University Transportation Research Center - Region 2

# Final Report

## Data Collection and Econometric Analysis of the Demand for Nonmotorized Transportation

Performing Organization: Cornell University

January 2014

Sponsor:
University Transportaton Research Center - Region 2

## University Transportation Research Center - Region 2

The Region 2 University Transportation Research Center (UTRC) is one of ten original University Transportation Centers established in 1987 by the U.S. Congress. These Centers were established with the recognition that transportation plays a key role in the nation's economy and the quality of life of its citizens. University faculty members provide a critical link in resolving our national and regional transportation problems while training the professionals who address our transportation systems and their customers on a daily basis.

The UTRC was established in order to support research, education and the transfer of technology in the field of transportation. The theme of the Center is "Planning and Managing Regional Transportation Systems in a Changing World." Presently, under the direction of Dr. Camille Kamga, the UTRC represents USDOT Region II, including New York, New Jersey, Puerto Rico and the U.S. Virgin Islands. Functioning as a consortium of twelve major Universities throughout the region, UTRC is located at the CUNY Institute for Transportation Systems at The City College of New York, the lead institution of the consortium. The Center, through its consortium, an Agency-Industry Council and its Director and Staff, supports research, education, and technology transfer under its theme. UTRC's three main goals are:

**Research**

The research program objectives are (1) to develop a theme based transportation research program that is responsive to the needs of regional transportation organizations and stakeholders, and (2) to conduct that program in cooperation with the partners. The program includes both studies that are identified with research partners of projects targeted to the theme, and targeted, short-term projects. The program develops competitive proposals, which are evaluated to insure the mostresponsive UTRC team conducts the work. The research program is responsive to the UTRC theme: "Planning and Managing Regional Transportation Systems in a Changing World." The complex transportation system of transit and infrastructure, and the rapidly changing environment impacts the nation's largest city and metropolitan area. The New York/New Jersey Metropolitan has over 19 million people, 600,000 businesses and 9 million workers. The Region's intermodal and multimodal systems must serve all customers and stakeholders within the region and globally.Under the current grant, the new research projects and the ongoing research projects concentrate the program efforts on the categories of Transportation Systems Performance and Information Infrastructure to provide needed services to the New Jersey Department of Transportation, New York City Department of Transportation, New York Metropolitan Transportation Council , New York State Department of Transportation, and the New York State Energy and Research Development Authorityand others, all while enhancing the center's theme.

**Education and Workforce Development**

The modern professional must combine the technical skills of engineering and planning with knowledge of economics, environmental science, management, finance, and law as well as negotiation skills, psychology and sociology. And, she/he must be computer literate, wired to the web, and knowledgeable about advances in information technology. UTRC's education and training efforts provide a multidisciplinary program of course work and experiential learning to train students and provide advanced training or retraining of practitioners to plan and manage regional transportation systems. UTRC must meet the need to educate the undergraduate and graduate student with a foundation of transportation fundamentals that allows for solving complex problems in a world much more dynamic than even a decade ago. Simultaneously, the demand for continuing education is growing – either because of professional license requirements or because the workplace demands it – and provides the opportunity to combine State of Practice education with tailored ways of delivering content.

**Technology Transfer**

UTRC's Technology Transfer Program goes beyond what might be considered "traditional" technology transfer activities. Its main objectives are (1) to increase the awareness and level of information concerning transportation issues facing Region 2; (2) to improve the knowledge base and approach to problem solving of the region's transportation workforce, from those operating the systems to those at the most senior level of managing the system; and by doing so, to improve the overall professional capability of the transportation workforce; (3) to stimulate discussion and debate concerning the integration of new technologies into our culture, our work and our transportation systems; (4) to provide the more traditional but extremely important job of disseminating research and project reports, studies, analysis and use of tools to the education, research and practicing community both nationally and internationally; and (5) to provide unbiased information and testimony to decision-makers concerning regional transportation issues consistent with the UTRC theme.

**UTRC-RF Project No:** 49997-35-24

**Project Date:** January 2014

**Project Title:** Data Collection and Econometric Analysis of the Demand for Nonmotorized Transportation

**Project's Website:**
http://www.utrc2.org/research/projects/data-collection-and-econometric-analysis

**Principal Investigator:**

Dr. Ricardo A. Daziano
Assistant Professor
School of Civil and Environmental Engineering
Cornell University
305 Hollister
Ithaca, NY 14850
Tel: 607 255-2018
Email: ra477@cornell.edu

**Co-Author:**

Dr. Yutaka Motoaki
School of Civil and Environmental Engineering
Cornell University
301 Hollister, Ithaca, NY 14850
Email: ym323@cornell.edu

**Performing Organizations:** Cornell University

**Sponsor:**

University Transportation Research Center - Region 2, A Regional University Transportation Center sponsored by the U.S. Department of Transportation's Research and Innovative Technology Administration

To request a hard copy of our final reports, please send us an email at utrc@utrc2.org

**Mailing Address:**

University Transportation Reserch Center
The City College of New York
Marshak Hall, Suite 910
160 Convent Avenue
New York, NY 10031
Tel: 212-650-8051
Fax: 212-650-8374
Web: www.utrc2.org

## Board of Directors

The UTRC Board of Directors consists of one or two members from each Consortium school (each school receives two votes regardless of the number of representatives on the board). The Center Director is an ex-officio member of the Board and The Center management team serves as staff to the Board.

**City University of New York**
*Dr. Hongmian Gong - Geography*
*Dr. Neville A. Parker - Civil Engineering*

**Clarkson University**
*Dr. Kerop D. Janoyan - Civil Engineering*

**Columbia University**
*Dr. Raimondo Betti - Civil Engineering*
*Dr. Elliott Sclar - Urban and Regional Planning*

**Cornell University**
*Dr. Huaizhu (Oliver) Gao - Civil Engineering*
*Dr. Mark A. Turnquist - Civil Engineering*

**Hofstra University**
*Dr. Jean-Paul Rodrigue - Global Studies and Geography*

**Manhattan College**
*Dr. Anirban De - Civil & Environmental Engineering*
*Dominic Esposito - Research Administration*

**New Jersey Institute of Technology**
*Dr. Steven Chien - Civil Engineering*
*Dr. Joyoung Lee - Civil & Environmental Engineering*

**New York Institute of Technology**
*Dr. Nada Marie Anid - Engineering & Computing Sciences*
*Dr. Marta Panero - Engineering & Computing Sciences*

**New York University**
*Dr. Mitchell L. Moss - Urban Policy and Planning*
*Dr. Rae Zimmerman - Planning and Public Administration*

**Polytechnic Institute of NYU**
*Dr. John C. Falcocchio - Civil Engineering*
*Dr. Kaan Ozbay - Civil Engineering*

**Rensselaer Polytechnic Institute**
*Dr. José Holguín-Veras - Civil Engineering*
*Dr. William "Al" Wallace - Systems Engineering*

**Rochester Institute of Technology**
*Dr. J. Scott Hawker - Software Engineering*
*Dr. James Winebrake -Science, Technology, & Society/Public Policy*

**Rowan University**
*Dr. Yusuf Mehta - Civil Engineering*
*Dr. Beena Sukumaran - Civil Engineering*

**Rutgers University**
*Dr. Robert Noland - Planning and Public Policy*

**State University of New York**
*Michael M. Fancher - Nanoscience*
*Dr. Catherine T. Lawson - City & Regional Planning*
*Dr. Adel W. Sadek - Transportation Systems Engineering*
*Dr. Shmuel Yahalom - Economics*

**Stevens Institute of Technology**
*Dr. Sophia Hassiotis - Civil Engineering*
*Dr. Thomas H. Wakeman III - Civil Engineering*

**Syracuse University**
*Dr. Riyad S. Aboutaha - Civil Engineering*
*Dr. O. Sam Salem - Construction Engineering and Management*

**The College of New Jersey**
*Dr. Thomas M. Brennan Jr. - Civil Engineering*

**University of Puerto Rico - Mayagüez**
*Dr. Ismael Pagán-Trinidad - Civil Engineering*
*Dr. Didier M. Valdés-Díaz - Civil Engineering*

## UTRC Consortium Universities

The following universities/colleges are members of the UTRC consortium.

City University of New York (CUNY)
Clarkson University (Clarkson)
Columbia University (Columbia)
Cornell University (Cornell)
Hofstra University (Hofstra)
Manhattan College
New Jersey Institute of Technology (NJIT)
New York Institute of Technology (NYIT)
New York University (NYU)
Polytechnic Institute of NYU (Poly)
Rensselaer Polytechnic Institute (RPI)
Rochester Institute of Technology (RIT)
Rowan University (Rowan)
Rutgers University (Rutgers)*
State University of New York (SUNY)
Stevens Institute of Technology (Stevens)
Syracuse University (SU)
The College of New Jersey (TCNJ)
University of Puerto Rico - Mayagüez (UPRM)

** Member under SAFETEA-LU Legislation*

## UTRC Key Staff

**Dr. Camille Kamga:** *Director, UTRC*
*Assistant Professor of Civil Engineering, CCNY*

**Dr. Robert E. Paaswell:** *Director Emeritus of UTRC and Distin*guished Professor of Civil Engineering, The City College of New York

**Herbert Levinson:** *UTRC Icon Mentor, Transportation Consultant and Professor Emeritus of Transportation*

**Dr. Ellen Thorson:** *Senior Research Fellow, University Transportation Research Center*

**Penny Eickemeyer:** *Associate Director for Research, UTRC*

**Dr. Alison Conway:** *Associate Director for New Initiatives and Assistant Professor of Civil Engineerin*g

**Nadia Aslam:** *Assistant Director for Technology Transfer*

**Dr. Anil Yazici:** *Post-doc/ Senior Researcher*

**Nathalie Martinez:** *Research Associate/Budget Analyst*

TECHNICAL REPORT STANDARD TITLE PAGE

| 1. Report No. | 2.Government Accession No. | 3. Recipient's Catalog No. |
|---|---|---|
| 4. Title and Subtitle<br>Data collection and econometric analysis of the demand for non-motorized transportation | | 5. Report Date<br>January 2014<br>6. Performing Organization Code |
| 7. Author(s)<br>Dr. Ricardo A. Daziano and Dr. Yutaka Motoaki | | 8. Performing Organization Report No. |
| 9. Performing Organization Name and Address<br>Cornell University<br>School of Civil & Environmental Engineering<br>220 Hollister Hall<br>Ithaca, NY 14853 | | 10. Work Unit No.<br>11. Contract or Grant No.<br>49997-35-24 |
| 12. Sponsoring Agency Name and Address<br>University Transportation Research Center<br>City College of New York-Marshak 910<br>160 Convent Avenue<br>NewYork, NY 10031 | | 13. Type of Report and Period Covered<br>Final Report<br>14. Sponsoring Agency Code |
| 15. Supplementary Notes | | |

16. Abstract

In this project, we derived a latent class model with a class assignment mechanism based on the latent bicycle status of the respondent. Two segments were identified: more-skilled and experienced cyclists, versus less-skilled- and non-cyclists. The two segments have different sensitivities to the factors that may encourage or discourage riding a bike. For instance, slope inclination is considered almost 3 times as bad by less-skilled cyclists. Heavy traffic affects twice as much to less-skilled cyclists, who also consider rain to be 2.4 times more bothersome (and snow almost 4 times more bothersome) than more-skilled cyclists. On the other hand, bike lanes are 1.6 times more appreciated by less-skilled cyclists. Because in cycling route decisions there is no direct monetary cost involved, to analyze differences in the taste parameters we have proposed to use the ratio of the marginal rate of substitution with respect to travel time. In addition, we measured the diminishing negative effect of a hilly topography (slope inclination) as a function of the physical condition of the cyclist.

In terms of policy recommendations, our results suggest that the provision of bike lanes may encourage an increase in the modal share of cycling, especially among those individuals using a bike infrequently, or mostly for recreational purposes.

