-to-hub choices in the real world.

1.3 THESIS OUTLINE

Including this introduction, this thesis consists of four chapters. Chapters 2 and 3 respectively contain two stand-alone research papers, and Chapter 4 briefly summarizes major findings and contributions.

Chapter 2 determines that the most preferred route between each hub pair can be regarded as a dominant route given the value of Gn and compares it with the corresponding shortest path based on distance in terms of route attributes based on t statistics. The comparison shows that dominant routes are significantly longer than their shortest distance routes, indicating that most bike share cyclists are willing to detour for some routes attributes, such as bicycle facilities and low traffic volumes, which may be accompanied by the increase of some negative features such as turns and intersections.

Chapter 3 examines the determinants of bike share cyclist route choice behavior using modelling Path-size Logit. A new method of generating alternative routes for route choice sets using actual cyclists routes within a BSP is introduced and confirmed to be effective. All the models for datasets with different Gn levels demonstrated that cyclists are willing to detour for some positive features, such as bicycle facilities, but they also try to avoid detouring too much, steep slopes over 4%, and high traffic volumes.

Chapter 4 reviews major findings and contributions of these two research papers, followed by a section discussing the limitations of this research. Finally, this thesis concludes with suggestions for future studies.

REFERENCES

Andersen, L. B., Harro, M., Sardinha, L. B., Froberg, K., Ekelund, U., Brage, S., & Anderssen, S. A. (2006). Physical activity and clustered cardiovascular risk in children: a cross-sectional study (The European Youth Heart Study). *The Lancet*, 368: 299-304.

Biddle, S. J., Gorely, T., & Stensel, D. J. (2004). Health-enhancing physical activity and sedentary behaviour in children and adolescents. *Journal of sports sciences*, 22: 679-701.

Bull, F. C., Armstrong, T. P., DixON, T., Ham, S., NEiMAN, A. N. D. R. E. A., & Pratt, M. (2004). Physical inactivity. *Comparative Quantification of Health Risks Global and Regional Burden of Disease Attributable to Selected Major Risk Factors. Geneva: World Health Organization*, 729-881.

De Hartog, J. J., Boogaard, H., Nijland, H., & Hoek, G. (2010). Do the health benefits of cycling outweigh the risks? *Environmental Health Perspectives*, 118: 1109.

De Nazelle, A., Nieuwenhuijsen, M. J., Ant, J. M., Brauer, M., Briggs, D., Braun-Fahrlander, C., ... & Hoek, G. (2011). Improving health through policies that promote active travel: a review of evidence to support integrated health impact assessment. *Environment International*, 37: 766-777.

Fishman, E. (2016). Bikeshare: A review of recent literature. *Transport Reviews*, 36: 92-113.

Flynn, M. A. T., McNeil, D. A., Maloff, B., Mutasingwa, D., Wu, M., Ford, C., & Tough, S. C. (2006). Reducing obesity and related chronic disease risk in children and youth: a synthesis of evidence with best practicerecommendations. *Obesity Reviews*, 7: 7-66.

Giles-Corti, B., Foster, S., Shilton, T., & Falconer, R. (2010). The co-benefits for health of investing in active transportation. *New South Wales Public Health Bulletin*, 21: 122-127.

Kunzli, N., Kaiser, R., Medina, S., & Studnicka, M. (2000). Public-health impact of outdoor and traffic-related air pollution: a European assessment. *The Lancet*, 356: 795-801.

R., Kobayashi,S. S., Beuthe, M., Gasca, J., Greene, D., Lee, D. S., Muromachi, Y., Newton, P. J., Plotkin, S., Sperling, D., Wit, R., & Zhou, P. J. (2007). ransport and its infrastructure. Metz,B., Davidson,O. R., Bosch,P. R., Dave,R., Meyer, L.A.(Eds.). *Climate Change 2007: Mitigation Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, Cambridge and New York.

OECD. (2017). Obesity Update 2017. Retrieved August 18, 2017 from: http://www.oecd.org/els/health-systems/Obesity-Update-2017.pdf

Panter, J. R., Jones, A. P., & Van Sluijs, E. M. (2008). Environmental determinants of active travel in youth: A review and framework for future research. *International Journal of Behavioral Nutrition and Physical Activity*, 5: 34.

Prato, C. G. (2009). Route choice modeling: past, present and future research directions. *Journal of Choice Modelling*, 2: 65-100.

Pucher, J., Buehler, R., Bassett, D. R., & Dannenberg, A. L. (2010). Walking and cycling to health: a comparative analysis of city, state, and international data. *American Journal of Public Health*, 100: 1986-1992.

Rojas-Rueda, D., de Nazelle, A., Tainio, M., & Nieuwenhuijsen, M. J. (2011). The health risks and benefits of cycling in urban environments compared with car use: health impact assessment study. *British Medicial Journal*, 343: d4521.

Saunders, L. E., Green, J. M., Petticrew, M. P., Steinbach, R., & Roberts, H. (2013). What are the health benefits of active travel? A systematic review of trials and cohort studies. *PLoS One*, 8: e69912.

WHO. (2016). Obesity and overweight.

http://www.who.int/mediacentre/factsheets/fs311/en/ (Accessed August 18, 2017)

Chapter 2

UNDERSTANDING BIKE SHARE CYCLIST ROUTE CHOICE USING GPS DATA: COMPARING DOMINANT ROUTES AND SHORTEST PATHS

2.1 INTRODUCTION

Active travel, any form of human-powered transportation like bicycling and walking, benefits not only physical fitness (Merom et al., 2010; Sahlqvist et al., 2012; Saunders et al., 2013), but also social and cognitive development (Badland and Oliver, 2012). For this reason, policy makers and urban planners continue to seek ways of increasing the use of active commute modes (Sallis et al., 2006; St-Louis et al., 2014). It has been suggested that cycling is more likely to replace motorized travel modes than walking given its faster speed and capability of covering longer distances, though most of the focus recently has been placed on walking (Dill, 2009). Cycling can benefit not only the environment by reducing carbon emissions (Rissel, 2009; Woodcock et al., 2009), but also health by reducing obesity, chronic diseases and weight gain (Andersen et al., 2000; Oja et al., 2011; Pucher et al., 2010). Bike share programs (BSP), providing bikes that can be picked up and dropped off at self-serve docking stations, have grown rapidly in past years; for example, the number of participating cities has increased from 13 in 2004 to 855 in 2014 (Fishman, 2015). Hamilton, Ontario is one such city operating a BSP commonly referred to as SoBi (Social Bicycles) Hamilton, which, at the time of its official launch in March 2015, had 750 GPS (global positioning system)-equipped bicycles located at over 100 hubs. The GPS feature means that cyclist routes can be tracked in real-time, providing an opportunity for route choice analysis.

In general terms, route choice analysis is necessary to appraise perceptions of route attributes, to forecast future traffic conditions on road networks, to simulate travel behavior under hypothetical scenarios, and to look at response and adaptation to message sources (Prato, 2009). Government policy makers, researchers and professionals can understand individual travel preferences by analyzing the route choice decision-making process in an effort to identify related determinants in terms of route attributes and the demographic characteristics of travelers. In the context of walking and cycling, they can develop policies and build facilities for encouraging greater use of such active travel modes. For example, studies of cyclist routes can help to identify what types of regulations and cycling infrastructure programs are useful in promoting the use of the bicycle for utilitarian trips in order to reduce automobile usage (Aultman-Hall et al., 1997; Su et al., 2010).

Most route choice studies create alternative routes using choice set generation methods. However, this research generates routes for all the trips between SoBi hubs from GPS data using a GIS-based map-matching toolkit (Dalumpines and Scott, 2011), and extracts unique routes from those duplicate hub-to-hub trips using a link signature extraction tool developed for this study. As a result, for a hub pair, it is possible to create choice sets of observed and alternative routes from actual routes taken by SoBi users instead of creating alternative routes using various techniques. Unlike most previous research based on individuals, this study generates choice sets on the basis of hub pairs to control origins and destinations for routes and investigates characteristics of these hub-to-hub routes within the BSP. Between a hub pair, the unique route with a maximum number of trips on it is regarded as the dominant route. As such, this study presents a new and essential attempt to explore the spatial distribution of dominant routes, which visually provides planners with road segments suitable for developing bicycle facilities. Given the usage frequencies of unique routes, extraction of hub-to-hub dominant routes can help determine cyclists preferences. In this study, this is achieved by comparing attributes between dominant routes and shortest paths based on distance, thereby identifying factors affecting cyclist behavior. Although SoBi Hamilton provides a real-time app for its users to show the number of bicycles and docks available at each hub, the app is not able to recommend optimal cycling routes. In this case, the dominant route between a hub pair could be considered as the optimal route rather than a shortest path because it is frequently chosen by multiple SoBi users. Finally, this study illustrates the usefulness of using bike share data for understanding route choice decisions since bicycles from recent BSPs are usually equipped with GPS devices.

The remainder of this paper is organized as follows. Section 2 reviews methods to collect data for route choice analysis and discusses variables, including route attributes and cyclist characteristics, affecting cyclist route choice behavior. Section 3 briefly describes the study area and sources of both the cycling network and GPS dataset. Section 4 discusses the generation of route choice sets using the GIS-based map-matching toolkit and link signature tool, and the methods of statistical analysis involving the normalized Gini coefficient and paired t-test. Section 5 displays results derived from statistical analysis and spatial patterns of dominant routes between hubs. Section 6 summarizes major findings, significant limitations, and future implications of this research.

2.2 BACKGROUND

2.2.1 Data Collection Methods

Most previous bicycle route choice studies used stated preference (SP) surveys or revealed preference (RP) surveys as the data collection method (Aultman-Hall et al., 1997; Guttenplan & Patten, 1995; Howard & Burns, 2001; Hunt & Abraham, 2007;

Tilahun et al., 2007). Respondents for SP surveys make a choice among different facilities or different route options, forcing them to trade off some positive features (Broach et al., 2012). According to Abraham et al. (2002), SP surveys can collect a large sample size of data easily and cheaply, and avoid inter-correlations among attributes, but the ability of respondents to convert their usual routes and preferred facilities in mind to match the created choice set in the survey may lead to missing some important features for route choices. On the contrary, RP surveys gather information based on actual route choices made by participants, so the collected data can reveal preferences in a real choice environment. However, the tedious and time-consuming collection process limits the sample size, and the capability of participants to precisely recall routes influences the match between revealed routes and actual route networks (Stinson & Bhat, 2003). In order to accurately recall routes that participants choose, GPS devices that can automatically record traces have been used for data collection in many more recent route choice studies including Broach et al. (2012), Hood et al. (2011), Menghini et al. (2010), and Papinski and Scott (2013). The drawbacks to using GPS data for research include the high cost of equipping GPS devices and the transformation of points recorded by devices into actual traces that users take.

2.2.2 Potential Determinants of Cyclist Route Choice Behavior

Some work has regarded travel time or distance as the most important factor influencing the route choices of cyclists for commuting purposes (Aultman-Hall et al, 1997; Sener et al., 2009; Stinson & Bhat, 2003). However, Tilahun et al. (2007) discovered bicycle route preferences of trading off travel time for particular facilities, such as designated bike lanes, trails off street, and parking on the street side. Similarly, Winters et al. (2010b) found that utilitarian bicycle trips are 360 m longer than the shortest path route in Metro Vancouver because cyclists are willing to detour slightly to ride on routes with more bicycle facilities. Almost all the earlier studies explore the influence of bicycle facility type on commuter cyclist route choice. Broach et al. (2012) illustrated that off-street/separated bike paths that definitely have no motorized traffic are preferred, followed by bike boulevards that are neighborhood streets with traffic calming features. Simultaneously, on-street bike lanes can more or less mitigate the negative influence of traffic nearby, so they are more attractive than a heavy traffic street without a bike lane, but not preferred compared to a street with low traffic volume (Broach et al., 2012). Consistently, Winters & Teschke (2010) found the order of bicycle route preferences: off-street paths, physically separated routes adjacent to main streets, neighborhood routes, rural roads and routes on major roads.However, these findings are opposite to some studies that have found that bike lanes on streets are more attractive than separated bike paths followed by routes without bicycle facilities (Hood et al., 2011; Sener et al., 2009; Stinson & Bhat, 2003).

In addition to bicycle facilities, bicyclist route choice behavior can be affected by other route attributes. Utilitarian or purposeful cyclists generally prefer fewer stop signs, red lights, and major cross streets (Sener et al., 2009; Stinson & Bhat, 2003). Many previous studies have emphasized the obvious importance of slope, turn frequency and motorized traffic volume; that is, cyclists tend to avoid steep slopes, turns and exposure to heavy traffic volume. Statistically, cyclists will choose to avoid a turn with a cost of less than 0.17 km, and if the detour of avoiding climbing a hill 10 m high is no more than 0.59 km, cyclists will choose the detour (Hood et al., 2011).

With regard to bridges, commuter cyclists prefer those without automobiles, or those with a barrier between motorized and non-motorized traffic, or those without any special provisions for cyclists but connected to bike lanes (Stinson & Bhat, 2003). Few studies have explored the effects of on-street parking characteristics, such as presence of parking, parking type, parking occupancy rate, and length of parking area, on cyclist route choice. However, in Texas, Sener et al. (2009) found that cyclists prefer routes with no or minimal parking along the street, and discovered a preference of routes with angled parking among all the alternative routes with on-street parking.

Additional factors that affect cyclists routes in terms of cyclist characteristics, such as age, gender, income, and cycling experience, are also commonly explored in some studies. Lower income and younger respondents tend to make shorter commutes, and there is no significant relationship between gender and travel time, though males make many more bicycle trips than females (Shafizadeh & Niemeier, 1997). Nevertheless, Aultman-Hall et al. (1997) did not find any significant correlation between age and commuting distance for adults, though a few respondents under 18 years old in Guelph, Ontario rode a slightly shorter distance. As cycling experience and motor vehicle traffic volume are associated with safety, more experienced cyclists with better skills may prefer to ride in the street than on a bike path, while inexperienced/infrequent cyclists are more sensitive to safety concerns and traffic volumes or speed limits (Antonakos, 1994; Hunt & Abraham; 2007; Lott & Tardiff, 1978; Winters & Teschke, 2010). Consistently, experienced cyclists are more willing to ride on bike lanes rather than separated bike paths, and their familiarity with roads helps them mitigate delay and safety issues on utilitarian trips (Broach et al., 2012). In addition, females are found to have a higher probability to cycle on perceived safer routes than males (Tilahun et al., 2007). Similarly, females attach more importance to safety, few hills, and convenience for errands than males (Antonakos, 1994). As a reaction to a cycling accident, it is interesting to find that females are more likely to alter their routes than males (Howard & Burns 2001).