We also examined the performance of several ridership prediction models, including the Negative Binomial regression and time-series models such as SARIMA and SARIMAX. Using cycling counts for Portland, we show that the SARIMAX model that includes weather conditions (temperature and precipitation) as explanatory variables performs best in out-of-sample prediction. Future research in State Space models is needed for overcome the problems of SARIMAX when predicting ridership in periods with really poor weather.

In sum, both the discrete choice and time series analyses coincide in that poor weather conditions are indeed a main determinant for discouraging cycling as a transportation alternative.

| 17. Key Words | 18. Distribution Statement | | |
|---|---|---|---|
| Mobility, Economics, Pollution, Environment, Congestion, Bicycling, Infrastructure | | | |
| 19. Security Classif (of this report) | 20. Security Classif. (of this page) | 21. No of Pages | 22. Price |
| Unclassified | Unclassified | 79 | |

Form DOT F 1700.7 (8-69)

**Disclaimer**

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. The contents do not necessarily reflect the official views or policies of the UTRC or the Federal Highway Administration. This report does not constitute a standard, specification or regulation. This document is disseminated under the sponsorship of the Department of Transportation, University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

# 1. INTRODUCTION

Current mobility patterns in the U.S. are often characterized by automobile dependence, creating many societal, economic, health, and environmental problems. The negative externalities of automobile-dependent societies range from congestion and high levels of pollution to health issues due to lack of physical activity (Litman & Laube, 2002).

Although relatively little attention has been paid to cycling when compared with other modes of transport, increasing the level of bicycle transportation can be a solution to those health and environmental problems. To encourage the use of bicycles and to improve decision support of policies aiming at encouraging a sustainable modal shift, we need to better understand the motives underlying demand for cycling commute. Although econometric travel demand models are highly valuable for assessing the effect of policies, there are several challenges. The characteristics of bicycle travel are very different from those of motorized transportation modes in that the decision to bike does not involve direct monetary costs. Users of the transportation system may be motivated to cycle or walk not because of the tradeoff between cost and time, but because of other factors, such as weather and the physical effort needed (Wardman et al., 1997).

Several studies found that there are also more subtle social and psychological factors which act as determinants to cycle commuting, such as attitude toward biking, the perception of the effect of weather conditions, and the way cycle commuters are perceived socially in a car dominated-society (see, Nankervis, 1997, Heninen, et al, 2010, Wang, et al., 2013; Gatersleben & Appleton, 2007).

Bicycling improvements are particularly appropriate and effective for transportation management on university campuses for a number of reasons. University community consists of many young and physically active student commuters, and if they acquire the environmental transportation habit they will retain it after their graduation (Balsas, 2004). Nevertheless, if we expect that bicycling improvements will raise ridership in the future, careful planning and appropriate investment in bicycle infrastructure are necessary in order to accommodate future demand. Since accurate ridership prediction will be indispensable for those planning investment decisions, it is important to understand various ridership prediction models that are capable of identifying significant factors related to the motivation for people to bicycle.

## 1.1 Research Approach

The objective of this proposal is to **advance our understanding of the motives underlying the choice of non-motorized modes** in order to improve decision support of policies aiming at encouraging a sustainable modal shift. In pursuit of this goal, two relevant dimensions of the demand for cycling will be analyzed as two related research projects: derivation of a demand model for discrete choices of non-motorized travel, and development of a methodology for statistical analysis cycling-count series to capture the effect of weather on demand.

On the one hand, microeconomic demand models allow researchers to represent compensatory behavior, but in the context of non-motorized alternatives these models fail to properly explain consumers' evaluation of non-instrumental attributes. Discrete choice models consider tradeoffs among directly measurable attributes (travel cost and time), without consideration to other relevant dimensions such as lifestyle or environmental attitudes. In this project an Internet-based stated-preference survey is designed to test a discrete choice model aimed at better explaining sustainable mobility patterns and behaviors. Special focus is given to non-traditional factors that motivate or discourage individuals to cycle and/or walk instead of driving. Two instruments are used to collect data: focus groups and an Internet survey targeted at Cornell students. The data is used to test and validate new demand estimators based on innovative modeling extensions of discrete choice.

On the other hand, a traditional econometric tool for studying demand is time series analysis. The use of innovative sensors that monitor active transportation enables collecting massive amounts of data, entailing the need to revise current statistical modeling procedures. In effect, a major challenge in transportation engineering is how to fully exploit new data collection methods such as non-intrusive inductive loops that can count mixed traffic with high precision. Several cities are investing in automatic counting stations not only for motorized vehicles but also for bicycles, including New York City. As counts from more cities and more stations within a city become available, a good methodology for disaggregate time series analysis will be needed to understand the trends and fluctuations in cycling demand. The goal is to develop a comprehensive methodology to understand the effects of weather on cycling demand using automatic counts (cf. Brandenberg et al., 2007). This methodology is applied to counts in Portland, Oregon.

The remaining sections of this report are organized as follows: First, we review the literature in bicycle research and discuss the characteristics of transportation at the Cornell campus and its surrounding environment. Second, we discuss the survey instrument we developed for this project, as well as the findings from the two focus groups sessions conducted during the study. Third, we analyze the survey data using discrete choice models with a latent factor. Fourth, we analyze time-series data of bicycle ridership counts with a particular focus on examining forecasting capabilities of various econometric models. The last section discusses both our findings and future research opportunities.

# 2. BACKGROUND & LITERATURE REVIEW

Heinen, et al. (2010) and Sener, et al (2009) conducted a comprehensive review on previous bicycling research. In this chapter, we place our primary focus on recent studies that have analyzed the effects of weather, landscape, and the general bicycling situation, especially at university communities.

## 2.1. Bicycling trends in the U.S.

Bicycle travel has received increasing attention in recent years as planners and policymakers recognize the benefits of bicycling to communities, public health, and the environment (United States, 2000). There has been considerable growth in cycling over the past few decades — the total number of bike trips in the U.S. more than tripled between 1977 and 2009, while the bike share of total trips almost doubled, rising from 0.6% to 1.0% (Pucher, et al., 2011). As a result, some local governments have changed local zoning ordinances to encourage bicycle commuting (Buehler, 2012). However, the bike share of total trips is still very small. In the U.S., bicycles are used primarily for recreation and not for daily urban travel (Pucher, et al., 1999, Dill, 2009; NHTSA). The dependency on automobiles has resulted in land use patterns that favor automobile trips that undermined the safety and comfort of both pedestrians and bicyclists (Balsas, 2002). Bicycling is impeded by both the country's automobile based transport system, and the lack of a tradition of cycling for utilitarian purposes and cultural status of cyclists.  Balsas (2002) stresses that this was a result of planning agencies focusing too much on accommodating car traffic in cities and so little on non-motorized and environmentally friendly modes of transportation. Bike commuting is often said to be associated with individuals' socio-demographic characteristics. For example, Buehler (2012) found that bike commuting is more common for whites, males, and individuals with a high income. The share of bike commuters is greater in the urban core, at higher population densities, and in areas with more bike paths and lanes.

## 2.2. Transportation in U.S. University Communities

There is a growing literature examining travel patterns on college campuses because of the increasing adverse effects of motorized vehicles to health and environment in university communities (Wang, et al., 2013).  Universities campuses are no exceptions from an auto-dependent environment where much of the infrastructure is built for cars and other motorized vehicles rather than for bicycles and pedestrians. Akar & Clifton (2009) studied the impact of possible bicycle infrastructure improvements and policy at the University of Maryland (UMD), where the rate of cycling commute is low among students. Their findings suggest that both non-bicycle commuters and bicycle commuters at UMD agree that bicycle facilities such as lanes, trails, and paths would encourage them to ride a bike to campus. The respondents also identified the safety concern from vehicular traffic as one of the most important factors that keep them from bicycling. From the findings of the study, the authors suggested several policy recommendations: (1) construction of bike lanes and other bike infrastructure; (2) enforcement of stricter traffic rules against drivers; (3) providing a campus map showing bicycle routes; and (4) increasing security for students on campus, especially at night time. Wang, et al. (2013) examined spatial autocorrelation in commuters' mode choices – the interactions among the decision makers in a neighborhood – at Ohio State University community. They found a "neighborhood effect" among commuters, i.e. the higher the share of non-auto commuters, the more attractive the non-auto modes become to all commuters. They also found that students are more likely to choose non-motorized modes as compared to faculty and staff. Although students' higher willingness to bike than that of faculty and staff can be attributed to their youth and physical ability, Weiss (1996) points out that students' confidence and resulting lack of safety awareness is a serious concern for university bicyclists. Weiss observed bicyclists riding onto the University of Arizona campus in 1985, 1990, and 1994 and recorded the percentage of bicyclists who were wearing a helmet. He found that among university students the percentage of helmet usage is low; the highest value recorded in 1994 was only 25 %. A simple solution for this is a mandatory helmet requirement; however, a helmet requirement law may reduce cycling rate (Jong 2012).

### 2.3. The City of Ithaca and Cornell University Ithaca Campus

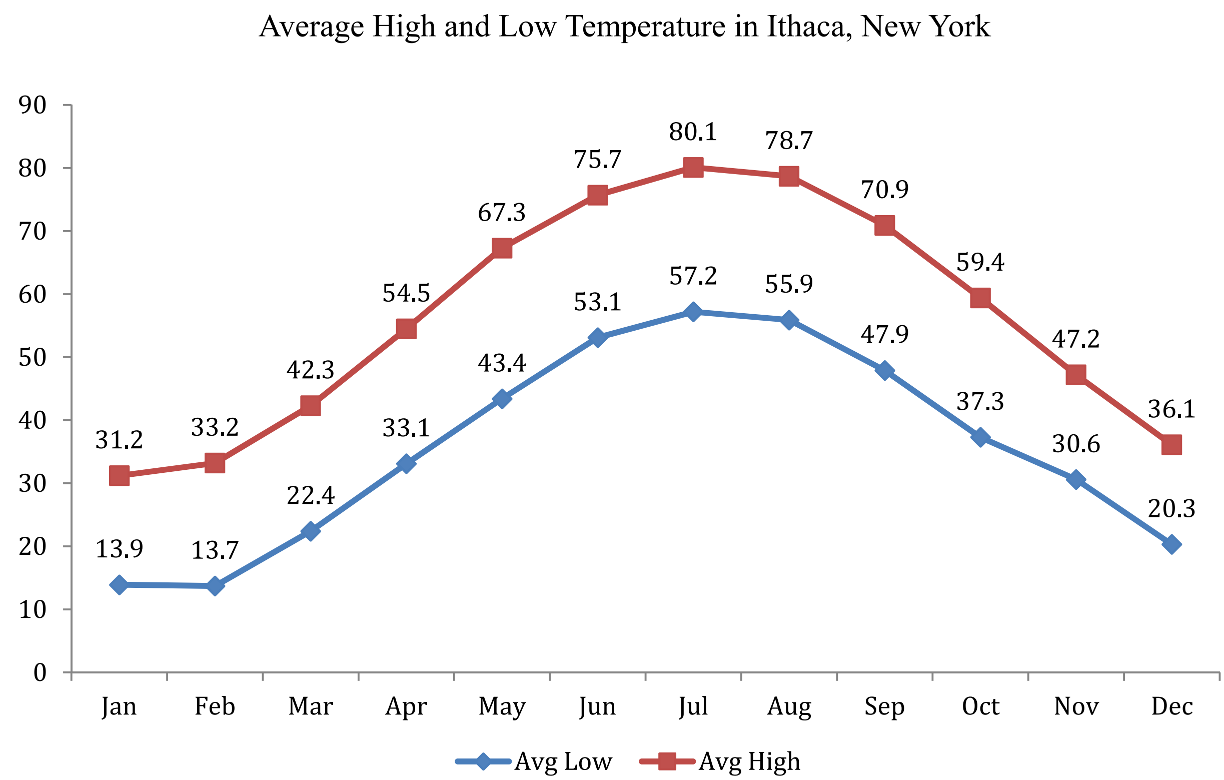
The Cornell University campus is located on a hill about 400 ft. above downtown of the city of Ithaca. Cornell's Ithaca campus is a community of approximately 21,000 students and 9,600 faculty and staff members (Cornell University 2012). The City of Ithaca is home to about 30,000 residents (U.S. Census). The climate of the area is characterized by hot and wet weather in summer, and cold and snowy weather in winter. Figure 2-1 shows the average monthly temperature, and Figure 2-2 shows the average monthly precipitation of Ithaca.