In summary, the most commonly used route attributes for analyzing cycling route choice include slopes, turn frequency, traffic volumes, and bicycle facilities, such as bike lanes, separated bike paths, and residential streets with traffic calming. Other route attributes that may affect choice include stop signs, red lights and major cross streets, and on-street parking characteristics. In terms of cyclist characteristics, some previous studies found that lower income and younger cyclists have a higher sensitivity to travel time, but a lower sensitivity to facility quality, and are more willing to use separated bike paths, sidewalks, and dirt trails, while other studies did not find any significant relationships. Experienced bicyclists have a stronger preference to ride on a bike lane than on a bike path, while inexperienced bicyclists care more about safety, traffic volumes and speed limits. Concerning gender, safety, few hills, convenience for errands, and quality of facilities are more important for women than men.

2.3 DATA

2.3.1 Study Area

Hamilton is a densely populated city in the province of Ontario with a population of 536,917 in 2016 (Statistics Canada, 2016). Public transit in the city depends on buses. In March 2015, Hamilton launched a bike share program called SoBi Hamilton, consisting of 750 bicycles and over 100 hubs with a designated distance of 300 to 600 meters between hubs. As shown in Figure 2.1, all of the hubs, represented by black dots, are located below the escarpment, in the vicinities of downtown Hamilton and McMaster University.

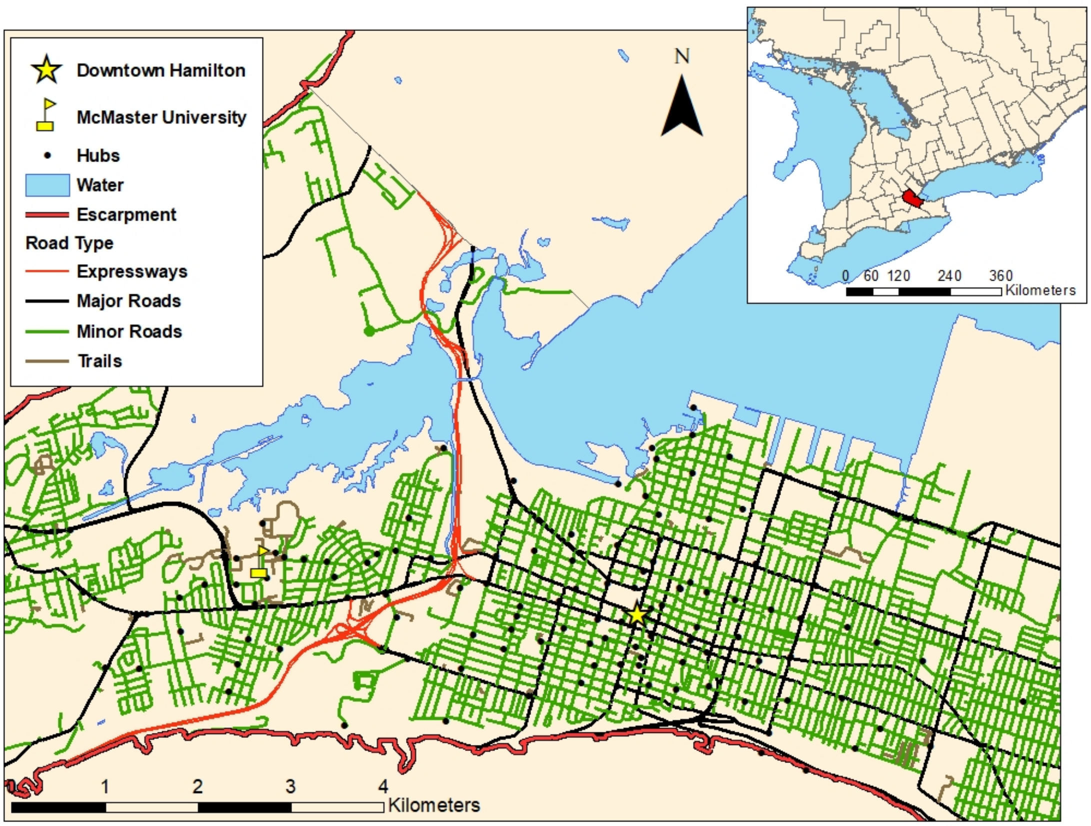

Figure 2.1: Study area of SoBi Hamilton.

2.3.2 Cycling Network

For this study, the cycling network was created from two data sources: a road network containing trails for the City of Hamilton (DMTI Spatial, 2015) and bikeways (Open & Accessible Data, 2016). Essentially, an attribute describing different types of bikeways was added to links comprising the road network, producing a more detailed network, consisting of 22,710 links and 16,731 junctions, containing 6 road types and 10 categories of bikeways. Cyclists, however, only traveled along three road types: major roads, minor roads, and trails (Figure 2.1). A digital elevation model with a resolution of 30 m was used for to calculate link slopes.

2.3.3 GPS Dataset

As mentioned earlier, SoBi bikes are GPS-equipped, meaning that the number of bikes at hubs can be determined in real time, informing users of bicycle and parking availability at start hubs and end hubs, respectively. More importantly, cycling routes can be tracked in real time, providing reliable data for route choice analysis. Although stated preference and revealed preference surveys have been widely used in past studies due to the ease and cost effectiveness of collecting data, this project uses GPS points to reveal the actual routes of users.

This study obtained from SoBi Hamilton 161,426 GPS trajectories describing the actual routes of bike share users over a 12-month period (April 1, 2015 to March 31, 2016). However, only 132,396 trips were between two different hubs. This study grouped trips by origin-destination (O-D) hub pairs and used a GIS-based map-matching toolkit to generate multiple route attributes for exploring determinants of cyclist route choice behavior.

2.4 METHODOLOGY

2.4.1 Data Generation

GIS-based map-matching toolkit

The GIS-based map-matching toolkit developed by Dalumpines and Scott (2011) was used to match the GPS trajectories with the cycling network and extract a series of attributes describing each route. The main principle of the toolkit is to use the shortest path algorithm, provided by the Network Analyst extension in ArcGIS, given basic inputs, including stops representing origin and destination points, and barriers constraining the shortest path algorithm, to follow the GPS trajectories. The map-matching process is shown in Fig. 2. Specifically, origin and destination points are identified as stops, and a polyline feature is created from a set of GPS points representing a hub-to-hub trip (Figure 2.2a). Then, a buffer with a default distance chosen by the user (in this case, 50 m) is created around the polyline feature to identify barriers by intersecting with road segments (Figure 2.2b). The observed route is reproduced because barriers force the shortest path algorithm to follow the stream of GPS trajectories (Figure 2.2c). Since the fit of GPS trajectories to a road network is sensitive to buffer distance, Dalumpines & Scott (2011) did an experiment testing the impact of different buffer distances on map-matching results and found that a 50 m buffer maximizes the accuracy of the map-matching algorithm.

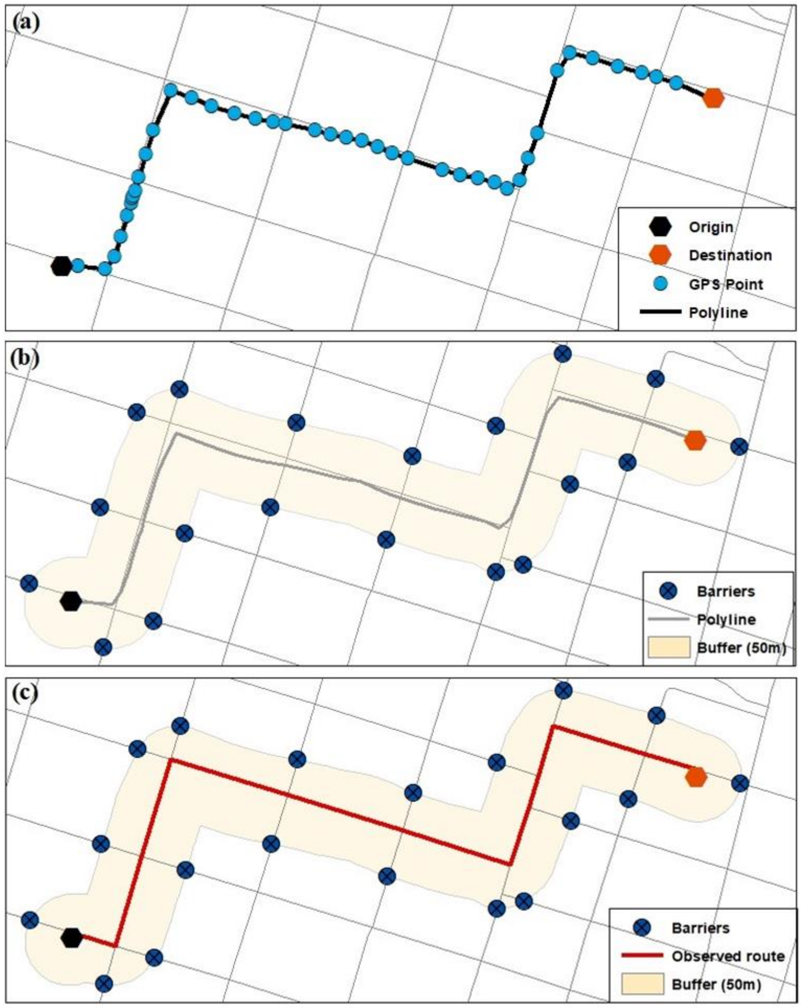

Figure 2.2: Conversion from GPS trajectories to an observed route using the map-matching toolkit.

Link signature extraction

In the cycling network, each road segment (link) has an ID. The link signature extraction tool, developed for this study, extracts all links comprising a route and combines their link IDs in ascending order using colons for separation. The list of ordered link IDs is the original link signature for one trip. As shown in Figure 2.3, overlapping trips may be slightly different at the origin and the destination, which can be caused by the delay of activating time. However, such trips should still be seen as exactly the same. Thus, the links at the start and end points are removed from the link ID list to generate an altered link signature for the core portion of each trip. As a result, unique routes from hub-to-hub are extracted from all the actual map-matched trips according to their core link signatures. At the same time, the use frequency of each unique route is calculated according to the number of trips traversing it.

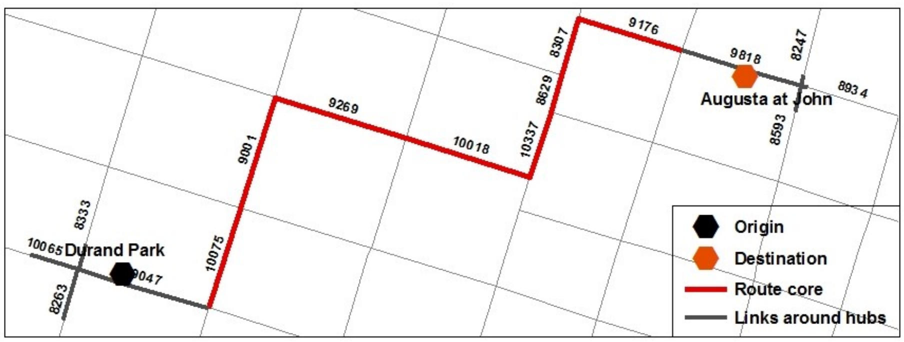

Figure 2.3: Generation of core link signatures for each hub-hub trip.

2.4.2 Statistical Analysis

In order to compare dominant routes with their shortest path counterparts based on distance, it is necessary to determine whether a dominant route exists between a hub pair; that is, whether the most popular route with the most trips on it between a hub pair can be regarded as the dominant route.

Normalized Gini coefficient

The Gini coefficient is a widely used statistical measure showing dispersion in a set of values. Its value ranges from 0 to $(1 - 1/N)$, where N refers to the sample size.

A Gini coefficient of 0 indicates perfect equality, while a value that is maximized indicates complete inequality. In the field of route choice analysis, the Gini coefficient is used to evaluate whether a preference exists among all the unique routes between O-D pairs. Lima et al. (2016) normalized the Gini coefficient to compare its value G for an O-D group of trips with the corresponding heterogeneous number of unique routes N (>1) by the following formula:

$$G_n = \frac{G}{1 - 1/N} \tag{2.1}$$

In this case, no matter how many unique routes are between an O-D group, a G_n of 1 indicates perfect inequality (e.g., for an O-D group with a large number of trips, where all the trips travel along the same route, while the other alternative routes have no trips, the normalized Gini coefficient will be 1). In this study, however, perfect inequality cannot be reached because only hub pairs containing at least 2 unique routes are considered in the analysis. A normalized Gini coefficient of 0 indicates perfect equality, where all unique routes have the same number of trips on them.

Paired *t*-test

The paired *t*-test is a parametric test that identifies differences between paired measurements respectively from two paired samples, and determines whether the mean difference is statistically significant. In this study, for hub pairs, the paired *t*-test is used to compare differences between dominant routes and corresponding shortest paths based on distance.

2.5 RESULTS

For this study, Table 2.1 lists general information about the SoBi dataset extracted from the GPS points. There are 7,437 hub pairs in total, but only 5,561 of them are used for this study. Excluded are hub pairs with round trips or only one unique route between them. The map-matching procedure produces 161,426 trips along the cycling network, where 82% are from hub-to-hub. Given the core link signature of each hub-to-hub trip, 49,120 unique routes are extracted from all the trips. Some of the unique routes are then removed from the entire choice set, including those not between hub pairs for study, those with portions on expressways likely due to GPS errors, and those far away from start and end hubs caused by GPS errors where 600 m is used due to the designated distance separating hubs. Statistically, the average number of hub-to-hub unique routes is 8, while the maximum is 77 and the minimum is 2. Concerning the number of trips on each unique route, dominant routes between hub pairs are identified, and only 299 of them follow shortest paths based on distance.

Table 2.1: GPS data processing summary.

Description	Number	Notes
Map-matched trips	161,426	
Trips between hubs	132,397	
Hub pairs	7,437	
Hub pairs for study	5,561	Excludes 1) Round trips, 2) One route
Unique hub-to-hub routes	49,120	
Unique hub-to-hub routes for study	41,369	Excludes 1) Not within 600 m of hub, 2) On expressways
Average # routes between hub pairs	8	
Maximum # routes between hub pairs	77	
Minimum # routes between hub pairs	2	
Dominant hub-to-hub routes	4,563	Excludes dominant routes with 1 trip
Dominant route = shortest path (m)	299	

In order to determine whether some routes are used more frequently than others, the normalized Gini coefficient G_n is used. Specifically, for a hub pair, the closer to 0 the G_n is, the more evenly routes are chosen; the closer to 1 the G_n is, the more biased route choice is toward one route. Generally, hub pairs have high a G_n with a median value of 0.5, indicating that a dominant route exists.

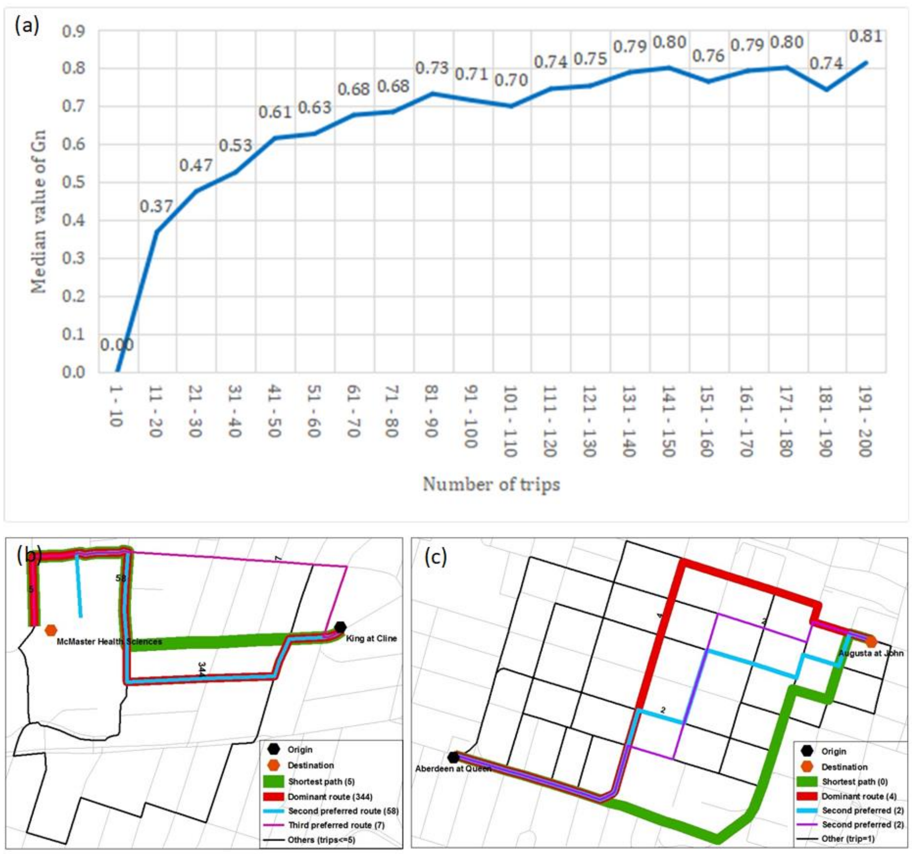

Figure 2.4: Cyclist route choices.

(a) The median values of G_n versus the number of hub-to-hub trips.

(b) Demonstration of a hub pair with a high G_n.

(c) Demonstration of a hub pair with a low G_n.

The relationship between median values of the normalized Gini coefficients for hub pairs with no less than two unique routes and the number of corresponding trips is plotted in Figure 2.4a. The results suggest that a route is much more preferred as the number of trips between hubs increases. In addition, the tendency of G_n shows suggests that if a unique route is repeated more than 25 times, a biased route tends to be the dominant route between a hub pair. Figure 2.4b shows an example of all the unique routes between a hub pair with a high normalized Gini coefficient of 0.91, where the dominant route has 344 trips on it, followed by the second preferred route with 58. The shortest distance route has only 5 trips on it. On the contrary, the example of the hub pair with a low G_n of 0.19 contains 20 unique routes and 26 trips, where only 4 trips travel along the dominant route and no trip are on the shortest distance route (Figure 2.4c). In general, the number of trips on dominant routes accounts for about 49% of the total number of trips between hub pairs for study.

The spatial pattern of SoBi users dominant routes along the entire cycling network shows that highly used road segments usually contain a series of hubs (Figure 2.5a). The road segments without bikeways containing a large number of dominant routes are likely suitable candidates for adding bicycle facilities (Figure 2.5b). At the same time, most road candidates for additional bicycle facilities are around a cluster of hubs and have relatively high accessibility to the existing bicycle infrastructure. Additionally, a visual comparison of Figure 2.5a and 2.5b shows that most frequently used roads are equipped with bicycle facilities, indicating that cyclists tend to travel along roads with bikeways.

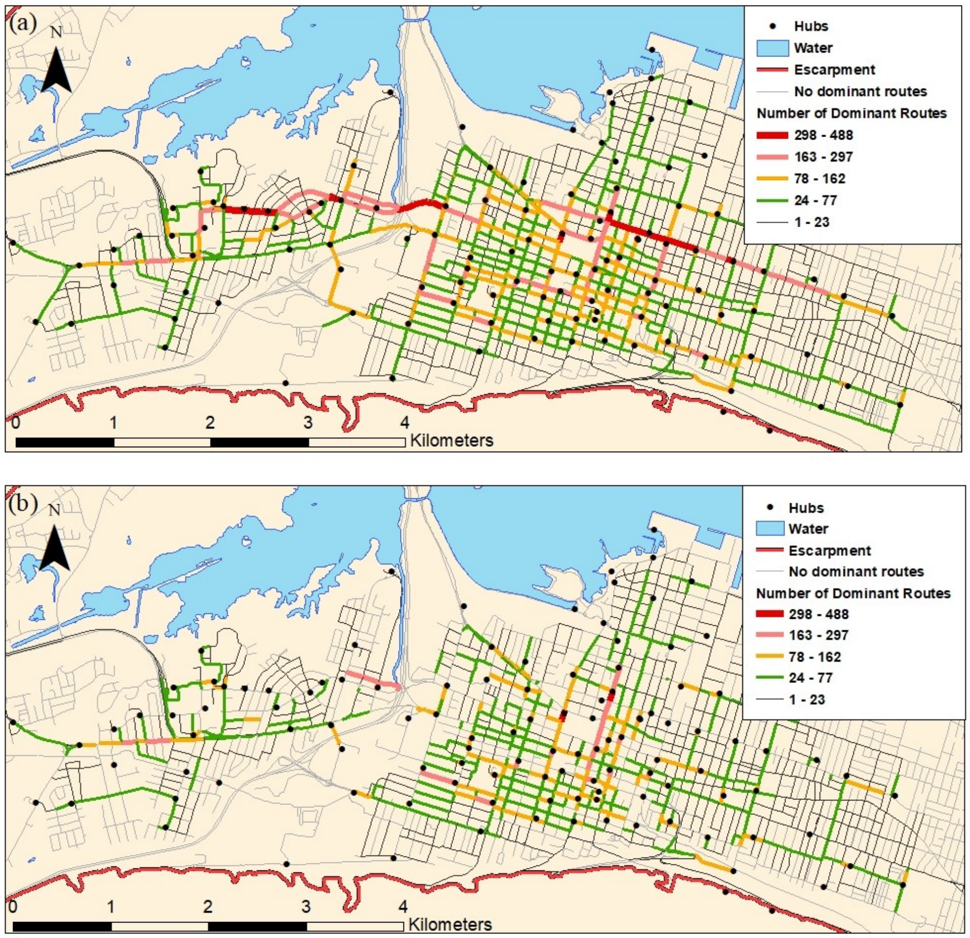

Figure 2.5: Spatial distribution of dominant routes.

(a) on all the road segments and (b) on roads without bikeways.

The paired t-test was used to explore differences between dominant routes and shortest paths based on distance. Table 2.2 shows summary statistics (mean values for route attributes), and t-statistics to identify whether differences between dominant-route attributes and shortest-path attributes are statistically significant. Specifically, bolded values suggest that differences are statistically significant at the 5% significance level.

Table 2.2: Route attributes of dominant routes compared to corresponding shortest paths (distance).
Bolded t-statistics imply differences significant at the 5% significance level.

Variables	**Dominant**	**Shortest**	**Difference**	***t* statistic**
Length (m)	2137.98	1971.35	166.63	**36.25**
RDI (compared to straight line)	1.36	1.26	0.1	**23.96**
Unique road segments	5.36	5.08	0.28	**6.34**
Intersections	22.83	22.59	0.24	**3.55**
Mean route distance btw intersections (m)	95.2	91.74	3.46	**13.86**
Longest leg length (m)	1009.72	1051.13	-41.41	**-5.24**
Turn statistics				
Left turn	2.01	1.9	0.11	**4.51**
Right turn	2.05	1.89	0.15	**6.51**
Sharpe left turn	0.13	0.11	0.02	**3.23**
Sharpe right turn	0.13	0.1	0.03	**5.89**
Total turns	4.31	4	0.3	**6.92**
% of route based on slope				
Slope 0-2%	65.16	66.44	-1.28	**-6.37**
Slope 2-4%	20.86	19.92	0.94	**6.16**
Slope 4-6%	5.93	6.46	-0.52	**-7.03**
% of route based on road type or bikeway type				
Major (%)	31.74	41.53	-9.78	**-22.83**
Minor (%)	66.42	56.7	9.72	**22.77**
Trail (%)	1.83	1.75	0.08	**1.99**
Designated BL (%)	26.92	20.36	6.55	**19.87**
Cautionary un-signed BR-HT (%)	2.8	5.36	-2.55	**-16.19**
Cautionary un-signed BR-MT (%)	3.39	2.1	1.29	**10.14**
Cautionary un-signed BR-LT (%)	1.12	1.06	0.06	1
Signed on-street BR-MHT (%)	0.28	0.26	0.02	0.72
Signed on-street BR-LT (%)	10.21	8.24	1.97	**9.22**
Separated bike path (%)	0.27	0.16	0.11	**3.64**

Note: Differences are calculated as attributes of dominant routes minus those of shortest distance routes. Thus, positive t statistics correspond to higher values for dominant route attributes, while negative t statistics suggest higher values for shortest path attributes.

Dominant routes are significantly longer than shortest paths based on distance, implying that cyclists may detour for other route characteristics. Consistently, the mean route directness index (RDI) of dominant routes is larger than that of shortest distance routes. The RDI, measuring the ratio of a routes distance to the straight-line distance between its origin and destination, shows the efficiency/circuity of a route. The mean RDI values of dominant routes and shortest paths are 1.36 and 1.26, respectively indicating that routes are 36% and 26% longer than the straight-line distances between hubs. In this case, dominant routes are, on average, 10% less efficient than corresponding shortest paths based on distance. Simultaneously, dominant routes have more unique road segments, more intersections, and longer mean route distance between intersections compared to shortest distance routes, while the average distance on the longest leg of dominant routes is significantly shorter than that of shortest distance routes.

In terms of turn frequency, significant differences are identified for the number of sharp turns and normal turns. The average number of turns is 4.31 for dominant routes with 2.01 left turns and 2.05 right turns, compared to 4.00 for shortest paths (distance) with 1.90 left turns and 1.89 right turns. Both dominant and shortest distance routes are found to avoid sharp turns over 90 degrees, though the number of sharp left and right turns are both slightly higher in dominant routes than in shortest distance routes. With regard to slopes, cyclists will feel a slope of 3-4%, and do not like climbing a grade over 4% (Transport Canada, 2010). Also, slopes over 5% are regarded as steep hills (Winters et al., 2010a). In this study, 92% of dominant routes on average are on slopes of no more than 6%, where 65% are on 0-2% gradients. Similarly, 93% of shortest paths based on distance are on slopes of no more than 6%, where 66% are on 0-2% gradients. Compared to shortest distance routes, cyclists prefer to take longer trips on 2-4% slopes, and avoid steep slopes of 4-6% as well as flat areas with slopes no more than 2%. A possible reason for avoiding flat areas for dominant routes may be a tradeoff for other attributes.

Dominant routes tend to follow minor roads covering 66% of a route and avoid major roads comprising 32%. In comparison, shortest distance routes contain fewer proportions on minor roads (57%) and more proportions on major roads (42%). At the same time, the distance of a dominant hub-to-hub route on trails on average is significantly longer than that of its corresponding shortest path. Concerning bicycle facilities, it is more likely for SoBi users to choose routes with longer designated on-street bike lanes, cautionary un-signed bike routes on streets with moderate traffic volume, signed on-street bike routes with mostly low traffic volume, and separated bike paths instead of shortest distance routes. Additionally, SoBi cyclists tend to avoid riding on cautionary un-signed bike routes on streets with high traffic volume compared with shortest paths based on distance.

2.6 CONCLUSIONS

Route choice analysis is commonly based on stated preference surveys or revealed preference surveys, which can impact sample size and result accuracy. Some updated studies used GPS devices to automatically record the actual routes taken by users and create alternative routes using some form of choice set generation (Broach et al., 2012; Hood et al., 2011; Menghini et al., 2010). However, none of them created choice sets using actual cyclists routes, which in this study were extracted from all the duplicate trips between an origin and a destination derived from GPS data. The usage of each unique route is identified afterwards. As a result, all the generated actual unique routes between an origin and destination (OD) pair can be regarded as an OD choice set. This paper used a GIS-based map-matching toolkit to produce all the trips along the cycling network using a GPS dataset, and introduced a new link signature extraction tool to do all three aforementioned jobs. Lima et al. (2016) introduced the idea of dominant routes for drivers frequent trips and found a tendency of each driver towards a preferred route between an OD pair. This study not only demonstrates that a dominant route is chosen by multiple cyclists instead of only individuals between a bike share hub pair, but also identifies route attributes contributing to such dominance by comparing with the corresponding shortest path based on distance.

Normalized Gini coefficients indicate that cyclists tend to have a dominant route between a hub pair, and the bias becomes much stronger as the number of trips increase between them. According to the core link signatures of routes, only 7% of dominant routes chosen by SoBi users are shortest paths based on distance. In this study, dominant routes are compared to their corresponding shortest distance routes using paired t-tests. Significant differences are identified in many route attributes, including mean route distance between intersections, the length of longest leg, the proportion of route along road segments with different slopes, with different road types, and with different bikeways, as well as the number of unique road segments, intersections, and turns. Similar comparisons can be conducted between other alternative routes, such as dominant routes and second preferred routes.

The most significant differences in this research are found in route distance, RDI, road types, and bicycle facilities. Hub-to-hub dominant routes are 10% less efficient than their corresponding shortest path based on distance according to RDI values, suggesting that shortest distance routes are not the optimal choice for cyclists in the real world. With respect to road type, minor roads and trails are preferred to major roads for cycling. Moreover, SoBi users are willing to detour for bicycle facilities without high traffic volumes.