Cornell instituted a Transportation Demand Management Plan in the 1990s, whose aim is "getting people out of single-occupancy vehicles and into alternative modes of transportation" (Cornell University 2008). A result was that Cornell managed parking and driving demand by eliminating a plan to build the additional parking spaces considered necessary at the time. Since then, Cornell University has continued to promote sustainable land use and environmentally friendly transportation plans, such as limiting growth within core campus boundaries, promoting mixed-use, and ensuring an easy walking and a cycling-friendly campus environment (Campus Planning 2008). In fact, Cornell imposes an extremely high annual parking permit cost of $725.76, much higher than the parking permit prices of nearby colleges and universities ($123 at Ithaca College and $135.55 at Binghamton University).

Cornell students use a variety of travel modes to commute to campus, including private cars, buses, bicycling, and walking. The public transit in Ithaca is operated by TCAT, Ithaca's local public transit agency, and widely used by Cornell students. The frequency of bus trips is high for a relatively small city. Downtown-campus circulator loop runs as frequent as every 5 minutes at peak hours, and both undergraduate and graduate students are given unlimited full year passes in their first year at Cornell. Figure 2-3 shows the modal split by undergraduate students, graduate students, and employees at Cornell. (Charts of modal split at University of Maryland are also presented for comparison.) Despite a majority of students (84% of graduate students and 97% of undergraduates) living within 5 miles of campus, the share of bicycle as a commuting transportation mode is very small. The 2006 Cornell University Travel Survey shows that the share of bicycle as primary mode to commute to campus was 1.4% for undergraduate students, 4.0% for graduate students, and 3.1% for employees.

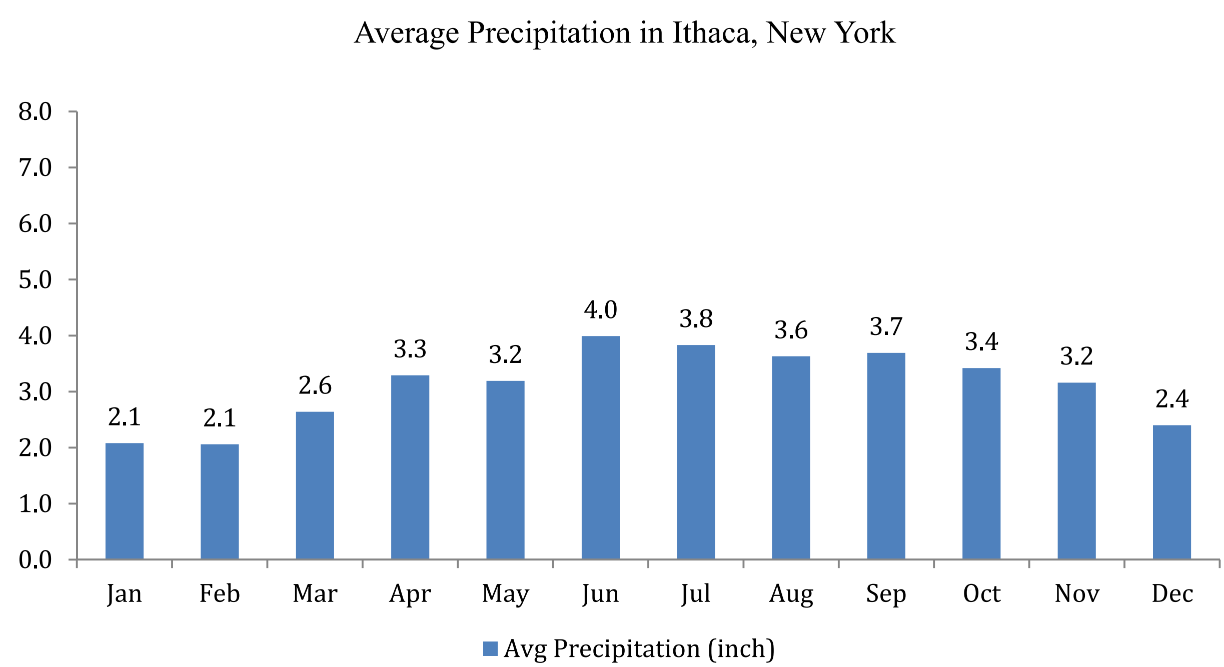
In the recent past, several initiatives have been taken to increase bicycle travel within campus. In 2010, Cornell students launched Big Red Bikes, a bicycle share program for students with several bicycle stations on campus available from March to November. The program allows for free bicycle sharing for 25 hours per week. In addition, Cornell University's Department of Transportation and Mail Services provides a campus map showing bicycle routes so that bicyclists can easily find routes that are safe for them to ride (Figure 2-4).

**Figure 2-1.**




***Data Source: Northeast Regional Climate Center***

**Figure 2-2.**




***Data Source: The Weather Channel***

## 2.4. Bicycle Environment, Landscape, and Bicycle Facilities

The costs and deterrence of a non-motorized mode such as bicycling are different from those of motorized modes. Whereas for car and transit users, landscape and altitude changes would not appear to be relevant, these factors require additional time and effort for bicycle travelling. The disutility of cycling is said to increase more than proportionally for longer distances, and the bicycle share decreases rapidly for distances longer than 5 km (van Wee et al., 2006). Pucher & Buehler (2006) identified both strict land use policies and high urban densities as the key factors that explain the higher levels of cycling in Canada compared to those in the U.S. Very short distances can also discourage bicycling. Keijer & Rietveld (2000), for example, show that for trips up to 2 km, bicycle is a less attractive mode of transportation because people would rather walk. The effects of topography are somewhat ambiguous. In general, one would expect that a hilly terrain would negatively impact bicycle choice. However, Moudon *et al.* (2005) found that the effects of the built environment (bicycle lanes, traffic speed, traffic volume, topographical conditions, etc.) on bicycle share for all trips are not significant.

Stinson & Bhat (2005) investigated the differences in preferences over characteristics of routes between bikers and non-bikers, and they found that experienced bikers actually prefer hilly environments to either flat or mountainous terrains for commuting, because hilly environments provide a higher level of exercise benefits compared to a flat ground. This evidence suggests that the purpose of bicycle uses and level of experience of bicyclists may influence the direction of impacts of bicycle environment on the decision to bike, e.g., recreational users and skilled bicyclists may prefer slopes due to the fact of a hilly environment being more physically demanding, whereas bicycle commuters and inexperienced bicyclists dislike slopes (Nelson & Allen, 1997).

Perhaps the most serious obstacle to bicycling conditions is the conflict between bicycles and automobiles sharing the same road, due to the lack of designated bicycle paths. Some studies suggest that presence of a bicycle facility is a very relevant determinant, and it is critically related to the safety of cyclists (Heinen, et al, 2010). From the bicyclists' point of view, two types of safety can be considered: objective and subjective safety. Objective safety is usually measured in terms of the number of bicycle-related incidents. Subjective safety refers to how individuals perceive safety, and is mostly measured in terms of stated safety reports (Heinen, et al, 2010). In order to increase the modal share of bicycles, it is said that subjective safety levels must be satisfied as positive perceptions of the availability of bike facility lead to more cycling and the desire to cycle more (Klobucar & Fricker 2007; Dill & Voros, 2007). Therefore, having more bicycle paths may result in a higher share of cycling (Barnes & Thompson, 2006; Pucher & Buehler, 2006). The type of bicycle infrastructure also matters. For example, bicyclists prefer bicycle paths to bicycle lanes and routes with more continuous facilities than discontinuous ones (Stinson & Bhat, 2005). Garrard et al. (2008) suggest that improving cycling infrastructure, which provides a high degree of separation from motor traffic, is likely to be important for increasing transportation cycling for women since female cyclists are likely to use routes with complete separation from motorized traffic. The causal relationships of bicycle infrastructures and modal share of bicycle is discussed further below.

**Figure 2-3**

**Mode Split of Transportation: Cornell University vs. University of Maryland**
***Source:*** Transportation Impact Mitigation Strategies: Draft Report (Cornell)
Akar & Clifton (2009)

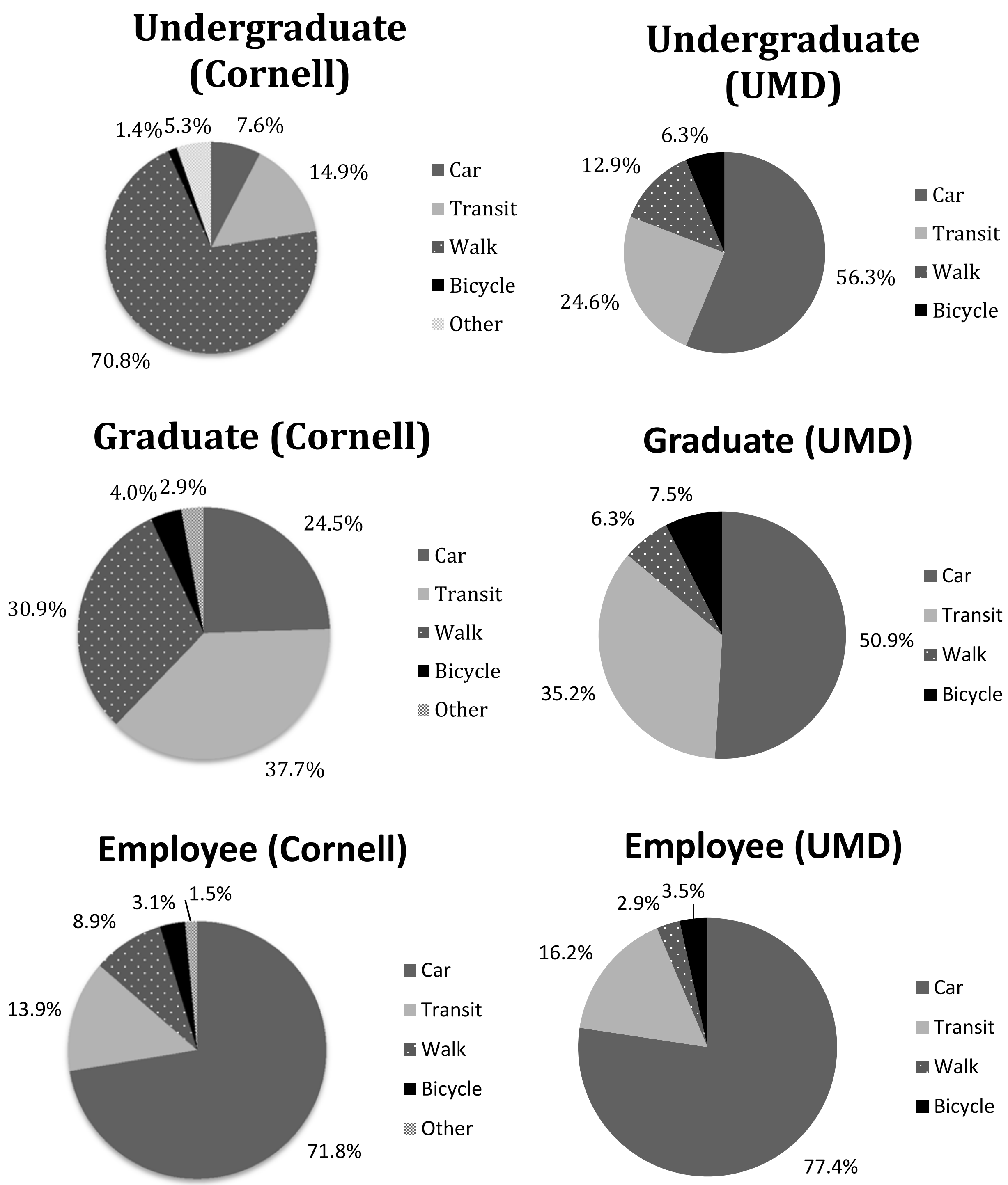
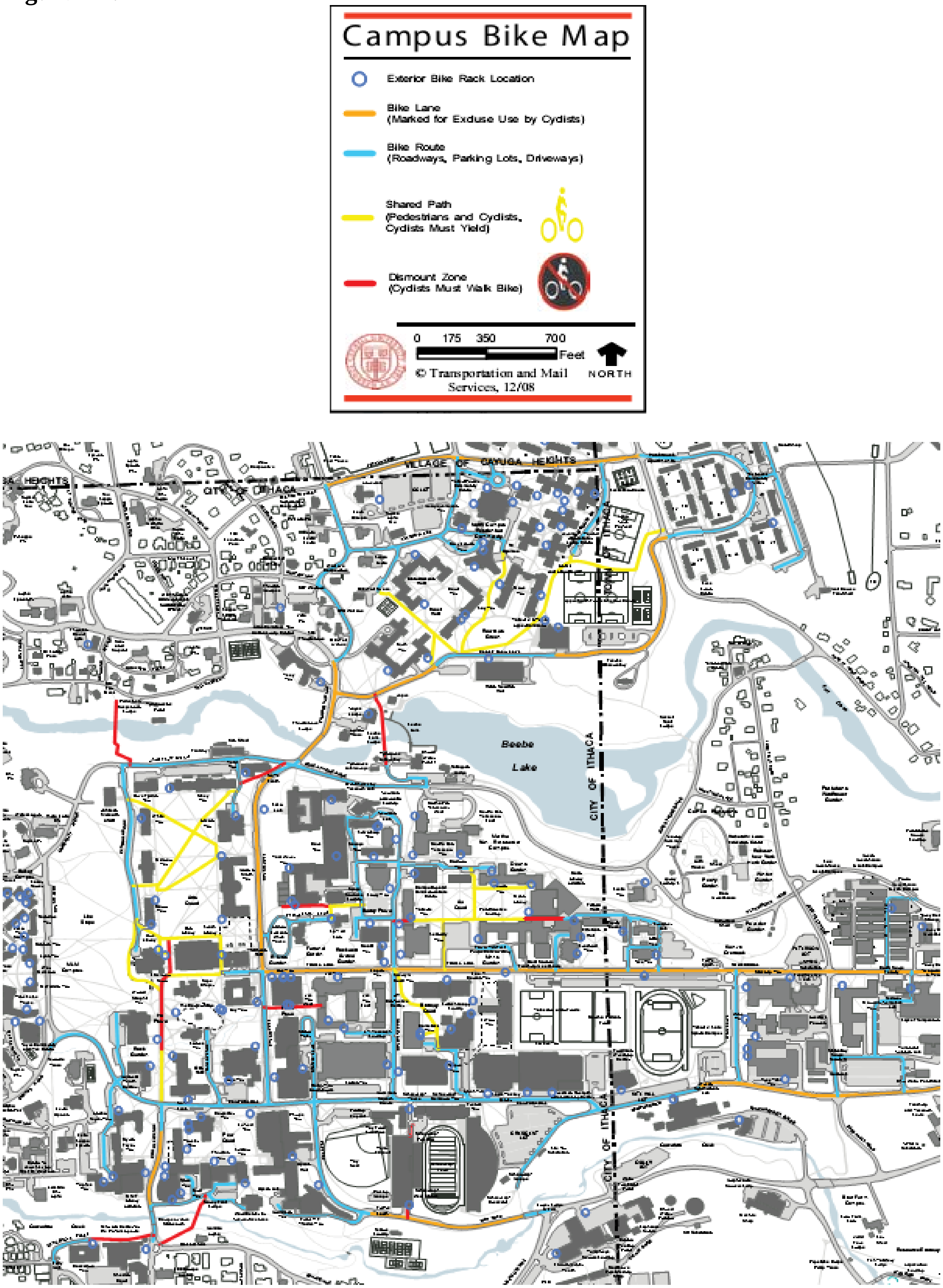


**Figure 2-4.**



Campus Bike Map
Exterior Bike Rack Location
Bike Lane
(Marked for Exclusive Use by Cyclists)
Bike Route
(Roadways, Parking Lots, Driveways)
Shared Path
(Pedestrians and Cyclists,
Cyclists Must Yield)
Dismount Zone
(Cyclists Must Walk Bike)
0 175 350 700
Feet
NORTH
© Transportation and Mail
Services, 12/08
VILLAGE OF CAYUGA HEIGHTS
CITY OF ITHACA
Beebe
Lake
CITY OF ITHACA

### *Effects of Weather*

Inclement weather is naturally considered as a strong disincentive to cycling. The unpredictable nature of daily weather can make cycling difficult or even impossible. Contradictory, Nankervis (1997) analyzed the effects of weather on bicycle commuting by students in Australian universities and found the association between commuter cycling and weather is less strong than was generally perceived. Nevertheless, the effects of weather may depend on the regional differences in the severity of weather. For example, Nankervis's study was conducted in Melbourne, Australia where, despite its erratic nature, weather is generally warm. In fact, studies conducted in colder regions find different results. Saneinejad, et al. (2012) examined the effects of wider range of weather (temperatures ranging from below 0 to above 35 degrees Celsius) on bicycle commuting. The authors found that in temperatures higher than 15°C bicyclists become insensitive to temperature, while for temperatures below 15°C the utility of cycling decreases. In sum, bicyclists are insensitive to temperature increases in warm weather but sensitive to temperature decreases in cold weather. They also found that the kind of precipitation has different effects on different types of bicyclists. Precipitation in the form of showers negatively impacts cyclists about twice as much as pedestrians, and this negative influence is greatest for young cyclists and it gradually improves for older cyclists. They did not examined the effect of snowfall on bicycle commute, and to our knowledge, it has never been examined in the literature. Nankervis (1997) suggests that the perception of the effect of weather conditions – besides the actual weather – may also have effects on the decision to bike. This was examined in Miranda-Moreno & Nosal (2011). Miranda-Moreno & Nosal conducted a time series analysis of the effect of weather on bicycle ridership in an urban Canadian environment using automatic bicycle counters that were installed in the City of Montreal. Based on both the bicycle traffic counts and historical weather data, the effects of weather (hourly values for temperature, precipitation, wind speed, and relative humidity) on hourly bicycle counts were examined. They found that the impact of precipitation is not only important in the current hour but also in any of the previous 3 hours. They also found that a temperature increase can impact ridership both positively and negatively depending on the reference temperature: they found that an increase of 5° from the average temperature has a positive effect during colder months but a strong negative effect during the warmer months. Rose et al. (2011) applied a similar methodology to analyze the regional differences in the bicyclists' sensitivity to weather on bicycle ridership in Melbourne, Australia and Portland, Oregon. Their findings show that cyclists in the two cities exhibit quite different sensitivities to weather. Cyclists in Melbourne are more sensitive to rain, and higher temperatures stimulated higher demand for cycling in Portland than in Melbourne. The authors attribute these regional differences to regional differences in culture and the type of rain, i.e., Portland's rain tends to be a gentle drizzle whereas Melbourne's rain tend to be an intense thunderstorm. In addition, Portland residents are accustomed to drizzles that are so prevalent in the region.

### *Modeling*

In the previous literature, several determinants regarding bicycle commuting have been examined and identified; however, it is often noted that the causal relationships between bicycle commuting and those determinants are ambiguous. On the one hand, several studies have found a positive statistical association between bicycle facilities and bicycle commuting, and these studies therefore suggest that an inadequate supply of bicycle facilities is the cause of low bicycle commuting (see, Nelson and Allen, 1997; Dill & Carr, 2003; Pucher & Buehler, 2006). Dill & Voros (2007), on the other hand, did not find such an association between bicycle facilities and the level of bicycle commuting. Other ambiguous causal relationships also exist in the literature. For example, the effect of neighborhood perceptions on cycling was discussed in Dill & Voros (2007). In their analysis, although the authors found such an association, they noted, "It is difficult to know the direction of the relationship between positive neighborhood perceptions and cycling, however. Do cyclists think that their neighborhood is better for cycling because they cycle and are more aware, or do their perceptions reflect the characteristics of the neighborhood and encourage them to cycle more?" These ambiguities in causal relationships between bicycle ridership and various environmental factors seem to severely limit the validity of study results and corresponding policy recommendations.

One of the causes of ambiguous causal relationships is the type of data used in the analysis and their quality. In the bicycle research literature, three kinds of data are commonly used: (1) Census household survey data; (2) bicycle count data (either traffic counts or parked bicycle counts); and (3) stated preference data. Census data is usually aggregated city-level data; the analysis using this data often involves a regression of bicycle counts, bicycle share, or bicycle mileage on characteristics of municipalities and their populations (see Nelson & Allen, 1997, Dill & Carr, 2003; Barnes & Thompson, 2006). Some studies have used data collected specifically for them. These data are often based on traffic/parking counts, travel diaries, and phone interviews (e.g., Buehler, 2012, Nankervis, 1997; Garrard et al., 2008). Buehler (2012), for example, collected revealed preference travel data of commuters in Washington D.C., including cyclists and non-cyclists to examine the effects of services for bike commuters provided by their employers. In phone interviews, respondents answered whether they bike to work under the conditions in which employers provide cyclist services for bicycle commuters such as showers, bike lockers, and bike parking. Buehler found that the combined supply of bike parking, clothes lockers, and cyclist showers has a statistically strong influence on bike commuting.

While (1) and (2) are based on revealed preferences – the data of how participants "actually behaved", (3) stated preference data differs in that the data is based on responses on what people "would do" under hypothetical situations. In revealed preference analysis, a limitation arises because only the final consumer choice is observed, making it difficult to learn what information on choices went into an individual's decision and how consumers came to their final decision (Tilahun, et al., 2007). Stated preference analysis gives the researcher a greater control over the environment in which respondents make a choice. While revealed preference data reflects actual behavior, information on alternatives that went into an individual's decision may not be fully known. In fact, alternative attributes are nonorthogonal and multicollinearity may be a problem. On the other hand, stated preference data offers a more flexible approach where the researcher controls the choices and attributes of the alternatives, i.e., the analyst is able to conduct "choice experiments". One of the problems with using stated preference data is that responses are based on behavioral intentions ("what a respondent would do under a given hypothetical circumstance") and may not be fully reflective of what a respondent actually would do in reality. Stated preference data is often preferred when revealed preference data is either difficult or costly to be collected.

Discrete choice experiments in transportation usually involve multiple transportation modes as alternatives; however, the inclusion of bicycle in the choice set is problematic: while motorized modes incur pecuniary costs, bicycle incurs costs of physical effort. Physical efforts are difficult to quantify. In addition, the tradeoff between pecuniary costs and physical efforts may be difficult to model. Therefore, discrete choice experiments in bicycle research are often based on either some arbitrary options of bicycling or route choice (see Figure 2-5). Table 2-1 lists selected stated preference studies on bicycling behavior from the past decade. An example of a route choice problem taken from Stinson & Bhat (2003) is presented in Figure 2-6. In this experiment, respondents were asked to “imagine that they had just begun a new job and that they would commute by bicycle to this new job” and to choose one of the two routes presented. The attributes associated with each route were travel time, facility continuity, number of stop signs, number of red lights, and number of major cross streets. Stinson & Bhat (2005) later examined the differences in the effects of those attributes on bicycle decisions between experienced bicyclists and inexperienced bicyclists using the same data. Experienced bicyclists turned out to have a strong bias to choose the routes with minimum travel time. Safety related attributes appeared as much less influential for experienced bicyclists compared to inexperienced bicyclists. As it can be imagined, respondents may experience difficulty in visualizing the hypothetical routes described in the survey, which may lead to bias in their decisions. Tilahun, et al. (2007) avoided this problem by presenting visual aids as both pictures of routes and first person videos of bicycling on those routes.