One of the most significant constraints affecting further interpretation of the results is the lack of demographic data. The findings could be more comprehensive if general socio-demographic information of SoBi users is collected (e.g., age, gender, income, cycle experience). In addition, the shortest-path alternatives based on time for cycling cannot be generated because the cycling speed varies between cyclists, and it will also change over time, especially for long distances.

The future implications of this study are twofold. On one hand, it provides a new tool for future studies to extract unique routes from a large dataset of trips between ODs and generate choice sets based on actual routes tracked by GPS devices. Also, since a series of attributes of unique routes have been produced by the GIS-based map-matching, the next step is to explore the contribution of those attributes to the use frequency of unique routes using models. In addition, since most BSPs recently contain the GPS technology, this study provides a brand new and cost effective way to explore cyclists route choice decision-making process using bike share users GPS trajectories. The findings from this research bring positive benefits to urban planning and policy making in the City of Hamilton by examining the spatial distribution of cyclists dominant routes and the influence of bicycle facilities, which can determine what types of bicycle facilities are suitable to be developed on which road segments. Simultaneously, for SoBi users, a recommended cycling route between hubs can be provided on the basis of generated dominant routes in this study.

REFERENCES

Abraham, J. E., McMillan, S., Brownlee, A. T., & Hunt, J. D. (2002). Investigation of cycling sensitivities. Presented at the 81st Annual Meeting of the Transportation Research Board, Washington, D.C. https://www.researchgate.net/publication/228706636_Investigation_of_Cycling_Sensitivities (Accessed August 24, 2017).

Andersen, L. B., Schnohr, P., Schroll, M., & Hein, H. O. (2000). All-cause mortality associated with physical activity during leisure time, work, sports, and cycling to work. *Archives of Internal Medicine*, 160: 1621-1628.

Antonakos, C. L. (1994). Environmental and travel preferences of cyclists. *Transportation Research Record: Journal of the Transportation Research Board*, 1438: 25-33.

Aultman-Hall, L., Hall, F., & Baetz, B. (1997). Analysis of bicycle commuter routes using geographic information systems: implications for bicycle planning. *Transportation Research Record: Journal of the Transportation Research Board*, 1578: 102-110.

Badland, H. M., & Oliver, M. (2012). Child independent mobility: making the case, and understanding how the physical and social environments impact on the behaviour. In: Turunen, E., Koskinen, A. (Eds.), *Urbanization and the Global Environment*. New York: NOVA Science Publishers, pp. 51-79.

Broach, J., Dill, J., & Gliebe, J. (2012). Where do cyclists ride? A route choice model developed with revealed preference GPS data. *Transportation Research Part A: Policy and Practice*, 46: 1730-1740.

Dalumpines, R., & Scott, D. M. (2011). GIS-based map-matching: development and demonstration of a postprocessing map-matching algorithm for transportation research. In Gertman, S., Reinhardt, W., Toppen, F. (Eds.), *Advancing Geoinformation Science for a Changing World,* Berlin: Springer, pp. 101-120.

Dill, J. (2009). Bicycling for transportation and health: the role of infrastructure. *Journal of Public Health Policy*, 30: S95-S110.

DMTI Spatial. (2015). CanMap Content Suite. https://www.dmtispatial.com/canmap/ (Accessed August 24, 2017)

Fishman, E. (2015). Bikeshare: a review of recent literature. *Transport Reviews*, 36: 1-22.

Guttenplan, M., & Patten, R. (1995). Off-road but on track: using bicycle and pedestrian trails for transportation. *TR News*, 178: 7-11.

Hood, J., Sall, E., & Charlton, B. (2011). A GPS-based bicycle route choice model for San Francisco, California. *Transportation Letters*, 3: 63-75.

Howard, C., & Burns, E. (2001). Cycling to work in Phoenix: route choice, travel behavior, and commuter characteristics. *Transportation Research Record: Journal of the Transportation Research Board*, 1773: 39-46.

Hunt, J. D., & Abraham, J. E. (2007). Influences on bicycle use. *Transportation*, 34: 453-470.

Lima, A., Stanojevic, R., Papagiannaki, D., Rodriguez, P., & Gonzlez, M. C. (2016). Understanding individual routing behaviour. *Journal of the Royal Society Interface*, 13: 20160021.

Lott, D. F., & Tardiff, T. (1978). Evaluation by experienced riders of a new bicycle lane in an established bikeway system. *Transportation Research Record: Journal of the Transportation Research Board*, 683: 40-46.

Menghini, G., Carrasco, N., Schssler, N., & Axhausen, K. W. (2010). Route choice of cyclists in Zurich. *Transportation Research Part A: Policy and Practice*, 44: 754-765.

Merom, D., van der Ploeg, H. P., Corpuz, G., & Bauman, A. E. (2010). Public health perspectives on household travel surveys: active travel between 1997 and 2007. *American Journal of Preventive Medicine*, 39: 113-121.

Oja, P., Titze, S., Bauman, A., de Geus, B., Krenn, P., RegerNash, B., & Kohlberger, T. (2011). Health benefits of cycling: a systematic review. *Scandinavian Journal of Medicine & Science in Sports*, 21: 496-509.

Open & Accessible Data. (2016). Bikeways. https://www.hamilton.ca/city-initiatives/strategies-actions/open-accessible-data (Accessed August 24, 2017)

Papinski, D., & Scott, D. M. (2013). Route choice efficiency: an investigation of home-to-work trips using GPS data. *Environment and Planning A*, 45: 263-275.

Prato, C. G. (2009). Route choice modeling: past, present and future research directions. *Journal of Choice Modelling*, 2: 65-100.

Pucher, J., Buehler, R., Bassett, D. R., & Dannenberg, A. L. (2010). Walking and cycling to health: a comparative analysis of city, state, and international data. *American Journal of Public Health*, 100: 1986-1992.

Sahlqvist, S., Song, Y., & Ogilvie, D. (2012). Is active travel associated with greater physical activity? The contribution of commuting and non-commuting active travel to total physical activity in adults. *Preventive Medicine*, 55: 206-211.

Sallis, J. F., Cervero, R. B., Ascher, W., Henderson, K. A., Kraft, M. K., & Kerr, J. (2006). An ecological approach to creating active living communities. *Annual Review of Public Health*, 27: 297-322.

Saunders, L. E., Green, J. M., Petticrew, M. P., Steinbach, R., & Roberts, H. (2013). What are the health benefits of active travel? A systematic review of trials and cohort studies. *PLoS One*, 8: e69912.

Sener, I. N., Eluru, N., & Bhat, C. R. (2009). An analysis of bicycle route choice preferences in Texas, US. *Transportation*, 36: 511-539.

Shafizadeh, K., & Niemeier, D. (1997). Bicycle journey-to-work: travel behavior characteristics and spatial attributes. *Transportation Research Record: Journal of the Transportation Research Board*, 1578: 84-90.

Statistics Canada. (2016). Census Profile, 2016. Census.http://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/index.cfm?Lang=E (Accessed August 24, 2017).

Stinson, M., & Bhat, C. (2003). Commuter bicyclist route choice: analysis using a stated preference survey. *Transportation Research Record: Journal of the Transportation Research Board*, 1828: 107-115.

St-Louis, E., Manaugh, K., van Lierop, D., & El-Geneidy, A. (2014). The happy commuter: a comparison of commuter satisfaction across modes. *Transportation Research Part F: Traffic Psychology and Behaviour*, 26: 160-170.

Su, J. G., Winters, M., Nunes, M., & Brauer, M. (2010). Designing a route planner to facilitate and promote cycling in Metro Vancouver, Canada. *Transportation Research Part A: Policy and Practice*, 44: 495-505.

Tilahun, N. Y., Levinson, D. M., & Krizek, K. J. (2007). Trails, lanes, or traffic: valuing bicycle facilities with an adaptive stated preference survey. *Transportation Research Part A: Policy and Practice*, 41: 287-301.

Transport Canada. (2010). Bicycle end-of-trip facilities: a guide for Canadian municipalities and employers. Ottawa: Transport Canada. Catalogue No: T22-194/2010E-PDF.

Winters, M., Brauer, M., Setton, E. M., & Teschke, K. (2010a). Built environment influences on healthy transportation choices: bicycling versus driving. *Journal of Urban Health*, 87: 969-993.

Winters, M., & Teschke, K. (2010). Route preferences among adults in the near market for bicycling: findings of the Cycling in Cities study. *American Journal of Health Promotion*, 25: 40-47.

Winters, M., Teschke, K., Grant, M., Setton, E., & Brauer, M. (2010b). How far out of the way will we travel? Built environment influences on route selection for bicycle and car travel. *Transportation Research Record: Journal of the Transportation Research Board*, 2190: 1-10.

Woodcock, J., Edwards, P., Tonne, C., Armstrong, B. G., Ashiru, O., Banister, D., ... Franco, O. H. (2009). Public health benefits of strategies to reduce greenhouse-gas emissions: urban land transport. *The Lancet*, 374: 1930-1943.

Chapter 3

DETERMINANTS OF BIKE SHARE CYCLIST ROUTE CHOICE BEHAVIOR USING GPS DATA

3.1 INTRODUCTION

Walking and cycling, two forms of active travel, can improve a persons health by reducing the likelihood of obesity (Gordon-Larsen et al., 2009) and chronic disease (Frank et al., 2006), and the environment by reducing carbon emissions and noise pollution (De Nazelle et al., 2011). According to Dill (2009), although most contemporary studies have focused on walking, cycling, rather than walking, has a greater potential to replace automobiles because of higher speed and the ability of traveling longer distances. A bike share program (BSP), an automated rental program providing bicycles that can be picked up and locked at hubs, is a cost-effective form of public transit that is healthy for users and can help improve traffic congestion and air pollution (Kisner, 2011). According to Davis (2014), newer BSP bicycles are usually equipped with global positioning system (GPS) devices, and technologies are developed to provide mobile applications and/or websites that can reveal bicycle availability at hubs. Although the global number of bike share programs has rapidly increased over the last decade, BSPs have been around for about 50 years (Shaheen et al., 2010). Fishman (2016) found that the number of cities with BSPs has risen from 13 in 2004 to 855 in 2014. In 2015, the City of Hamilton, Ontario joined this group by launching a BSP called SoBi (Social Bicycles) Hamilton with 750 bicycles equipped with GPS devices and over 100 hubs. The availability of bicycles at hubs and the actual traces of each SoBi user are recorded. Such information benefits route choice analysis by providing origins, destinations, and observed routes using GPS points.

Route choice is a decision-making process in which a person selects a specific route, based on its attributes, from a specified origin to a destination (Bapat et al., 2017). Route choice modeling consists of two parts: one is the generation of alternative routes for the corresponding observed route to form a choice set; the other is the determination of the likelihood of choosing one route from a choice set (Ben-Akiva et al., 2004). Prato (2009) states that it is necessary to analyze route choice because it can help researchers, policy makers, and urban planners evaluate travelers perceptions of route attributes, predict commuter travel behavior, learn about commuter reaction and adaptation to information sources, and estimate future traffic volumes on road networks. Therefore, the analysis of cyclist route choice behavior can contribute to promoting the use of bicycles by understanding their preferences for bicycle infrastructure and other route attributes.

Usually, choice set generation methods are used to create alternative routes for an observed route from an origin to a destination. However, this research regards each hub pair as a cluster and all the actual routes between the two hubs forming the pair as members of a choice set. Simultaneously, the weight of each route depends on the number of trips taken along the route. Unlike previous studies, this project does not derive choice sets using traditional choice set generation techniques, but instead uses actual cyclists routes and controls origins and destinations for multiple participants instead of individuals. This means that preferences for route attributes are evaluated by cyclists route choices in the real world. A GIS-based map-matching toolkit is used to transform GPS points into trips along the cycling network; then, unique routes are extracted from those duplicate trips based on their core link signatures. As a result, a large dataset containing all the actual routes from hub-to-hub taken by SoBi users is created. In order to determine whether a particular preference toward one route exists among all the unique routes between a hub pair, a normalized Gini coefficient (G_n) is calculated. For comparison, this study estimates a global model, a mild preference (medium G_n) model, and a strong preference (high G_n) model to identify the contribution of each route attribute to cyclist route choice decisions.

This paper mainly consists of six sections including this introduction and the list of references. Section 2 discusses different data collection methods, choice set generation techniques, and travel behavior models for route choice analysis. Section 3 describes the collection and processing of GPS data, including converting GPS trajectories to map-matched trips and creating choice sets for SoBi by extracting unique routes from trips, as well as the principles of G_n and route choice modelling Path-Size Logit (PSL) model. Section 4 presents descriptive results derived from G_n and three models with different preference levels using PSL. Section 5 concludes significant findings, contributions, and future implications of this research.

3.2 BACKGROUND

The most commonly used data collection methods for many previous studies in the realm of route choice analysis have been stated preference (SP) surveys or revealed preference (RP) surveys (Howard & Burns, 2001; Hunt & Abraham, 2007; Tilahun et al., 2007; Guo & Loo, 2013; Razo & Gao, 2013; Tseng et al., 2013). The SP surveys are attractive because participants only need to rank their preference of different facilities or route options, which can easily and quickly collect a large dataset without concerning inter-correlations among attributes for further analysis (Abraham et al., 2002). RP surveys are appealing because actual routes made by participants are recalled, which reveals preferences in the real world (Stinson & Bhat, 2003). SP surveys may, however, lose some significant attributes because participants are required to make choices among different facilities and route characteristics given by the choice set within a survey. Similarly, the participants of time-consuming RP surveys may not be able to recall accurately their routes. Some route choice studies used GPS devices to automatically record actual routes taken by participants, which maximizes accuracy, but can be costly in terms of equipping GPS devices and can present difficulties with respect to converting GPS data to actual routes within a road network (Broach et al., 2012; Hood et al., 2011; Papinski and Scott, 2013).

In route choice analysis, choice set generation is a common and essential step that creates alternative routes for observed routes between an origin-destination (O-D) pair. According to Prato (2009), there are four types of choice set generation methods: deterministic shortest path-based methods, stochastic shortest path-based techniques, constrained enumeration algorithms, and probabilistic approaches. K-shortest path is the simplest approach that provides multiple alternative routes by minimizing a generalized cost function. However, this technique is not effective due to the high probability of generating highly overlapping routes. Path labeling, the most widely used approach that can avoid the overlapping problem creates optimized routes based on different generalized costs for variables, known as labels. The labels, specified by researchers, can be different for each case study. Howard and Burns (2001), for example, generate alternative routes by extracting optimal paths between O-D pairs in terms of distance, directness, and safety. Prato and Bekhor (2006) applied the labeling approach by optimizing routes for distance, free-flow time, travel time, and delay that represents the congestion level; however, they found path labeling inadvisable because of high dependence on the definition of labels, which requires prior understanding about individual travel preferences. Broach et al. (2012) used an improved version of route labeling, where alternatives were created by maximizing individual criteria with distances controlled by multiple distance constrained values within a pre-specified range. For instance, alternative routes could be generated by maximizing the proportion of a route on bicycle facilities within a distance constraint. For bike share programs, cyclists can choose among multiple routes between an O-D hub pair. By focusing on hub pairs, the consideration of producing alternative routes to represent potential choices of cyclists using traditional choice set generation techniques may be unnecessary because multiple actual routes taken by BSP members between a pair can be treated as alternatives.

With respect to travel behavior modeling, the Multinomial Logit (MNL) and Nested Logit (NL) are perhaps the most commonly used discrete choice models, but neither of them is appropriate for route choice modeling (Prato, 2009). Specifically, NL is not suitable due to its assumption that each alternative route is solely assigned to one nest, but the actual routes may share links with others. The MNL is unable to account for similarities among alternatives, so it is highly likely to over-predict the probabilities of routes with overlapping links, which results in overwhelming estimated traffic volumes on these shared links (Ben-Akiva et al., 2012; Dalumpines & Scott, 2017). In order to solve the problem of overlapping links, many studies have applied the Path-Size Logit (PSL) model, which adds a correction factor, an attribute called the path size that measures the level of overlap, into the MNL formulation (Ben-Akiva & Bierlaire, 1999; Broach et al., 2012; Frejinger et al., 2009; Hood et al., 2013; Menghini et al., 2010; Ramming, 2001). The PSL is the most widely used model for route choice analysis because of the simple utility formulation and high efficiency of estimation using conventional software (Bekhor et al., 2006).

3.3 METHODS

3.3.1 GPS Data Collection and Post-processing

GPS data recording cyclists routes were collected from GPS devices equipped on the bicycles provided by SoBi Hamilton. The service area of the BSP, containing over 100 hubs with a designated distance of 300 to 600 meters between them is shown by black dots in Figure 3.1. The GPS device is always running and never turned off so that both SoBi operators and users can always track the location of each bicycle.

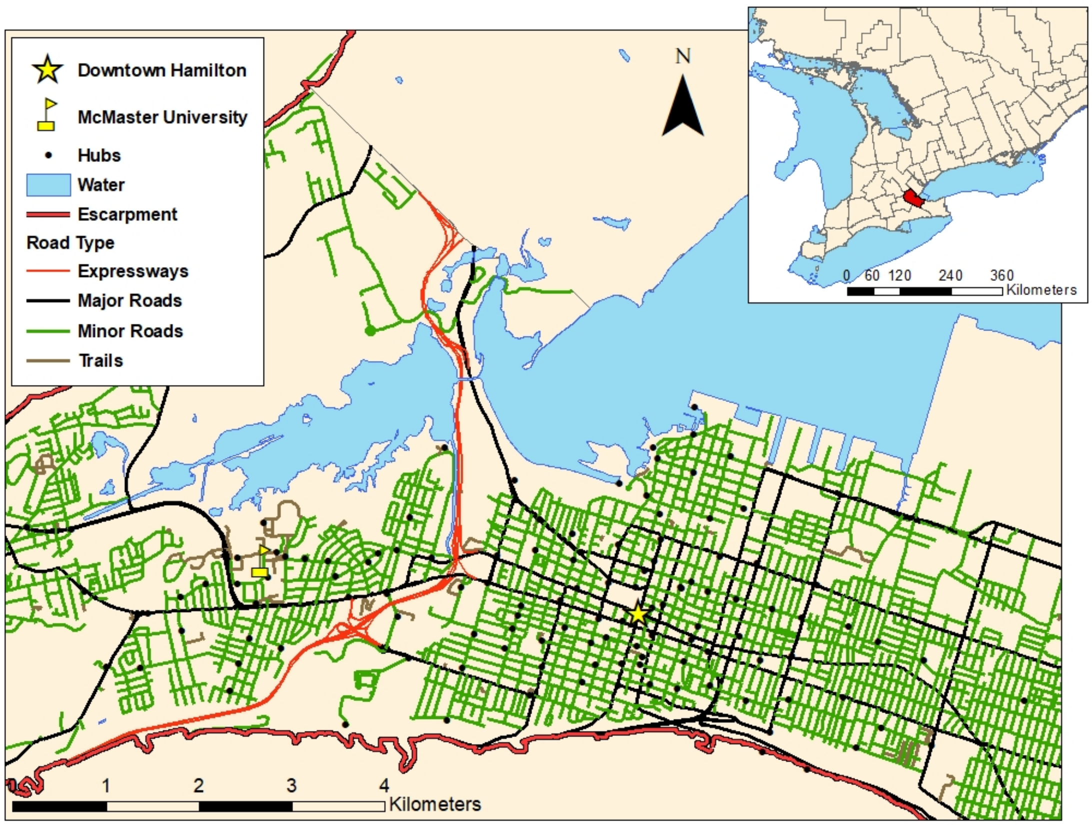

Figure 3.1: Study area and road network of SoBi Hamilton.

From April 1, 2015 to March 31, 2016, 161,426 GPS trajectories describing SoBi users observed routes were collected, where 132,396 cases travelled between hubs. In this study, only hub-to-hub trips were taken into account and were grouped by origin-destination (O-D) hub pairs. A GIS-based map-matching toolkit was used to match the GPS traces to the cycling dataset, and also to generate a series of attributes describing characteristics of the routes. The cycling network was created based on a bikeway dataset from the City of Hamilton (Open & Accessible Data, 2016) and a road line dataset containing all road types, such as major roads, minor roads, expressways, and trails, for Hamilton obtained from CanMap Content Suite (DMTI Spatial, 2015). As a result, a detailed cycling network composed of 22,710 links and 16,731 junctions with 6 road types and 10 categories of bikeways was created. However, only trails within McMasters campus, two road types including major and minor roads (Figure 3.1), and nine type of bikeways (Figure 3.2) were actually travelled by cyclists. A high-resolution digital elevation model (DEM) with a pixel size of 3030m was used to calculate the gradient gain and loss at about 30m increments along each road segment.

The actual cycling trips along the cycling network were generated by a GIS-based map-matching toolkit, which was developed using the Python scripting language. The toolkit generates a series of route attributes (Table 3.1), such as route distance, route directness, number of turns, mean distance between intersections, and length of longest leg (Dalumpines & Scott, 2011). The core principle of this toolkit is to match GPS points to the network using the shortest path algorithm and two basic inputs in ArcGIS (Network Analyst extension). One basic input is origin and destination points, and the other input is barriers that constrain the shortest path algorithm to create a map-matched route following the GPS trajectories representing an actual trip. In brief, the start and end points of a GPS trajectory are regarded as stops, and a 50 m buffer around the polyline converted by the GPS trajectory is created to identify the barriers. As a result, the shortest path algorithm generates the observed route, which is actually the shortest route between the origin and destination (stops) following the GPS points constrained by barriers.

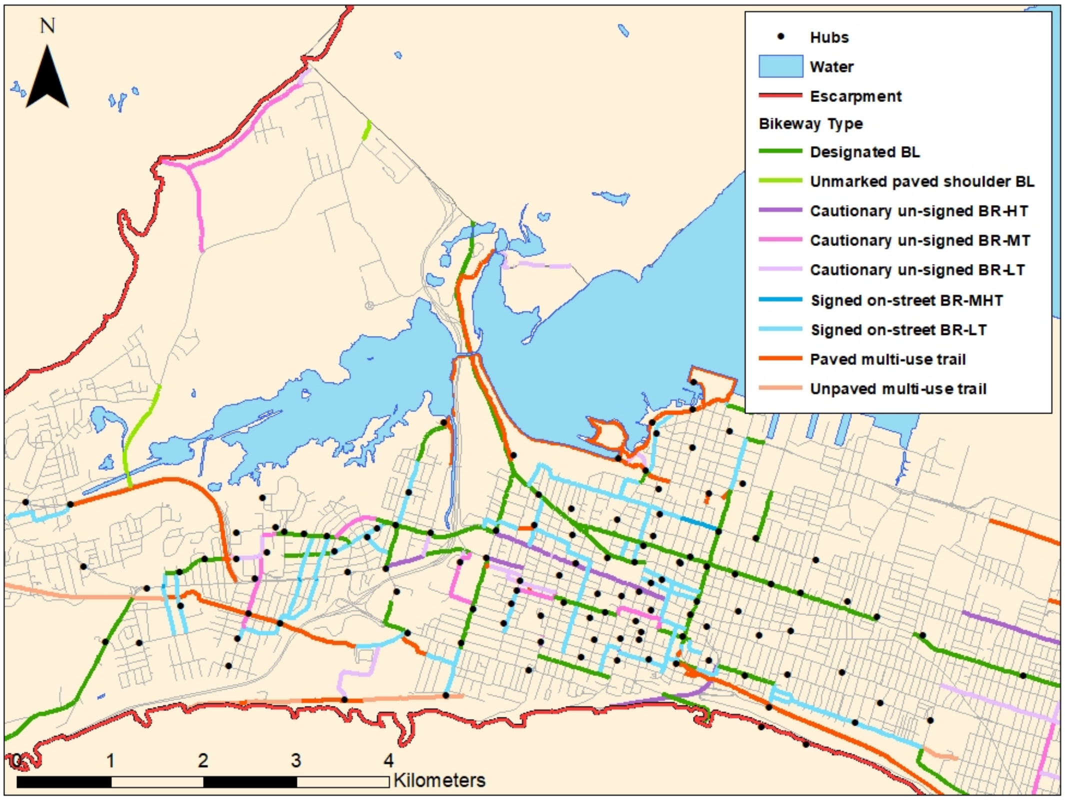

Figure 3.2: Bikeway distribution in study area.

3.3.2 Choice Set Generation

In the 12-month period SoBi dataset, 132,396 hub-to-hub trips were collected and map-matched; however, many of them follow identical routes. In order to extract unique routes between hub pairs from all the actual trips, core link signatures for trips are created to identify whether those trips travel along the same unique routes. Since each road segment has a link ID in the cycling network, all the links making up a trip are extracted, and their corresponding IDs are combined together in ascending order with colons separating each link ID. Repeated trips may have a small difference at the origin and destination points due to the instability of GPS devices or the delay in activating time, but they should be regarded as exactly the same. Thus, the list of ordered link IDs without the links at the origin and the destination is regarded as the core link signature. As a result, the trips with the same core link signature are considered to follow the same route, and the use frequency of each unique route is calculated based on the number of trips along it. Unlike other studies using choice set generation to create alternative routes for an observed route, this research regards all the actual unique routes taken by SoBi users as both observed and alternative routes between each hub pair, and the weight of each route refers to the number of trips on it. In this case, hub pairs with only one unique route between them are removed from the study; specifically, there are 7,962 hub pairs, but 20% of them are excluded because of only containing one unique route and 2% are excluded due to round trips. In general, the number of hub-to-hub unique routes ranges from 2 to 77, and on average each hub pair includes 14 unique routes.

3.3.3 Normalized Gini Coefficient

The Gini coefficient, normally used for measuring income inequality, is a well-known index to explore dispersion in a series of values. In route choice analysis, the Gini coefficient can identify whether there is a preference for a route between an origin and a destination. It varies between 0 (perfect equality) and $1 - 1/N$ (complete inequality), where N represents number of unique routes between a hub pair. In order to normalize the Gini coefficient by comparing the value of G for an O-D group to its number of unique routes N (>1), the formula of the normalized Gini coefficient is (Lima et al., 2016):

$$G_n = \frac{G}{1 - 1/N} \tag{3.1}$$

where a G_n value of 1 always indicates perfect inequality no matter how many unique routes exist within a group, suggesting perfect bias towards one route, while a G_n of 0 still refers to perfect equality, where all the O-D unique routes are evenly used by cyclists. Therefore, a value of G_n for a hub pair close to 0 indicates that routes are chosen more evenly, while the G_n close to 1 implies a route is more biased and the alternatives are less likely to be chosen.

3.3.4 Path-Size Logit Model

The multinomial logit (MNL) model is a method used to model the relationship between a polytomous dependent variable and a set of independent variables (Ben-Akiva & Lerman, 1985). In other words, MNL can predict the probabilities for alternative outcomes of a polytomous response variable given a series of explanatory variables. However, for route choice modeling containing overlapping routes, MNL is inappropriate as it cannot account for similarities among route alternatives (Bekhor et al., 2006; Prato, 2009). To correct for the overlapping routes, the Path-Size Logit (PSL) model, easily adding a correction factor (i.e., path size attribute) to the path utility under the structure of the MNL model, is suitable for studies exploring the factors influencing route choice behavior (Bekhor et al., 2006). There are multiple formulations to generate the path size attribute. Frejinger & Bierlaire (2007) stated that the original PS formulation (Ben-Akiva and Bierlaire, 1999) with a theoretical support could provide intuitive results. Thus, the original formulation was chosen in this study. PS_{in}, representing the path size (PS) attribute of path i taken by individual n, for the j alternatives in choice set C_n is defined as:

$$PS_{in} = \sum_{a \in \Gamma_i} \frac{L_a}{L_i} \frac{1}{\sum_{j \in C_n} \delta_{aj}} \tag{3.2}$$

where Γ_i is all the links of path i, L_a is the length of link a, L_i is the length of path i, and δ_{aj} is the link-route incidence dummy that equals 1 if link a is a part of j and 0 otherwise. The path size always has values less than or equal to 1 (Prato & Bekhor, 2007). That is, a sole route without overlapping links in a choice set is given a PS of 1, while a route containing partial overlaps is given a PS less than 1. The final formula of the PSL model is:

$$P\gamma(i|C_n) = \frac{e^{V_{in} + InPS_{in}}}{\sum_{j \in C_n} e^{V_{jn} + InPS_{jn}}} \tag{3.3}$$

where $P\gamma(i|C_n)$ is the probability of choosing alternative route i given the choice set of alternative routes C_n for observation n, V_{in} and V_{jn} are respectively the deterministic utilities of routes i and j. There is no adjustment to deterministic utility for a sole route with ln(PS) = 0, while the alternative routes in the choice set containing overlapping links with ln(PS) < 0 lead to the decreased deterministic utility (Dalumpines & Scott, 2017).