**Table 2-1.**

| Study | Data | Analysis Method |
|---|---|---|
| Stinson & Bhat (2003) | Web-based survey conducted at University of Texas at Austin | Binary Logit (Route Choice) |
| Stinson & Bhat (2005) | Web-based survey conducted at University of Texas at Austin | Binary Logit (Route Choice) |
| Hunt, J., & Abraham, J. (2007) | Mail-back survey questionnaire distributed to cyclists in Alberta, Canada | Multinomial Logit |
| Akar & Clifton (2009) | Campus Transportation Survey at University of Maryland | Principal component analysis |
| Sener, et al. (2009) | Web-based survey conducted at University of Texas at Austin | Binary Logit (Route Choice) |
| Kamargianni & Polydoropoulou (2013) | High school students' travel behavior and travel patterns in Cyprus | Hybrid Choice Model |
| Wang, et al. (2013) | Campus Transportation Survey at Ohio State University | Principal component analysis |

**Figure 2-5**

| Route 1 | Route 2 |
|---|---|
| Minor arterial | Residential street |
| Wide (14 ft) right-hand lane | Wide (14 ft) right-hand lane |
| Continuous | Discontinuous |
| 1 or 2 red lights | No red lights |

Which route would you choose?

Route 1
Route 2

***Attitude and Incorporation of Attitude in Choice Models***

A trip by bicycle has unique characteristics in that while factors, such as time and cost of mode, constitute an objective appraisal, there are psychological factors which may act as determinants of the decision to cycle. Past research has used either factor analysis or principal component analysis to examine these psychological characteristics of individuals' regarding bicycling. Heninen et al. (2010), for example, conducted an exploratory factor analysis to identify the main attitudes regarding bicycling. The main attitudes they identified were labeled "direct trip-based benefit", "awareness", and "safety". The variable, "direct trip-based benefit", was constructed of time-saving and comfort. "Awareness" was constructed of environmental benefits, health benefits, and "mentally relaxing". The "safety" variable was constructed of social safety and traffic safety.

Akar & Clifton (2009) identified three factors that affect commuters' mode choice by using principal component analysis. The first factor is associated with people who see walking and bicycling as an opportunity for exercise. The second factor is associated with people who feel safe walking and biking on campus after dark, and the third factor is associated with people who find the car parking costs on campus high and think that they do not have many options to get to campus. Akar & Clifton found that the first two factors are positively associated with bicycling and walking, and the third factor is negatively associated with bicycling and walking.

Attitudes, measured in terms of factors, can be incorporated into a choice model, resulting in a model called the Hybrid Choice Model (HCM). HCM integrates discrete choice and latent variables models, taking into account the impact of attitudes and perceptions on the decision process (McFadden, 1986). Hybrid choice models extend choice modeling by combining the following important modeling extensions: heterogeneity through flexible error structures; the combination of revealed and stated preference data; the presence of latent classes explaining underlying market segments (through a latent class model); and the integration of latent (unobserved) constructs according to an Integrated Choice and Latent Variable (ICLV) model. Kamargianni & Polydoropoulou (2013) collected both revealed preference and stated preference data for a hybrid choice model to study the effect of teenagers' attitudes toward walking or cycling to school in Cyprus. From the survey responses about attitudes toward walking and cycling, they defined a latent variable called "willingness to walk or cycle" as an explanatory variable. The authors found that this "willingness to walk or cycle" variable has a positive impact on the bicycle choice and walk choice, and has a negative impact on car choice.

People's attitude toward bicycling may be different depending on their experience in bicycling.

People who have experienced bicycle commute may have different attitudes than people who have never contemplated bicycle commuting. Prochaska's transactional model of behavior change (Prochaska and DiClemente, 1984) considers behavior change as a process characterized by a series of stages of change (see, Table 2-2). Gatersleben & Appleton (2007) applied Prochaska's transactional mode of behavior change to bicycle actions. They conducted two studies: a survey study and an action study. In the action study, the authors conducted an experiment in which non-cyclists were asked to cycle to work for two weeks and subsequently interviewed the participants about their expectations and experiences. The experiment showed that 8 out of 22 participants said that they would continue to cycle to work even after the study. The authors concluded that as people move from the pre-contemplation stage to the action stage, their attitudes toward cycling become more positive, and their perceptions of various personal and external barriers change. They further argue that this suggests that different strategies are necessary to move people in different stages of change to action and maintenance (many of those who have never seriously considered cycling believe that they would feel strange on a bicycle, and that others would also perceive it as strange if they cycled).

Our literature review revealed two important topics in bicycle research that have not been fully investigated. First, there is an ambiguous relationship between bicycle facilities and bicycle ridership. This is important in that if policy makers hope to increase bicycle ridership, they need to know whether building bicycle facilities would create incentives for people to bike. Second, the effects of weather have not been investigated in the context of discrete choice modeling. That is, the relative marginal sensitivities of an individual to temperature, rainfall, and snowfall respectively have not been examined. We will therefore investigate those two avenues of research in this project.

**Table 2-2.** Prochaska's transactional mode of behavior change, taken from Gatersleben and Appleton (2007)

| Stage | Characteristics |
|---|---|
| Pre-contemplation | Unaware of problems, no intention to change Increase general problem awareness |
| Contemplation | Aware of problems, thinking about change Motivate, encourage specific action |
| Prepared for action | Intention to change in next 6 months Assist in developing specific plans |
| Action | Action being taken Feedback, social support, reinforcement |
| Maintenance | Has maintained action for 6 months or more Reminders, feedback, social support |

# 3. DISCRETE CHOICE MODELING

## 3.1 Introduction

This chapter presents a comprehensive approach for identifying potential factors influencing bicycle ridership and subsequently making policy recommendations to increase bicycle ridership. The approach uses structural equation modeling (SEM) to identify a latent factor that measures an individual's skills and experiences in bicycling and subsequently Hybrid Choice models and Latent Class models are used to examine how those latent factors influence decisions to bicycle.

In bicycle research, researchers are often interested in examining the differences in preferences between bikers and non-bikers. The underlying hypothesis is that people who regularly bicycle may have different tastes, behavior, and attitude toward bicycling than people who don't bicycle. By revealing their differences and segmenting the sample, researchers are able to determine the factors affecting their decisions to bike. For example, Stinson and Bhat (2003) investigated the differences in preferences over types of routes between bikers and non-bikers, and found experienced bikers in fact prefer hilly environments to flat or mountainous terrains for commuting. However, there are some issues with the segmentation methods used in the past research.

First of all, in order to segment the population into bicyclists and non-bicyclists, we need to define "bicyclists". In the past research, the definition of bicyclists has been quite arbitrary. In some studies, definition of bicyclists is simply not presented (e.g., Hunt & Abraham, 2007). Some researchers had set a certain threshold, such as frequency of bicycle use, to determine person's bicyclist status (Tilahun, et al., 2006; Noland & Kunreuther, 1995; Sener & Bhat, 2009). For example, Akar (2009) grouped the respondents of the survey as bicyclists if they rode bicycle within a year and those who did not as non-bicyclists. A self-reporting of a bicyclist experience is sometimes used. Stinson & Bhats (2004), for example, asked respondents to choose the statement that best described their bicyclist status. Those statements were: (1) I bicycle to work regularly (or at least when I can tolerate the weather), (2) I have experience in bicycling to work, but currently do not bicycle to work, (3) I am not very experienced in bicycling to work, but I might bicycle to work in the future, and (4) I am not very experienced in bicycling to work, and I am not interested in trying it. In their study, they combined categories (1) and (2) into an "experienced bicycle commuter" and the other two segments respectively as "Inexperienced, but interested group" and "Inexperienced and uninterested group". Although this method has some merits, the decision rule to classify bicyclists and non-bicyclists is still arbitrary and subjective, i.e., whether a person who rides a bicycle once a month should be considered as a bicyclist is largely a subjective decision. Because of this, inferences based on bicyclists vs. non-bicyclists will largely differ depending on how the sample was segmented into bikers and non-bikers. We therefore suggest that if we want to classify people based on their experiences and skills in bicycling, we need to estimate a continuous latent variable using multiple measurements of experiences and skills in bicycling. This way, we can examine various heterogeneity among people with different bicycle skills and experience.

In light of these, in this study, we will set three goals: (1) investigate a flexible yet thorough way of defining one's "bicyclist status" (what makes them "bicyclists") (2) examine how people's "bicycle status" affects their decision to utilize bicycles as transportation mode for commuting (3) investigate the heterogeneity between bicyclists and non-bicyclists on their preference on bicycle facilities.

## 3.2 Methodology

### *Data*

In this project we used a web-based survey to analyze bicycle route choice by evaluating (1) the trade-offs among the route facility attributes such as existence of bike lane and travel time, and (2) the effects on weather conditions on route choice decisions. We designed the survey instrument in multiple stages – two focus groups, one pilot, two samples for full data collection – to reflect specific needs to the Cornell community (see Appendix for details about survey development). We first reviewed several survey instruments to identify questions for desired topics, such as travel patterns, environmental factors, and perceptions associated with bicycling (Akar and Clifton, 2009; Stinson and Bhat, 2003). We chose an online survey instead of a paper survey because of the high rate of Internet usage in the target population.

The survey instrument consisted of 23 survey items, which fell into six categories. The first category, **Travel characteristics**, asked about basic travel patterns, including bicycle use. The second category, **Obstacles**, was concerned with factors that discourage respondents to commute to campus by bicycle. The third category, **Improvement**, was concerned with factors that encourage the use of bicycle. Included in the second and third categories were factors primarily related to the natural and built environments and the lack of bicycle facilities. The fourth category, **Behavior and perceptions**, included seven statements about subjective perceptions related to bicycling and self-evaluation of physical ability using a five-point Likert Scale (Strongly Disagree, Disagree, Neither Disagree or Agree, Agree, and Strongly Agree). Since bicycling necessarily requires physical efforts, we also speculate that the decision to bicycle is also related to one's willingness and confidence to exercise; thus we included three questions asking about a respondent's willingness to exercise and his/her perception of own physical ability. The fifth category consisted of 6 binary route choice experiments, which will be discussed in the next section. Finally, the sixth category consisted of questions regarding respondents' basic socio-demographic information.