3.4 RESULTS

There were 132,396 hub-to-hub trips taken by its SoBi Hamilton users between April 1, 2015 and March 31, 2016, where most trips were taken along the same unique routes. From these trips, 41,318 hub-to-hub unique routes were extracted, which constitutes the global dataset, where for each hub pair, the number of alternative routes ranges from 2 to 77 with an average of 14. Globally, a high normalized Gini coefficient with a median value of 0.5 for hub pairs is found, implying a tendency of a dominant route between a hub pair. Figure 3.3 displays two examples respectively referring to all the unique routes between a hub pair with a high G_n of 0.91 (a), where 75% trips are travelling along the dominant route, and the other with a low G_n of 0.19 (b), where all the unique routes are almost evenly used. With regard to hub pairs with a mild preference towards a route among all the route choices ($0.5 \leq G_n < 0.8$), only 1257 of the total hub pairs meet the specification and are included in the medium G_n dataset with 12,902 unique routes, and the number of alternative routes for each hub pair varies from 2 to 66 with an average of 17. Concerning the high preference dataset ($G_n \geq 0.8$) that contains 2472 routes derived from 250 hub pairs with highly biased dominant routes, the number of alternative routes ranges from 2 to 73 with a mean value of 14 for each hub pair.

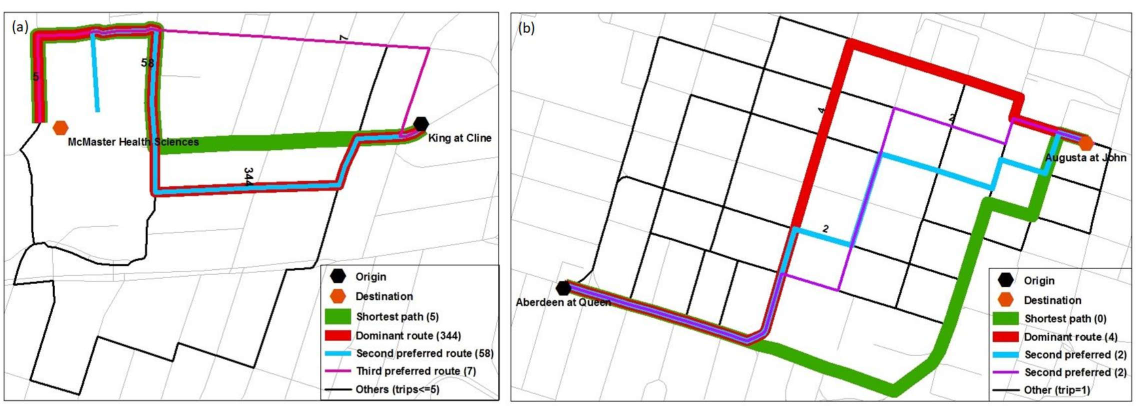

Figure 3.3: (a) Demonstration of a hub pair with a high G_n. (b) Demonstration of a hub pair with a low G_n.

3.4.1 Descriptive Statistics

Table 3.1 lists summary statistics, including the mean value and standard deviation of each route attribute, respectively for the global, medium G_n, and high G_n datasets. As expected, the average route distance of the high G_n dataset is shorter than the average distance of routes from the medium G_n dataset, followed by that of routes from the global dataset. Generally, around 99% of routes in all three datasets are no more than 6 km, and their route distance distribution are quite similar (Figure 3.4). Specifically, most of routes taken by cyclists from all the datasets are between 1 and 2 km. Compared to the global dataset, many more routes from the other two datasets are within 2 km, while the global dataset routes tend to dominate over the other two datasets from 2 km onward.

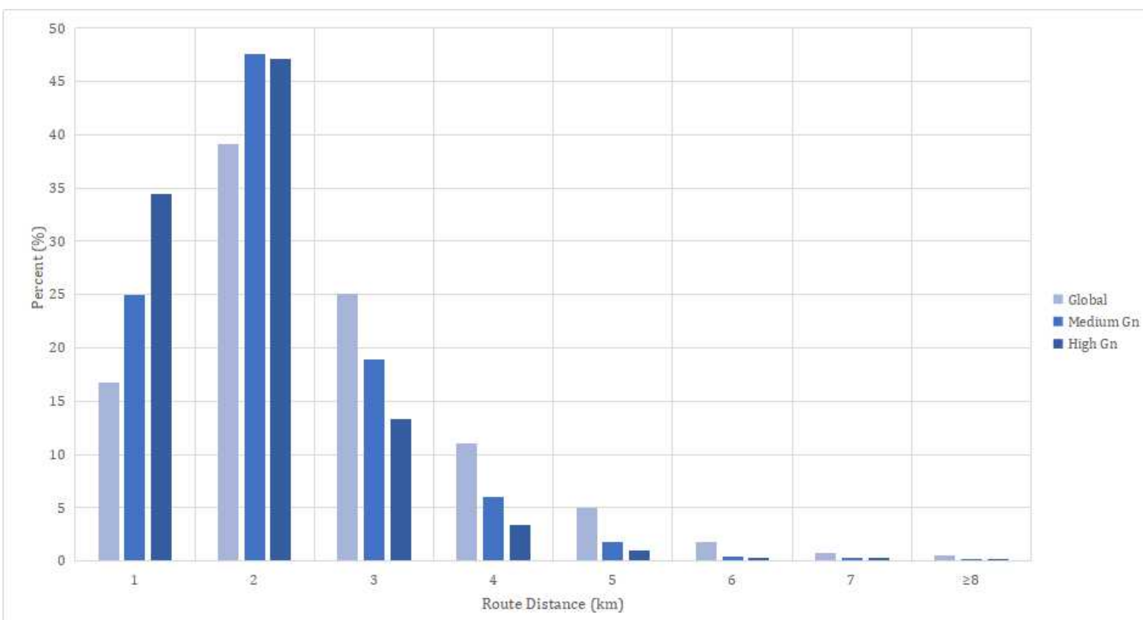

Figure 3.4: Route distance pattern for routes from three datasets: global, medium G_n, and high G_n.

In general, the mean values of route attributes for the medium G_n dataset are usually between those for the global and high G_n dataset (Table 3.1). The average route directness index (RDI), showing the efficiency of a route, is lowest in the high G_n dataset, which indicates that routes from the high G_n dataset are the least circuitous on average, followed by the medium G_n and the global datasets. Similarly, the number of road segments, total turns, and intersections confirm that routes in the high preference dataset are more direct than those in the other two datasets. Interestingly, the average route distances between intersections are almost the same for all three datasets. Compared to the medium G_n and global datasets, the high preference dataset contains routes with a larger proportion along flat roads with slope of 0-2%, trails, designated bike lanes, and cautionary un-signed bike routes on streets with moderate and low traffic volume on average. Given the average proportion of a route along some bikeways less than 0.3%, routes from all three datasets are almost unlikely to be taken on signed on-street bike routes with moderate to high traffic volume, paved multi-use trails shared with pedestrians, unpaved multi-use trails, and unmarked paved shoulder bike lanes. In addition, the average path size values suggest that routes in the high G_n dataset have more overlaps than those in the medium G_n and global datasets.

Table 3.1: Descriptive statistics of route attributes for global, medium G_n, and high G_n datasets.

Variable	Description	Global	Medium G_n	High G_n
		mean ± std.	mean ± std.	mean ± std.
Route distance	Distance of route in kilometers	2.11 ± 1.27	1.66 ± 0.99	1.43 ± 0.92
Straight-line distance	Straight-line distance between origin and destination in kilometers	1.5 ± 0.94	1.18 ± 0.71	1.01 ± 0.63
RDI	Route directness index (compare to straight line)	1.65 ± 13.69	1.56 ± 2.33	1.50 ± 0.81
Path size	A correction factor to path utility	0.38 ± 0.23	0.31 ± 0.22	0.28 ± 0.21
Unique road segments	Number of unique road segments that a route travel along	6.22 ± 3.16	5.32 ± 2.60	4.37 ± 2.54
Left turn	Number of left turns	2.46 ± 1.71	2.05 ± 1.51	1.74 ± 1.52
Right turn	Number of right turns	2.48 ± 1.70	2.09 ± 1.51	1.69 ± 1.45
Total turns	Total number of turns	5.20 ± 3.23	4.40 ± 2.74	3.66 ± 2.76
Intersections	Number of intersections	23.13 ± 15.11	18.06 ± 11.84	15.44 ± 10.66
Distance btw intersections	Average route distance between intersections in kilometers	0.1 ± 0.02	0.10 ± 0.02	0.10 ± 0.03
Longest leg length	Length of longest leg in kilometers	0.91 ± 0.58	0.78 ± 0.47	0.75 ± 0.44
Prop. Slope 0-2% (%)	Proportion of route along road segments with slope of 0-2%	65.98 ± 24.36	68.92 ± 25.09	72.81 ± 25.59
Prop. Slope 2-4% (%)	Proportion of route along road segments with slope of 2-4%	20.87 ± 14.48	19.59 ± 15.27	17.39 ± 14.57
Prop. Slope 4-6% (%)	Proportion of route along road segments with slope of 4-6%	5.80 ± 7.16	5.01 ± 7.39	3.61 ± 6.33
Prop. Slope 6%+	Proportion of route along road segments with slope over 6 %	7.35 ± 11.02	6.48 ± 11.42	6.19 ± 13.74
Major road (%)	Proportion of route along major roads	29.48 ± 27.27	26.66 ± 28.56	22.15 ± 30.44
-Major roads without BL (%)	Proportion of route along major roads without designated bike lanes	22.63 ± 24.37	20.41 ± 25.59	15.77 ± 25.29
-Major roads with BL (%)	Proportion of route along major roads with designated bike lanes	6.85 ± 13.43	6.26 ± 14.07	6.38 ± 16.83
Minor road (%)	Proportion of route along local roads	68.34 ± 27.01	70.07 ± 28.38	69.35 ± 30.20
-Minor roads without BL (%)	Proportion of route along minor roads without designated bike lanes	52.09 ± 27.31	53.40 ± 28.48	48.25 ± 30.86
-Minor roads with BL (%)	Proportion of route along minor roads with designated bike lanes	16.25 ± 20.31	16.67 ± 21.70	21.10 ± 27.41
Trail (%)	Proportion of route along trails	2.18 ± 7.72	3.27 ± 9.55	8.50 ± 15.39
Designated BL (%)	Proportion of route on designated bike lanes	23.30 ± 23.26	23.20 ± 24.80	28.31 ± 30.15
Cautionary un-signed BR-HT (%)	% of route on cautionary un-signed bike routes on street with high traffic volume	2.70 ± 8.47	1.93 ± 7.40	0.88 ± 4.72
Cautionary un-signed BR-MT (%)	% of route on cautionary un-signed bike routes on street with moderate traffic volume	3.47 ± 8.86	4.05 ± 9.98	6.99 ± 14.00
Cautionary un-signed BR-LT (%)	% of route on cautionary un-signed bike routes on street with low traffic volume	1.20 ± 5.28	1.38 ± 5.91	4.12 ± 11.19
Signed on-street BR-MHT (%)	% of route on signed on-street bike routes with moderate to high traffic volume	0.29 ± 2.21	0.23 ± 2.16	0.23 ± 2.12
Signed on-street BR-LT (%)	% of route on signed on-street bike routes with mostly low traffic volume	10.72 ± 14.28	11.05 ± 15.63	10.05 ± 16.89
Paved multi-use trail (%)	% of route on paved multi-use trails (shared with pedestrians)	0.23 ± 1.72	0.19 ± 1.55	0.11 ± 0.82
Unpaved multi-use trail (%)	% of route on unpaved (softer surface) multi-use trails (shared with pedestrians)	0.0001 ± 0.02	0.0003 ± 0.03	0
Unmarked paved shoulder BL (%)	% of route on unmarked paved shoulder bike lanes	0.0003 ± 0.03	0.0004 ± 0.04	0

3.4.2 Contribution of Route Attributes

Before modeling, multi-collinearity should be examined to make sure that no correlation exists among independent variables because collinearity makes the coefficients unstable. Given the correlation coefficients between independent variables, it can be found that the straight-line distance, the number of intersections, and the total route distance are positively correlated with each other with indices larger than 0.8. Identically, the number of road segments, left turns, right turns, and total turns are highly correlated with each other. Also, the proportion of a route along major roads without designated bike lanes (BLs) is strongly negatively correlated to that along minor roads without designated BLs. Simultaneously, proportion of route along a minor road with designated BLs has a very high positive correlation with that along the overall designated BLs because most designated BLs are on the minor roads in the study area. Therefore, the straight-line distance, the proportion of a route along minor roads without designated BLs, the overall designated BLs, the number of intersections, road segments, left turns, and right turns are removed from the final models due to multi-collinearity. Additionally, the proportion of a route along slopes over 6% is also excluded from the model because it is mutually complementary with the sum of the other three groups of slopes, and around 93% of routes are on road segments with slopes lower than 6%. Furthermore, proportions of a route along unpaved multi-use trails and unmarked paved shoulder BL are removed from the final model because such bicycle facilities are at the edge of SoBi service area and few trips occur there ($< 0.0005\%$ of a route is along these two facilities).

For all three models estimated from the aforementioned datasets, the probability of choosing a route is negatively correlated to route distance if there are no other route attributes, indicating that cyclists tend to choose shorter routes without considering other factors. However, the estimation results of the PSL models displayed in Table 3.2 show that route distance has a positive effect, implying a detour for other significant attributes. The estimates for all three models suggest that cyclists prefer less circuitous routes given a negative sign of RDI and a positive sign of the longest leg length. Similarly, the significant negative determinant, the total number of turns, and the significant positive determinant, the average distance between intersections, in the global and high G_n models tell the same story because turns and intersections may make routes more circuitous. In terms of elevation, it has no significant influence on the medium G_n model, but both global and high G_n models find significant negative effects for 0-2% and 4-6% slopes, and it is much less likely to travel on 4-6% slopes than on 0-2% slopes given the values of the coefficients. Interestingly, in the high G_n dataset containing hub pairs with a high preference towards dominant routes, the slope of 2-4% has a significant positive contribution. In this case, avoiding flat areas may be a tradeoff for other route attributes.

With regards to road types and bicycle facility types, all three models show a preference for trails that are mainly within McMasters campus, followed by cautionary un-signed bike routes on streets with low traffic volumes, and a tendency of avoiding those un-signed bike routes on streets with high traffic volumes and paved multi-use trails that are a small subset requiring a significant detour. At the same time, in the global model, cautionary un-signed bike routes on street with moderate traffic volume is a significant positive variable, while it is a negative determinant for the high preference dataset. In addition, both the global and medium G_n models find a preference for signed on-street bike routes with mostly low traffic volumes while the high G_n model finds it insignificant.

Table 3.2: PSL model estimation results for global, medium G_n, and high G_n datasets.

Attributes	Global		Medium G_n		High G_n	
	coeff.	*t*-stat	coeff.	*t*-stat	coeff.	*t*-stat
Route distance	0.21	**19.71**	0.72	**28.86**	2.48	**49.64**
RDI	-0.01	**-3.06**	-0.46	**-27.94**	-2.25	**-48.81**
Total turns	-0.01	**-4.48**	0.0003	0.08	-0.07	**-10.35**
Distance btw intersections	1.37	**8.43**	-0.11	-0.40	1.16	**5.08**
Longest leg length	0.65	**39.23**	0.87	**31.82**	2.12	**41.29**
Prop. Slope 0-2%	-0.21	**-3.37**	-0.14	-1.59	-0.53	**-3.12**
Prop. Slope 2-4%	-0.02	-0.30	0.02	0.15	0.62	**3.17**
Prop. Slope 4-6%	-0.48	**-3.88**	0.10	0.55	-5.96	**-13.84**
Major roads without BL	0.13	**5.03**	0.08	**2.22**	-1.44	**-16.43**
Major roads with BL	0.27	**5.57**	0.34	**4.76**	0.25	1.81
Minor roads with BL	-0.08	**-2.85**	0.03	0.76	-2.47	**-28.79**
Trails	1.41	**21.84**	1.30	**12.49**	1.61	**12.97**
Cautionary un-signed BR-HT	-0.66	**-9.19**	-0.46	**-3.94**	-1.69	**-4.67**
Cautionary un-signed BR-MT	0.32	**7.35**	-0.01	-0.21	-0.22	**-2.37**
Cautionary un-signed BR-LT	1.14	**16.16**	0.94	**7.14**	0.94	**7.29**
Signed on-street BR-MHT	-0.06	-0.28	0.04	0.09	0.18	0.23
Signed on-street BR-LT	0.19	**6.21**	0.16	**3.48**	0.12	1.44
Paved multi-use trail	-1.39	**-4.04**	-1.30	**-2.41**	-13.98	**-7.08**
ln(Path Size)	-1.34	**-140.93**	-1.50	**-104.70**	-2.68	**-96.37**
Number of observations		41318		12902		2472
Pseudo R^2		0.054		0.069		0.195

Bolded values imply attributes significant at the 5% level.

As expected, without considering other attributes, SoBi cyclists tend to pedal along major roads with or without designated BLs, and minor roads with designated BLs possibly because the most highly demanded hubs are along major roads (Figure 3.5). Surprisingly, minor roads with designated BLs change to a significant negative factor in the global model, and for the high G_n model, the coefficients of both major roads without designated BLs and minor roads with designated BLs switch to negative after including ln (Path Size) in these models probably because most route overlaps occur upon designated BLs, especially designated BLs on minor roads. As shown in

Figure 3.6, a hub pair with a high G_n of 0.86 from Cannon at Sherman to York at Macnab contains 9 unique routes, and the majority of overlaps lies on Cannon Street East that is classified as a minor road with designated BLs, which may cause an interaction with the route path size. In Table 3.2, the ln (Path Size) estimates for all three models are negative, which indicates that SoBi users have tendency to choose hub-to-hub routes sharing links with other alternatives, especially between hub pairs with a high preference for dominant routes. Concerning the model performance, the high G_n model performs the best, followed by the medium G_n model and the global model, which is reasonable because the high G_n model examines hub pairs with a clear bias rather than the other two datasets where some alternatives may mitigate the effects of some significant attributes.

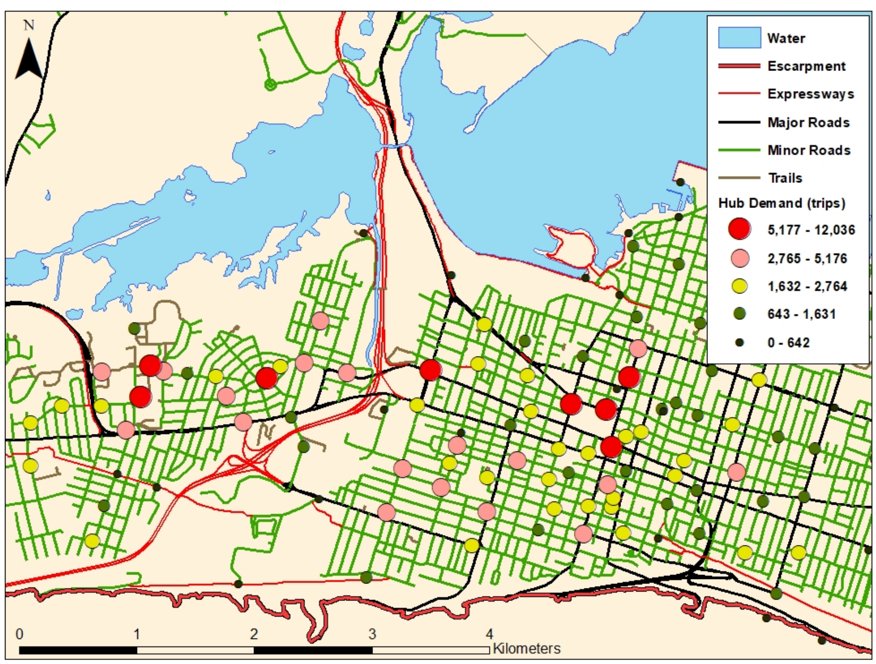

Figure 3.5: Demand pattern of SoBi hubs.

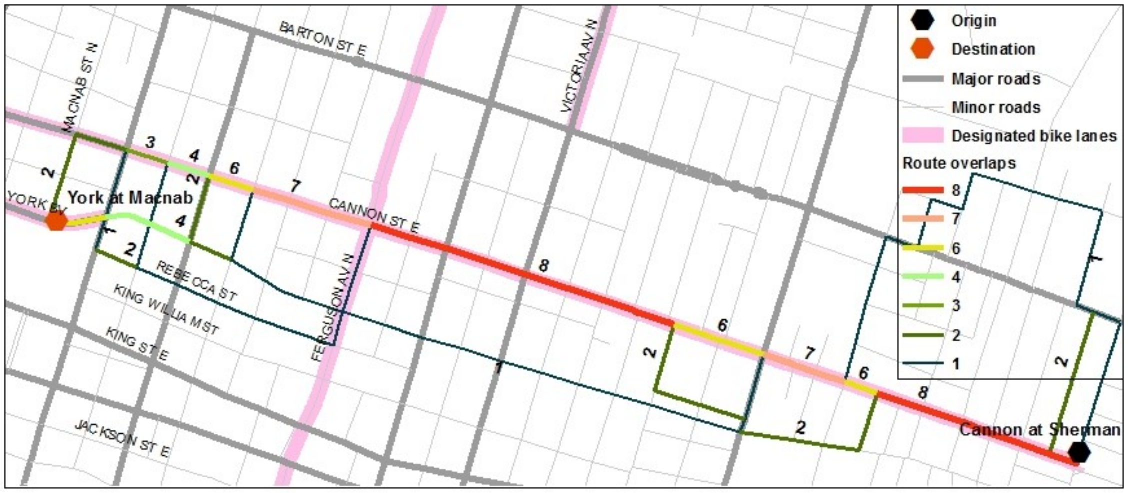

Figure 3.6: Demonstration of routes between a hub pair mainly along Cannon Street.

3.5 CONCLUSIONS

Like some other route choice analysis studies (Broach et al., 2012; Hood et al., 2011; Menghini et al., 2010), this paper used data from GPS devices to accurately track the actual routes taken by participants. However, it did not make use of traditional choice set generation methods, such as path labeling and K-shortest path, to create alternative routes because a choice set consisting of actual cyclists routes can be generated for a bike share program (BSP). In detail, multiple unique routes between each hub pair are extracted from repeated trips generated by the GIS-based map-matching toolkit, and the weight of each unique route was identified by the number of trips on it. In terms of route choice modeling, similar to previous studies, PSL models are found to improve the performance of MNL models tested in this paper (Bekhor et al., 2006; Frejinger & Bierlaire, 2007).

Generally, cyclists are willing to slightly endure more time to use bicycle facilities (Broach et al., 2012; Tilahun et al. 2007; Winters et al., 2010; Winters & Teschke, 2010), such as designated bike lanes, trails off street, separated bike paths, and residential streets with traffic calming, and to avoid some other route attributes, involving stop signs, red lights, major cross streets (Sener et al., 2009; Stinson & Bhat, 2003), turn frequency (Aultman-Hall et al., 1997; Hood et al., 2011), steep slopes and high traffic volumes (Broach et al., 2012; Hood et al., 2011). The contribution of most route attributes to SoBi cyclists route choice behavior is similar to the aforementioned studies; specifically, the BSP users also slightly detour on less circuitous routes to avoid turns, changing roads, steep slopes (i.e., slopes over 4%), and high traffic volumes. Concerning off-street trails, the PSL models illustrated a strong preference for them; however, usage occurs on McMaster Universitys campus as travelling on campus has few additional choices. In terms of other road types and bicycle facilities, the PSL models illustrated a preference for cautionary un-signed bike routes on streets with low traffic volumes, followed by designated BLs, and signed on-street bike routes with mostly low traffic volumes, and an avoidance of cautionary un-signed bike routes on streets with high traffic volumes and paved multi-use trails far away from main roads. The major differences relied on the major roads without designated BLs and minor roads with designated BLs; the global model found a bias on major roads without designated BLs and a tiny negative effect of minor roads with designated BLs, while the high preference model revealed an avoidance for both of them probably because of their correlation with the path size attribute. Although a positive correction factor for link overlaps is expected, all three models in this paper find that ln (Path Size) has a significant negative contribution, which is similar to

Dalumpines and Scott (2017) illustrating that ln (Path Size) is regarded as a utility probably containing hidden attractive features from overlapping links instead of a correction factor. The negative ln (Path Size) estimate is reasonable in this study as actual routes make up the choice set, so cyclists prefer overlapping links, which may contain some advantages, such as covering bicycle facilities and high connectivity to other routes.

These findings can help transportation professionals and urban planners understand cyclists preference for route choice. In order to promote cycling, the findings suggest that bicycle infrastructure should be added on the major roads and roads with low traffic volumes, but also other factors should be considered (e.g., roads with steep slopes over 4% should be avoided; paved multi-use trails should be built in more accessible areas). Also, negative ln (Path Size) estimates suggest that future studies can further explore the highly overlapped links to identify the hidden features that attract BSP cyclists. In addition, this paper proposes a new way to establish a BSP choice set by extracting actual hub-to-hub routes from repeated trips taken by cyclists, which provides a new alternative route generation technique for BSP route choice analysis. Furthermore, these outcomes can be used to build a new function within the SoBi mobile application to provide recommended routes for cyclists given the effects of significant determinants. The main limitation of this paper is lacking some route attributes that may be significant, such as stop signs, red lights, and pavement quality, as well as some demographic data, such as age, gender, income, and cycling experience, which can make the analysis of BSP cyclists route choice behavior more comprehensive.

REFERENCES

Abraham, J. E., McMillan, S., Brownlee, A. T., & Hunt, J. D. (2002). Investigation of cycling sensitivities. Presented at the 81st Annual Meeting of the Transportation Research Board, Washington, D.C.https://www.researchgate.net/publication/228706636_Investigation_of_Cycling_Sensitivities (Accessed August 24, 2017).

Aultman-Hall, L., Hall, F., & Baetz, B. (1997). Analysis of bicycle commuter routes using geographic information systems: implications for bicycle planning. *Transportation Research Record: Journal of the Transportation Research Board*, 1578: 102-110.

Bapat, S. V., Zala, D. R. L. B., & Amin, A. A. (2017). Route choice analysis: a case study of NH 8 and NE 1. *International Journal of Recent Trends in Engineering & Research*, 3: 216-227.

Bekhor, S., Ben-Akiva, M. E., & Ramming, M. S. (2006). Evaluation of choice set generation algorithms for route choice models. *Annals of Operations Research*, 144: 235-247.

Ben-Akiva, M., & Bierlaire, M. (1999). Discrete choice methods and their applications to short term travel decisions. *Handbook of Transportation Science*, 23: 5-33.

Ben-Akiva, M. E., Gao, S., Wei, Z., & Wen, Y. (2012). A dynamic traffic assignment model for highly congested urban networks. *Transportation Research Part C: Emerging Technologies*, 24: 62-82.

Ben-Akiva, M. E., & Lerman, S. R. (1985). Discrete choice analysis: theory and application to travel demand. *Estimation of Switching Models from Revealed Preferences and Stated Intentions*, 9: 7-58.

Ben-Akiva, M. E., Ramming, M. S., & Bekhor, S. (2004). Route choice models. *Human Behaviour and Traffic Networks* (pp. 23-45). Springer, Berlin, Heidelberg.

Broach, J., Dill, J., & Gliebe, J. (2012). Where do cyclists ride? A route choice model developed with revealed preference GPS data. *Transportation Research Part A: Policy and Practice*, 46: 1730-1740.

Dalumpines, R., & Scott, D. M. (2011). GIS-based map-matching: Development and demonstration of a postprocessing map-matching algorithm for transportation research. Cartwright, W., Gartner, g., Meng, L., & Peterson M. P. (Eds.), *Advancing Geoinformation Science for A Changing World* (pp. 101-120). Springer, Berlin, Heidelberg.

Dalumpines, R., & Scott, D. M. (2017). Determinants of route choice behavior: A comparison of shop versus work trips using the Potential Path Area-Gateway (PPAG) algorithm and Path-Size Logit. *Journal of Transport Geography*, 59: 59-68.

Davis, L. S. (2014). Rolling along the last mile: Bike-sharing programs blossom nationwide. *Planning*, 80: 10-16.

De Nazelle, A., Nieuwenhuijsen, M. J., Ant, J. M., Brauer, M., Briggs, D., Braun-Fahrlander, C., ... Hoek, G. (2011). Improving health through policies that promote active travel: a review of evidence to support integrated health impact assessment. *Environment International*, 37: 766-777.

Dill, J. (2009). Bicycling for transportation and health: the role of infrastructure. *Journal of Public Health Policy*, 30: S95-S110.

DMTI Spatial. (2015). CanMap Content Suite. https://www.dmtispatial.com/canmap/ (Access August 24, 2017).