We distributed the survey to students, faculty, and staff of the Cornell University's Ithaca campus during the spring semester in 2013. Response to the survey was successful as we received more than 600 responses within 2 weeks. Two versions of the discrete choice experiment led to two samples, the main sample (Sample 1) focused on cold temperatures, and the secondary sample (Sample 2) focused on warm and hot temperatures.

| Respondent characteristic | Total $N = 599$ | | Sample 1 $N_1 = 544$ | | Sample 2 $N_2 = 55$ | |
|---|---|---|---|---|---|---|
| | Total | % | Total | % | Total | % |
| Male | 250 | 42% | 220 | 40% | 30 | 55% |
| Access to bike (yes=1) | 323 | 54% | 291 | 53% | 32 | 58% |
| Advanced, confident cyclist | 127 | 21% | 110 | 20% | 30 | 55% |
| Intermediate cyclist | 195 | 33% | 176 | 32% | 19 | 35% |
| Cycling commute: never | 497 | 83% | 451 | 83% | 46 | 84% |
| Cycling commute: less than once a week | 32 | 5% | 31 | 6% | 1 | 2% |
| Cycling commute: 1-2 days a week | 29 | 5% | 31 | 6% | 1 | 2% |
| Cycling commute: 3-4 days a week | 18 | 3% | 18 | 3% | 0 | 0% |
| Cycling commute: 5+ days a week | 15 | 3% | 15 | 3% | 0 | 0% |
| Commute mostly by car | 66 | 11% | 58 | 11% | 8 | 15% |
| Commute mostly by bus | 168 | 28% | 145 | 27% | 23 | 42% |
| Live on campus | 164 | 27% | 156 | 29% | 8 | 15% |
| Distance to campus: within 1 mile | 276 | 46% | 253 | 47% | 23 | 42% |
| Distance to campus: 1-5 miles | 118 | 20% | 101 | 19% | 17 | 31% |
| Distance to campus: 5-10 miles | 18 | 3% | 13 | 2% | 5 | 9% |
| Age: 23-27 | 116 | 19% | 91 | 17% | 25 | 45% |
| Age: 28-40 | 78 | 13% | 70 | 13% | 8 | 15% |
| Age: 40+ | 54 | 9% | 51 | 9% | 3 | 5% |
| Exercise frequency: never | 37 | 6% | 37 | 7% | 0 | 0% |
| Exercise frequency: once a week | 90 | 15% | 79 | 15% | 11 | 20% |
| Exercise frequency: 2-3 times a week | 205 | 34% | 198 | 36% | 7 | 13% |
| Exercise frequency: daily | 128 | 21% | 121 | 22% | 7 | 13% |
| Undergraduate student | 350 | 58% | 335 | 62% | 15 | 27% |
| Graduate student | 184 | 31% | 148 | 27% | 36 | 65% |
| Faculty | 36 | 6% | 36 | 7% | 0 | 0% |
| Staff | 29 | 5% | 25 | 5% | 4 | 7% |

The discrete choice experiment is the key component of the survey. The experiments were based on binary route choice for bicycling, and the final experimental attributes and levels were decided after the results of the focus groups.

To examine effects of weather on a respondent's choice of route we included information regarding weather conditions of the day (``sun", ``rain", and ``snow"), including temperature and expected depth of precipitation in inches (for rain and snow). In fact, each choice situation started with a screenshot of the weather conditions similar to how information is displayed in smartphones (Figure 3-1).

The experimental route attributes were travel time, slope (grade), presence of a bike lane, and traffic volume. The slope of the route was described with a triangle sign indicating the percent grade. Text descriptions of the route grade levels were ``flat surface", ``moderate slope", and ``steep slope". Traffic volume was presented as text (``light" and ``heavy") and in the form of a road sign. Additionally, we presented pictures describing the two routes for each choice situation in the experiment (Figure 3-1). Table 3-1 summarizes the experimental attribute and attribute levels. These levels were combined using a D-efficient design with 9 choice situations.

Respondents were asked to choose between the two unlabeled routes, but we also included an ``opt-out" option in each choice situation to represent the decision not to bicycle.

**Table 3-1:** Attribute Levels

| Attribute | Attribute levels | Attribute | Attribute levels |
|---|---|---|---|
| Weather | 1. Sunny<br>2. Rain<br>3. Snow | Bike lane | 1. Yes<br>2. No |
| Hilliness | 1. 0 %<br>2. 10 %<br>3. 20 % | Temperature | 1. Cold<br>2. Moderate<br>3. Warm |
| Time | 1. 10 minutes<br>2. 15 minutes<br>3. 20 minutes | Traffic | 1.Heavy<br>2. Light |

**Figure 3-1**


***Structural Equation Modeling***

Structural equation modeling (SEM) is a modeling technique that can handle a large number of endogenous and exogenous variables, as well as unobserved (latent) variables specified as linear combinations of the observed variables. Regression, simultaneous equations, path analysis, and variations of factor analysis can be considered as special cases of SEM. SEM provides a way to test the specified set of relationships among observed and latent variables as a whole, and allow theory testing even when experiments are not possible. SEM has been applied in many research fields such as psychology, sociology, educational research, political science, and market research. Use of SEM has rapidly expanded as user-friendly software, such as Amos and Mplus, has become available. Several SEM applications in travel behavior research have been conducted in the past (e.g., Tardiff, 1976; Allaman et al., 1982).

A SEM with latent variables is usually composed of a measurement model for the endogenous variables and a structural model. However, a SEM model can consist of a structural model without any measurement models (if all variables are observable), or a measurement model alone (as in factor analysis). In SEM, the primary measure of parsimony is the degrees of freedom of the model, which equal to the difference between the number of free parameters in the model and the number of known quantities. The number of known quantities in covariance analysis is equal to the number of free elements in the variance–covariance matrix of the variables. Constructing a SEM consists in specifying an over-identified model in which only some of the possible parameters are free and many are restricted to zero, such that the model is a reasonable representation of the phenomena under study. Theory and good sense must guide model specification.

Estimation of SEM is performed using the covariance analysis method, and a Chi-square test can be used to determine if a model specified by the researcher is consistent with the variance–covariance in the data. Approximate fit indices (AFI) based on the direct comparison of the sample and model implied variance–covariance matrices are usually used for checking the model's fit. These include: (1) The root mean square residual of approximation (RMSEA), (2) the goodness-of-fit index (GFI). The use of approximate fit (AFI) indices has been the common practice among researchers to check model adequacy in SEM; however, the reliance of AFIs has recently become the target of a sharp criticism (Barrett, 2007, McIntosh and Markland, 2007).

In this study, we are interested in examining how a person's skills and experiences in bicycling can be represented as a latent factor and how that factor would affect his or her decision to commute by bicycle. Our methodology differs from the past studies in that instead of defining person's "bicyclist status" by a discrete threshold on a certain scaled measures, we estimated "bicycle status" as a continuous latent variable using multiple measurements.

The definition of the latent variables is based on the attitudinal responses to the following questions:

1. What encourages/would encourage you to ride a bike to/ from campus?
   a. Dedicated bike lanes on roads
   b. Bike pathways physically separated from the roadway
   c. Regulating car traffic on roads
   d. A campus map showing bicycle routes
   e. More convenient bike parking
   f. More secure or covered bike parking
   g. A convenient place to shower/ change clothes
   h. A bicycle station on campus providing repairs/ supplies

i. Bike racks on buses
j. Priority given by law to use road over vehicular traffic

2. Please indicate whether you agree or disagree with the following statements:
   a. Motor vehicle drivers seem to care little about bikers on road
   b. Bicyclists seem to care little about vehicular traffic on road
   c. Bicyclists seem to care little about pedestrians on street
   d. I do not like to share road with bikers when I am driving
   e. I have strong motivation to exercise.
   f. I am confident about my physical fitness
   g. I enjoy outdoor activities (camping, fishing, jogging, etc.)

3. What keeps/would keep you from riding a bicycle more to/ on campus?
   a. I am not interested in biking
   b. I live too far
   c. I need to change clothes
   d. Lack of adequate bicycle parking
   e. I am worried about accidents
   f. I do not feel safe about the possibility of crime
   g. Lack of bike lanes on road
   h. I am worried about possible mechanical problems that may occur, such as a flat tire
   i. Severe weather conditions
   j. Poor road conditions

| What encourages/would encourage you to ride a bike to/ from campus? | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| # | Question | Strongly Disagree | Disagree | Neither Agree nor Disagree | Agree | Strongly Agree | Total Responses | Mean |
| 1 | Dedicated bike lanes on roads | 11 | 40 | 133 | 336 | 119 | 639 | 3.80 |
| 2 | Bike pathways physically separated from the roadway | 9 | 38 | 117 | 283 | 192 | 639 | 3.96 |
| 3 | Regulating car traffic on roads | 17 | 86 | 249 | 225 | 62 | 639 | 3.36 |
| 4 | A campus map showing bicycle routes | 30 | 128 | 213 | 214 | 54 | 639 | 3.21 |
| 5 | More convenient bike parking | 23 | 114 | 227 | 222 | 53 | 639 | 3.26 |
| 6 | More secure or covered bike parking | 18 | 103 | 188 | 235 | 95 | 639 | 3.45 |
| 7 | A convenient place to shower/ change clothes | 29 | 135 | 210 | 200 | 65 | 639 | 3.21 |
| 8 | A bicycle station on campus providing repairs/ supplies | 22 | 109 | 200 | 232 | 76 | 639 | 3.36 |
| 9 | Bike racks on buses | 25 | 126 | 237 | 196 | 55 | 639 | 3.20 |
| 10 | Priority given by law to use road over vehicular traffic | 30 | 106 | 241 | 197 | 65 | 639 | 3.25 |

| Please indicate whether you agree or disagree with the following statements. | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| # | Question | Strongly Disagree | Disagree | Neither Agree nor Disagree | Agree | Strongly Agree | Total Responses | Mean |
| 1 | Motor vehicle drivers seem to care little about bikers on road | 15 | 144 | 199 | 230 | 51 | 639 | 3.25 |
| 2 | Bicyclists seem to care little about vehicular traffic on road | 21 | 147 | 203 | 221 | 47 | 639 | 3.20 |
| 3 | Bicyclists seem to care little about pedestrians on street | 24 | 183 | 219 | 165 | 48 | 639 | 3.05 |
| 4 | I do not like to share road with bikers when I am driving | 49 | 153 | 150 | 212 | 75 | 639 | 3.17 |
| 5 | I have strong motivation to exercise. | 16 | 93 | 152 | 247 | 131 | 639 | 3.60 |
| 6 | I am confident about my physical fitness | 12 | 115 | 163 | 238 | 111 | 639 | 3.50 |
| 7 | I enjoy outdoor activities (camping, fishing, jogging, etc.) | 10 | 44 | 119 | 265 | 201 | 639 | 3.94 |

| What keeps/would keep you from riding a bicycle more to/ on campus? | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| # | Question | Strongly Disagree | Disagree | Neither Agree nor Disagree | Agree | Strongly Agree | Total Responses | Mean |
| 1 | I am not interested in biking | 192 | 173 | 120 | 104 | 50 | 639 | 2.45 |
| 2 | I live too far | 242 | 217 | 75 | 69 | 36 | 639 | 2.12 |
| 3 | I need to change clothes | 103 | 150 | 132 | 205 | 49 | 639 | 2.92 |
| 4 | Lack of adequate bicycle parking | 157 | 216 | 166 | 91 | 9 | 639 | 2.34 |
| 5 | I am worried about accidents | 98 | 167 | 112 | 193 | 69 | 639 | 2.95 |
| 6 | I do not feel safe about the possibility of crime | 174 | 253 | 138 | 62 | 12 | 639 | 2.19 |
| 7 | Lack of bike lanes on road | 82 | 133 | 141 | 214 | 69 | 639 | 3.09 |
| 8 | I am worried about possible mechanical problems that may occur, such as a flat tire | 161 | 248 | 141 | 77 | 12 | 639 | 2.27 |
| 9 | Severe weather conditons | 30 | 60 | 72 | 280 | 197 | 639 | 3.87 |
| 10 | Poor road conditions | 65 | 124 | 152 | 208 | 90 | 639 | 3.21 |

Many specifications were tested, but we present the results of a Multiple Indicator and Multiple Causes (MIMIC), which is a confirmatory factor analysis model with explanatory variables (causal indicators), with three latent variables. The underlying concepts are: (1) bicycle status; (2) bike anxiety; and (3) physical condition. Table 3-2 displays the estimates of the MIMIC model that we used as final specification for this project.