Fishman, E. (2016). Bikeshare: A review of recent literature. *Transport Reviews*, 36: 92-113.

Frank, L. D., Sallis, J. F., Conway, T. L., Chapman, J. E., Saelens, B. E., & Bachman, W. (2006). Many pathways from land use to health: associations between neighborhood walkability and active transportation, body mass index, and air quality. *Journal of the American Planning Association*, 72: 75-87.

Frejinger, E., & Bierlaire, M. (2007). Capturing correlation with subnetworks in route choice models. *Transportation Research Part B: Methodological*, 41: 363-378.

Frejinger, E., Bierlaire, M., & Ben-Akiva, M. (2009). Sampling of alternatives for route choice modeling. *Transportation Research Part B: Methodological*, 43: 984-994.

Gordon-Larsen, P., Boone-Heinonen, J., Sidney, S., Sternfeld, B., Jacobs, D. R., & Lewis, C. E. (2009). Active commuting and cardiovascular disease risk: the CARDIA study. *Archives of Internal Medicine*, 169: 1216-1223.

Guo, Z., & Loo, B. P. (2013). Pedestrian environment and route choice: evidence from New York City and Hong Kong. *Journal of Transport Geography*, 28: 124-136.

Guttenplan, M., & Patten, R. (1995). Off-road but on track: using bicycle and pedestrian trails for transportation. *TR News*, 178: 7-11.

Hood, J., Sall, E., & Charlton, B. (2011). A GPS-based bicycle route choice model for San Francisco, California. *Transportation Letters*, 3: 63-75.

Howard, C., & Burns, E. (2001). Cycling to work in Phoenix: route choice, travel behavior, and commuter characteristics. *Transportation Research Record: Journal of the Transportation Research Board*, 1773: 39-46.

Hunt, J. D., & Abraham, J. E. (2007). Influences on bicycle use. *Transportation*, 34: 453-470.

Kisner, C. (2011). Integrating bike share programs into a sustainable transportation system. *Center for Research and Innovation*, National League of Cities, Washington, D.C.

Lima, A., Stanojevic, R., Papagiannaki, D., Rodriguez, P., & Gonzlez, M. C. (2016). Understanding individual routing behaviour. *Journal of the Royal Society Interface*, 13: 20160021.

Menghini, G., Carrasco, N., Schssler, N., & Axhausen, K. W. (2010). Route choice of cyclists in Zurich. *Transportation Research Part A: Policy and Practice*, 44: 754-765.

Open & Accessible Data. (2016). Bikeways. https://www.hamilton.ca/city-initiatives/strategies-actions/open-accessible-data (Accessed August 24, 2017).

Papinski, D., & Scott, D. M. (2013). Route choice efficiency: an investigation of home-to-work trips using GPS data. *Environment and Planning A*, 45: 263-275.

Prato, C., & Bekhor, S. (2006). Applying branch-and-bound technique to route choice set generation. *Transportation Research Record: Journal of the Transportation Research Board*, 1985: 19-28.

Prato, C., & Bekhor, S. (2007). Modeling route choice behavior: How relevant is the composition of choice set? *Transportation Research Record: Journal of the Transportation Research Board*, 2003: 64-73.

Prato, C. G. (2009). Route choice modeling: past, present and future research directions. *Journal of Choice Modelling*, 2: 65-100.

Razo, M., & Gao, S. (2013). A rank-dependent expected utility model for strategic route choice with stated preference data. *Transportation Research Part C: Emerging Technologies*, 27: 117-130.

Ramming, M. S. (2001). Network knowledge and route choice. *Unpublished Ph. D. Thesis, Massachusetts Institute of Technology.*

Sener, I. N., Eluru, N., & Bhat, C. R. (2009). An analysis of bicycle route choice preferences in Texas, US. *Transportation*, 36: 511-539.

Shaheen, S. A. (2012). Public bikesharing in North America: early operator and user understanding. *Mineta Transportation Institute*, Report 11-19, San Jose, CA.

So, Y., & Kuhfeld, W. F. (1995). Multinomial logit models. *SUGI 20 Conference Proceedings* (pp. 1227-1234).

Stinson, M., & Bhat, C. (2003). Commuter bicyclist route choice: Analysis using a stated preference survey. *Transportation Research Record: Journal of the Transportation Research Board*, 1828: 107-115.

Tilahun, N. Y., Levinson, D. M., & Krizek, K. J. (2007). Trails, lanes, or traffic: Valuing bicycle facilities with an adaptive stated preference survey. *Transportation Research Part A: Policy and Practice*, 41: 287-301.

Tseng, Y. Y., Knockaert, J., & Verhoef, E. T. (2013). A revealed-preference study of behavioral impacts of real-time traffic information. *Transportation Research Part C: Emerging Technologies*, 30: 196-209.

Winters, M., & Teschke, K. (2010). Route preferences among adults in the near market for bicycling: findings of the Cycling in Cities study. *American Journal of Health Promotion*, 25: 40-47.

Winters, M., Teschke, K., Grant, M., Setton, E., & Brauer, M. (2010). How far out of the way will we travel? Built environment influences on route selection for bicycle and car travel. *Transportation Research Record: Journal of the Transportation Research Board*, 2190: 1-10.

Chapter 4

CONCLUSIONS

4.1 INTRODUCTION

This thesis has used a combination of descriptive statistics and path-size logit models in order to explore and analyze the spatial distribution of cyclists dominant routes and determinants of bike share cyclist route choice behavior. The following four objectives have been addressed:

- Gnerate cyclists actual trips from hub to hub using GPS trajectories and extract unique routes to constitute a choice set for each hub pair.
- Dtermine whether a dominant route exists between a hub pair according to normalized Gini coefficients and explore the spatial distribution of bike share cyclists dominant routes.
- Compare dominant routes with shortest paths based on distance and examine differences.

- Identify determinants that contribute to the SoBi cyclist route choice decision making process in terms of route attributes based on path-size logit models.

This chapter summarizes major findings and contributions of this thesis with respect to the aforementioned objectives. Then, the limitations of this research are discussed, followed by conclusions and recommendations for future research.

4.2 SUMMARY OF FINDINGS AND RESEARCH CONTRIBUTIONS

This thesis has explored the spatial distribution of bike share cyclists dominant routes and identified factors accounting for their route choice decisions. Chapter 2 confirmed the existence of dominant routes between hub pairs, examined their spatial patterns, and compared them with shortest paths based on distance. Chapter 3 further introduced Path-Size Logit (PSL) modeling to create models reflecting the influence of each factor on SoBi cyclist route choice behavior. In general, the first three objectives are addressed by Chapter 2, while the last objective is addressed by Chapter 3.

Findings from both Chapters 2 and 3 are similar to most previous studies associated with the correlation between route attributes and cyclists route choices. First, cyclists are willing to slightly detour to use bicycle facilities (Broach et al., 2012; Tilahun et al. 2007; Winters et al., 2010; Winters & Teschke, 2010). Next, although detouring may come with more turns and intersections, cyclists still try to avoid both of them (Aultman-Hall et al., 1997; Hood et al., 2011; Sener et al., 2009; Stinson & Bhat, 2003). In addition, both steep slopes and high traffic volumes are found as negative features for cyclists route choice decisions (Broach et al., 2012; Hood et al., 2011). More details of findings and contributions of Chapters 2 and 3 are respectively discussed below.

Similar to Lima et al. (2016) who found that drivers tend to have a dominant route among all the routes between an origin and a destination, Chapter 2 demonstrates that bike share cyclists also have a preferred route between a hub pair according to the value of normalized Gini coefficients (G_n). A G_n closer to 1 suggests a stronger bias toward one route, which can be considered as a dominant route; while a G_n closer to 0 indicates an even use of all the routes between a hub pair. To the authors knowledge, this is the first instance of extracting unique routes from all the hub-to-hub trips taken by multiple cyclists and introducing the idea of dominant routes in the bike share context. Then, around 93% of dominant routes are found not to be shortest paths based on distance. That is, shortest distance routes are not the optimal route choices for bike share cyclists in the real world. Specifically, Chapter 2 finds that cyclists tend to endure 10% more distance than shortest distance routes on average to use bicycle facilities without high traffic volumes, which may result in a slight increase in turns and intersections. In addition, Chapter 2 also visually displays the spatial pattern of dominant routes and proposes candidates of road segments suitable for additional bicycle facilities (Figure 2.5).

Chapter 3 introduced three PSL models based on three datasets with different G_n levels. In other words, the global model estimates all the hub-to-hub routes; the medium G_n model estimates routes belonging to hub pairs with G_n at least 0.5 and smaller than 0.8; while the high G_n model estimates routes between hub pairs with G_n no less than 0.8. These three models imply a similar tradeoff between the route distance and other route attributes. Bike share cyclists are willing to detour on not too circuitous routes in order to avoid turns, steep slopes ($> 4\%$), and high traffic volumes. Simultaneously, the specific contribution of each bicycle facility type is examined. The PSL models show a strong preference for off-street trails that are mainly on McMaster Universitys campus, which has few additional choices for cycling. At the same time, SoBi cyclists prefer to travel along cautionary un-signed bike routes on streets with low traffic volumes, followed by designated BLs, and signed on-street bike routes with mostly low traffic volumes, and try to avoid cautionary un-signed bike routes on streets with high traffic volumes and paved multi-use trails far away from main roads. Unexpectedly, a significant negative contribution of ln (Path Size) has been found, which is similar to findings from Dalumpines and Scott (2017). Instead of a correction factor, ln (Path Size) in this case is treated as a utility probably with some hidden attractive features from overlapping links. As a route choice study utilizing choice sets consisting of actual routes, the negative ln (Path Size) estimate makes sense because route overlapping occurs with some positive features, such as bicycle facilities.

4.3 RESEARCH LIMITATIONS

A common and significant constraint of both studies limiting further interpretation of results is the lack of demographic data and some other route attributes that could probably be significant, such as stop signs, red lights, and pavement quality. More comprehensive findings could be achieved if SoBi cyclists socio-demographic information, such as age, gender, income, and cycling experience, and more route attributes had been collected. Simultaneously, some routes are removed from this research due to GPS errors, which is unavoidable for all the route choice studies using GPS technology. Additionally, for Chapter 2, it would be interesting to include the comparison between dominant routes and shortest paths based on travel time as well, which is difficult to generate as the cycling speed changes among different cyclists, and it may also vary over time, especially for long distances. With respect to Chapter 3, the correlation between designated BLs and path size attribute cannot be eliminated, so the contributions of major roads without BLs and minor roads with BLs changed when the path size attribute was added into those models, especially for the high G_n model. This may be because designated BLs usually have high route overlaps upon them.

4.4 CONCLUSIONS AND FUTURE RESEARCH

According to descriptive statistics and models generated within this thesis, it can be found that most bicycle facilities, especially those developed along major roads, have positive contributions while turns, high traffic volumes, and steep slopes have negative contributions to bike share cyclist route choice behavior. As a result, in order to promote the use of BSPs, this thesis helps policy makers and urban planners explore and propose a new plan to add more bicycle facilities probably on major roads or roads with low traffic volumes after examining road segment candidates for more bicycle facilities derived from the spatial distribution of SoBi users dominant routes. Moreover, a new section within the SoBi mobile application can be developed to provide recommended routes for cyclists given the impact of each significant determinant and hub-to-hub dominant routes.

With regard to future studies in the field of cycling route choice analysis, this thesis made several contributions. First, it provides a new tool to extract unique routes from a large dataset of trips between origin-destination pairs for future research using GPS data. Second, it proposes a new logic to explore cyclists route choices using BSP data, which can provide accurate routes without concerning the high expense of equipping GPS devices for researchers because recent BSPs usually contain GPS technology. At the same time, it is the first research to generate variables for each BSP route to analyze cyclist route choice behavior. Additionally, a brand new method of generating choice sets based on cyclists actual routes is introduced because multiple unique routes between each hub pair were taken by cyclists in the real world. In this case, the identified preference for route attributes is based on actual choices from multiple BSP cyclists instead of individuals. Furthermore, this thesis focused only on one BSP: SoBi Hamilton. Although this thesis did determine the contributions of factors to BSP users route choices, we recommend for planners and professionals that additional analysis of other BSP should be conducted in order to solidify these findings in the cycling route choice context. Similar descriptive statistics and models to those described above in this thesis need to be produced for other BSPs to better understand the determinants of cyclist route choice behavior in general.

REFERENCES

Aultman-Hall, L., Hall, F., & Baetz, B. (1997). Analysis of bicycle commuter routes using geographic information systems: implications for bicycle planning. *Transportation Research Record: Journal of the Transportation Research Board*, 1578: 102-110.

Broach, J., Dill, J., & Gliebe, J. (2012). Where do cyclists ride? A route choice model developed with revealed preference GPS data. *Transportation Research Part A: Policy and Practice*, 46: 1730-1740.

Dalumpines, R., & Scott, D. M. (2017). Determinants of route choice behavior: A comparison of shop versus work trips using the Potential Path Area-Gateway (PPAG) algorithm and Path-Size Logit. *Journal of Transport Geography*, 59: 59-68.

Hood, J., Sall, E., & Charlton, B. (2011). A GPS-based bicycle route choice model for San Francisco, California. *Transportation Letters*, 3: 63-75.

Lima, A., Stanojevic, R., Papagiannaki, D., Rodriguez, P., & Gonzlez, M. C. (2016). Understanding individual routing behaviour. *Journal of the Royal Society Interface*, 13: 20160021.

Sener, I. N., Eluru, N., & Bhat, C. R. (2009). An analysis of bicycle route choice preferences in Texas, US. *Transportation*, 36: 511-539.

Stinson, M., & Bhat, C. (2003). Commuter bicyclist route choice: Analysis using a stated preference survey. *Transportation Research Record: Journal of the Transportation Research Board*, 1828: 107-115.

Tilahun, N. Y., Levinson, D. M., & Krizek, K. J. (2007). Trails, lanes, or traffic: Valuing bicycle facilities with an adaptive stated preference survey. *Transportation Research Part A: Policy and Practice*, 41: 287-301.

Winters, M., & Teschke, K. (2010). Route preferences among adults in the near market for bicycling: findings of the Cycling in Cities study. *American Journal of Health Promotion*, 25: 40-47.

Winters, M., Teschke, K., Grant, M., Setton, E., & Brauer, M. (2010). How far out of the way will we travel? Built environment influences on route selection for bicycle and car travel. *Transportation Research Record: Journal of the Transportation Research Board*, 2190: 1-10.