**Table 3-2:** MIMIC model estimates

| Measurement Equation: Effect Indicator (R squared) | estimate | s.e. | t-stat | p-value |
|---|---|---|---|---|
| Bicycle Status | Reliability: 0.755 ($\alpha$), 0.765 ($\omega$) | | | |
| Frequency of cycle commuting (0.803) | 1.000 | - | - | - |
| Frequency of recreational cycling (0.449) | 0.661 | 0.051 | 12.893 | 0.000 |
| Self-evaluation as a cyclist (0.661) | 0.861 | 0.056 | 15.329 | 0.000 |
| Interest in cycling (0.501) | 0.711 | 0.050 | 14.277 | 0.000 |
| Bike Anxiety | Reliability: 0.728 ($\alpha$), 0.733 ($\omega$) | | | |
| Worried about accidents (0.544) | 1.000 | - | - | - |
| Worried about mechanical problems, flat tire (0.518) | 0.975 | 0.057 | 17.071 | 0.000 |
| Worried about crime (0.453) | 0.908 | 0.055 | 16.486 | 0.000 |
| Physical Condition | Reliability: 0.778 ($\alpha$), 0.780 ($\omega$) | | | |
| Strong motivation to exercise (0.834) | 1.000 | - | - | - |
| Confidence about physical fitness (0.722) | 0.898 | 0.050 | 17.880 | 0.000 |
| Enjoy outdoor activities (0.436) | 0.644 | 0.036 | 17.685 | 0.000 |

| Structural Equation: Causal Indicator | estimate | s.e. | t-stat | p-value |
|---|---|---|---|---|
| Bicycle Status (0.578) | | | | |
| Male | 0.355 | 0.090 | 3.927 | 0.000 |
| Age: 23-27 | 0.341 | 0.120 | 2.842 | 0.004 |
| Age: 28-40 | 0.488 | 0.147 | 3.318 | 0.001 |
| Age: 40+ | 0.455 | 0.164 | 2.771 | 0.006 |
| Distance to campus: 1-5 miles | 0.278 | 0.121 | 2.306 | 0.021 |
| Latent Physical Condition | 0.211 | 0.046 | 4.609 | 0.000 |
| Latent Bike Anxiety | -0.427 | 0.062 | -6.919 | 0.000 |
| Access to bike | 0.989 | 0.102 | 9.692 | 0.000 |
| Commute mostly by car | -0.389 | 0.135 | -2.872 | 0.004 |
| Commute mostly by bus | -0.239 | 0.098 | -2.435 | 0.015 |
| Bike Anxiety (0.105) | | | | |
| Male | -0.306 | 0.074 | -4.122 | 0.000 |
| Age: 23-27 | -0.204 | 0.105 | -1.950 | 0.051 |
| Age: 28-40 | -0.295 | 0.123 | -2.404 | 0.016 |
| Access to bike | -0.213 | 0.074 | -2.884 | 0.004 |
| Commute mostly by car | 0.173 | 0.118 | 1.471 | 0.141 |
| Commute mostly by bus | 0.174 | 0.084 | 2.068 | 0.039 |
| Physical Condition (0.426) | | | | |
| Male | 0.195 | 0.088 | 2.222 | 0.026 |
| Exercise frequency: never | -0.535 | 0.185 | -2.888 | 0.004 |
| Exercise frequency: once a week | 0.568 | 0.147 | 3.863 | 0.000 |
| Exercise frequency: 2-3 times a week | 1.155 | 0.126 | 9.137 | 0.000 |
| Exercise frequency: daily | 1.889 | 0.145 | 13.015 | 0.000 |

| Fit | |
|---|---|
| p-value (Chi-square) | 0.000 |
| Comparative Fit Index (CFI) | 0.941 |
| Tucker-Lewis Index (TLI) | 0.927 |
| Root Mean Square Error of Approximation | 0.057 |
| p-value RMSEA ≤ 0.05 | 0.049 |

The latent bicycle status summarizes the cycling skills and experience of the respondent, and is measured by the frequency of cycle for commuting, the frequency of recreational cycling, sel-evaluation as a cyclist by the respondent, and stated interest in cycling. The latent bike anxiety variable is measured by problems that may prevent the respondent from cycling, namely being worried about accidents, mechanical problems, and the possibility of crime. Finally, the latent physical condition of the respondent summarizes overall fitness and is manifested by the strength of the motivation to exercise, confidence about physical fitness, and stated satisfaction with outdoor activities.

The measurement equations reduce the dimensionality of the effect indicators, but it is the structural equations that provide a causal relationship that helps to explain how the latent variables are built. Bicycle status is explained by gender, age, distance to campus, access to bike, commute mostly by car, commute mostly by bus, physical condition, and bike anxiety. For example, a higher bicycle status is expected for men, people living 1-5 miles from campus (a shorter distance encourages walking and more than 5 miles discourages cycling), respondents having access to a bike, and for those having a higher (latent) physical condition. Commuting by motorized modes and a higher degree of (latent) bike anxiety reduce the bicycle status of the respondent. A higher level of bike anxiety is expected for individuals that prefer motorized modes, but for men and for people aged 23-40 years old bike anxiety is lower. Finally, the latent physical condition is explained by gender and the frequency of exercising.

### *Discrete/Hybrid Choice Modeling*

Discrete choice models describe decision makers' choices among alternatives (Train, 2003). They are usually derived under an assumption of utility-maximizing behavior by the decision maker. In random utility maximization (RUM) models, a decision maker $n$ faces a choice among $J$ alternatives. The decision maker obtains a certain level of utility or disutility from each alternative. A decision maker is assumed to choose the alternative that provides the largest utility among all alternatives in the choice set.

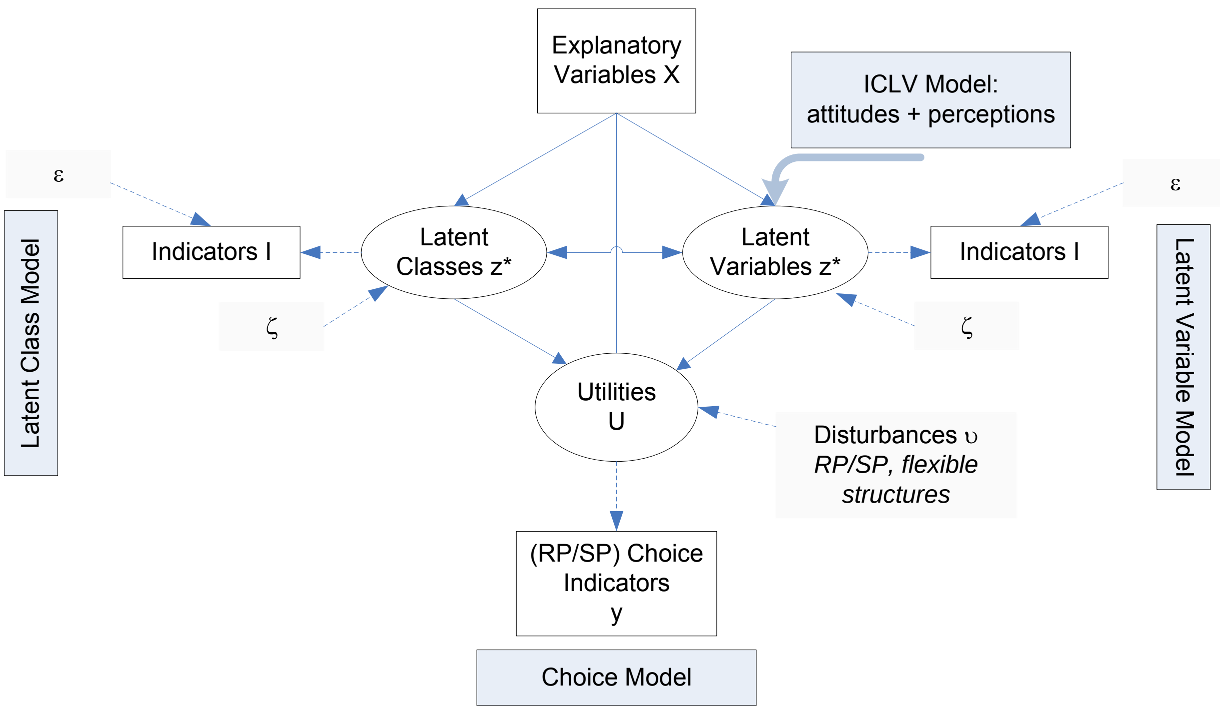
Under RUM, utility consists of a deterministic part ($V_{nj}$) and utility that is not observed ($\varepsilon_{nj}$). Therefore, the probability of individual $n$ choosing alternative $i$ can be written as:

$$P_{ni} = \mathrm{P}\left(\mathrm{U}_{ni} > \mathrm{U}_{nj}\right) = P\left(V_{ni} + \varepsilon_{ni} > V_{nj} + \varepsilon_{nj}\right) = P\left(\varepsilon_{ni} - \varepsilon_{nj} > V_{ni} - V_{nj}\right) \text{ for all j } \neq \text{i}$$

Different models can be derived by assuming different distributions for the error term $\varepsilon_{nj.}$

On the one hand, the conditional logit model –by far the most common discrete choice model– is derived by assuming that $\varepsilon_{nj}$ follows an extreme value type 1 distribution. On the other hand, hybrid choice models are a generalization of standard discrete choice models where different expanded models are considered simultaneously (Ben-Akiva et al., 2002). A hybrid choice model (HCM) expands on standard discrete choice modeling by considering the integration of latent (unobserved) constructs. The HCM framework permits the inclusion of latent attributes, which may be abstract hypothetical constructs, such as attitudes as well as qualitative attributes that do not have a natural order. By using HCM, we incorporated our SEM model above into the response to the route choice experiment.

**Figure 3-3**




General representation of the Hybrid Choice Model

For estimation purposes we worked with the validated sample of 599 observations, where each respondent answered 9 choice situations, for a total of 5391 pseudo-individuals. As described above, in each choice situation the respondent faced a hypothetical bicycle route selection problem with 2 alternatives, with the possibility of choosing not to bike.

In general, the utility function for each of the alternatives was specified as follows:

$$\begin{aligned}
V_A &= \beta_{bike} + \beta_{time} time_A + \beta_{slope} slope_A + \beta_{bike\ lane} bike\ lane_A + \beta_{traffic} traffic_A + \beta_{rain} rain \\
&\quad + \beta_{snow} snow + \beta_{T<75F} temperature(<75F) + \beta_{T\geq 75F} temperature(\geq 75F) \\
V_B &= \beta_{bike} + \beta_{time} time_B + \beta_{slope} slope_B + \beta_{bike\ lane} bike\ lane_B + \beta_{traffic} traffic_B + \beta_{rain} rain \\
&\quad + \beta_{snow} snow + \beta_{T<75F} temperature(<75F) + \beta_{T\geq 75F} temperature(\geq 75F) \\
V_O &= 0,
\end{aligned}$$

where $V_A$ represents the deterministic utility function of route A, $V_B$ represents that of route B, and $V_O$ represents the deterministic utility function of the opt-out alternative, which corresponds to “neither” option in the survey. The parameters provide the marginal effect of the attributes and weather-related variables on the utility valuation of the routes.

We first present the results of a standard conditional logit model. (Table 3-3.)

**Table 3-3:** standard conditional logit model

| Parameter | estimate | s.e. | t-stat | p-value | LB 95% CI | UP 95% CI |
|---|---|---|---|---|---|---|
| $\beta_{bike}$ | 2.4223 | 0.1996 | 12.1300 | 0.0000 | 2.0310 | 2.8136 |
| $\beta_{time}$ | -0.0568 | 0.0053 | -10.7400 | 0.0000 | -0.0671 | -0.0464 |
| $\beta_{slope}$ | -0.1633 | 0.0056 | -28.9600 | 0.0000 | -0.1744 | -0.1523 |
| $\beta_{bike\ lane}$ | 0.4044 | 0.0431 | 9.3900 | 0.0000 | 0.3200 | 0.4888 |
| $\beta_{traffic}$ | -0.8330 | 0.0418 | -19.9100 | 0.0000 | -0.9150 | -0.7510 |
| $\beta_{rain}$ | -0.6568 | 0.1030 | -6.3700 | 0.0000 | -0.8588 | -0.4549 |
| $\beta_{snow}$ | -0.9313 | 0.0477 | -19.5300 | 0.0000 | -1.0248 | -0.8379 |
| $\beta_{T<75F}$ | 0.0077 | 0.0044 | 1.7700 | 0.0770 | -0.0008 | 0.0163 |
| $\beta_{T\geq 75F}$ | -0.0055 | 0.0024 | -2.3000 | 0.0210 | -0.0102 | -0.0008 |
| Loglikelihood | -4732.3339 | | | | | |
| Adjusted ρ2 | 0.197 | | | | | |

The route characteristics time, slope, and traffic, as well as precipitation conditions (snow and rain) have a negative impact on the likelihood of bicycling. For instance, if a particular route has a steeper slope, then individuals will be less likely to choose that route, either by taking a different route or by choosing not to bike. If the amount of rain increases, then individuals are more likely to choose not to cycle (as rain affects all routes). The presence of a bike lane is appreciated and increases the probability of riding a bike. All coefficients, with the exception of colder temperatures, are statistically significant. However, the positive sign of the colder temperatures is an indication that individuals may be more likely to bike the more pleasant the temperature becomes, until it gets “too hot”.

Next, we estimated a hybrid choice model using a conditional logit model with latent variables. As a reminder, we identified three latent concepts, one that summarizes the physical condition of the respondents, another one that identifies in a single index their bicycle status, and a third one summarizing problems that may be encountered when cycling, labeled as bike anxiety. Both bicycle status and bike anxiety entered additively to the bike constant. For the latent physical condition, we considered an interaction with the slope of the route. The likelihood ratio test favors the hybrid choice model with respect to the standard conditional logit. The results of this hybrid choice model are shown in Table 3-4.

**Table 3-4:** conditional logit model with latent variables (HCM)

| Parameter | estimate | s.e. | t-stat | p-value | LB 95% CI | UP 95% CI |
|---|---|---|---|---|---|---|
| $\beta_{bike}$ | 2.1077 | 0.2049 | 10.2800 | 0.0000 | 1.7060 | 2.5094 |
| $\beta_{time}$ | -0.0569 | 0.0053 | -10.7400 | 0.0000 | -0.0672 | -0.0465 |
| $\beta_{slope}$ | -0.1862 | 0.0076 | -24.5500 | 0.0000 | -0.2010 | -0.1713 |
| $\rho_{slope\times phys.cond}$ | 0.0203 | 0.0043 | 4.7100 | 0.0000 | 0.0119 | 0.0288 |
| $\beta_{bike\ lane}$ | 0.4067 | 0.0432 | 9.4200 | 0.0000 | 0.3220 | 0.4913 |
| $\beta_{traffic}$ | -0.8365 | 0.0420 | -19.9300 | 0.0000 | -0.9187 | -0.7542 |
| $\beta_{rain}$ | -0.7053 | 0.1050 | -6.7200 | 0.0000 | -0.9111 | -0.4995 |
| $\beta_{snow}$ | -0.9638 | 0.0487 | -19.8100 | 0.0000 | -1.0592 | -0.8684 |
| $\beta_{T<75F}$ | 0.0079 | 0.0044 | 1.7900 | 0.0730 | -0.0007 | 0.0166 |
| $\beta_{T\geq 75F}$ | -0.0068 | 0.0024 | -2.7900 | 0.0050 | -0.0115 | -0.0020 |
| $\gamma_{bicycle\ status}$ | 0.3464 | 0.0400 | 8.6600 | 0.0000 | 0.2680 | 0.4248 |
| $\gamma_{bike\ anxiety}$ | 0.0451 | 0.0626 | 0.7200 | 0.4710 | -0.0775 | 0.1678 |
| Loglikelihood | -4658.8531 | | | | | |
| Adjusted ρ2 | 0.209 | | | | | |

The parameter estimates are in line with the findings of the standard logit model, but they provide additional insights. The positive interaction parameter $\rho_{slope\times phys.cond}$ indicates that the more fit the individual is, the lower the effect of the slope on his or her cycling decisions. In addition, the higher the bicycle status – which is a continuous measure of cyclists versus non cyclists – the more likely an individual is to ride the bike versus opting out. Bike anxiety has a counterintuitive positive sign, but the effect is not significant.

We also specified a mixed logit model with latent attributes. As discussed above, most empirical applications of hybrid choice models thus far have used a conditional logit kernel. This fact is due in part to availability of software, but also to numerical problems in finding the maximum likelihood estimator.

In Table 3-5 we present the estimates of a mixed logit model, where only the bike constant was assumed fixed. All other variables, with the exception of time, were assumed normally distributed. For time, we considered a log-normal distribution. Because of the assumptions we did over the heterogeneity distributions, we present the estimates of the mean and standard deviation for normally distributed parameters, and the mean, median, and standard deviation for the log-normally distributed time parameter.

**Table 3-5:** mixed logit model with latent variables (HCM)

| Parameter | estimate | s.e. | t-stat | p-value | LB 95% CI | UP 95% CI |
|---|---|---|---|---|---|---|
| Mean | | | | | | |
| $\beta_{bike}$ | 3.1853 | 0.3251 | 9.8000 | 0.0000 | 2.5482 | 3.8225 |
| $\beta_{time}$ | -0.0878 | 0.0073 | -11.9900 | 0.0000 | -0.1022 | -0.0735 |
| $\beta_{slope}$ | -0.2888 | 0.0153 | -18.9100 | 0.0000 | -0.3188 | -0.2589 |
| $\rho_{slope\times phys.cond}$ | 0.0336 | 0.0086 | 3.9300 | 0.0000 | 0.0168 | 0.0504 |
| $\beta_{bike\ lane}$ | 0.6389 | 0.0638 | 10.0200 | 0.0000 | 0.5139 | 0.7640 |
| $\beta_{traffic}$ | -1.2129 | 0.0652 | -18.6100 | 0.0000 | -1.3406 | -1.0852 |
| $\beta_{rain}$ | -1.4145 | 0.1604 | -8.8200 | 0.0000 | -1.7288 | -1.1002 |
| $\beta_{snow}$ | -2.0653 | 0.1270 | -16.2700 | 0.0000 | -2.3141 | -1.8164 |
| $\beta_{T<75F}$ | 0.0343 | 0.0085 | 4.0400 | 0.0000 | 0.0177 | 0.0509 |
| $\beta_{T\geq 75F}$ | -0.0052 | 0.0064 | -0.8100 | 0.4170 | -0.0178 | 0.0074 |
| $\gamma_{bicycle\ status}$ | -0.2436 | 0.2298 | -1.0600 | 0.2890 | -0.6940 | 0.2068 |
| $\gamma_{bike\ anxiety}$ | 0.6472 | 0.1612 | 4.0200 | 0.0000 | 0.3314 | 0.9631 |
| Median | | | | | | |
| $\beta_{time}$ | -0.0727 | 0.0073 | -9.9100 | 0.0000 | -0.0871 | -0.0583 |
| S.D. | | | | | | |
| $\beta_{time}$ | 0.0595 | 0.0060 | 9.8700 | 0.0000 | 0.0477 | 0.0714 |
| $\beta_{slope}$ | 0.1475 | 0.0118 | 12.5300 | 0.0000 | 0.1244 | 0.1706 |
| $\rho_{slope\times phys.cond}$ | 0.0067 | 0.0187 | 0.3600 | 0.7180 | -0.0299 | 0.0434 |
| $\beta_{bike\ lane}$ | 0.7342 | 0.0895 | 8.2000 | 0.0000 | 0.5588 | 0.9096 |
| $\beta_{traffic}$ | 0.7561 | 0.0903 | 8.3800 | 0.0000 | 0.5792 | 0.9330 |
| $\beta_{rain}$ | 0.5234 | 0.3439 | 1.5200 | 0.1280 | -0.1506 | 1.1974 |
| $\beta_{snow}$ | 1.4803 | 0.2329 | 6.3600 | 0.0000 | 1.0238 | 1.9368 |
| $\beta_{T<75F}$ | 0.0629 | 0.0055 | 11.5000 | 0.0000 | 0.0521 | 0.0736 |
| $\beta_{T\geq 75F}$ | 0.0285 | 0.0055 | 5.1600 | 0.0000 | 0.0177 | 0.0394 |
| $\gamma_{bicycle\ status}$ | -0.3366 | 0.3985 | -0.8400 | 0.3980 | -1.1176 | 0.4444 |
| $\gamma_{bike\ anxiety}$ | 0.8781 | 0.1127 | 7.7900 | 0.0000 | 0.6572 | 1.0990 |
| Loglikelihood | -3934.1438 | | | | | |
| Adjusted ρ2 | 0.332 | | | | | |

Again, most results are in line with the previous findings. All signs behave as expected. The marginal utility of colder temperatures now appears as being significantly different from zero, whereas warmer temperatures are no longer significant. However, the standard deviation of warmer temperatures is significant, indicating that some individual may prefer hotter temperatures, while others may think it is too hot for riding a bike. In addition, most of the estimates of the standard deviations are significant, and a likelihood ratio test rejects the previous models in favor of the mixed logit with latent variables. We conclude that the level of heterogeneity in tastes is high and significant.

One of the problems with the introduction of latent variables directly in the utility function is that the segmentation of users becomes rather difficult. With latent variables, latent clusters are continuous and based on an arbitrary measurement scale. The construction of segments is then complex, especially for those latent attributes entering linearly. Thus, in this project we propose to explore a discrete segmentation approach, where latent variables are used to identify the discrete segments of the population.

The latent class structure allows investigation into preference heterogeneity through the analysis of class membership, offering insights into a person's decision-making process (Olaru, et al., 2011). The assumption of latent class models is that individual heterogeneity depends on observable attributes and on latent heterogeneity that varies with factors that are unobserved by the researcher (Greene and Hensher, 2003). In a latent class model, individuals are sorted into $Q$ discrete classes; however, the class membership of each class is unknown to the researcher. The probability the respondent chooses alternative $i$ is conditional on belonging to a class $q$: $P_{nt|q}(j) = Prob(y_{nt} = j|class = q)$.

The unconditional choice probability is: $P_{jnt} = \sum_q P_{nt|q}(j)P_{nq}$, where $P_{nq}$ is the class assignment probability for individual $n$. In a latent class model, the marginal utility parameters are class-dependent. A convenient assumption for the class assignment model is the multinomial logit form:

$$P_{iq} = \frac{\exp(z_i'\theta_q)}{\sum_{q=1}^{Q} \exp(z_i'\theta_q)},$$

where $z_i$ are characteristics of respondents and $\theta_q$ are parameter estimates of membership. Consequently, the log-likehood is a product of probabilities:

$$lnL = \sum_{i=1}^{N} lnP_i = \sum_{i=1}^{N} \ln[\sum_{q=1}^{Q} P_{nq}(\sum_{t=1}^{T_n} P_{nt|q})]$$

We tested a hybrid choice model with latent classes based on

$$P_{iq} = \frac{\exp(z_i^{*\prime}\theta_q)}{\sum_{q=1}^{Q} \exp(z_i^{*\prime}\theta_q)},$$

where $z_i^*$ is a vector that contains latent characteristics of the individual. In particular, our hypothesis is that bicycle behavior depends on the different skills and experience of the users, which are summarized by the latent bicycle status variable. Results of the latent class model are reported in Table 3-6 below.

**Table 3-6:** latent class model with latent variables

| Parameter | estimate | s.e. | t-stat | p-value | LB 95% CI | UP 95% CI |
|---|---|---|---|---|---|---|
| Class 1 | | | | | | |
| $\beta_{bike}$ | 3.0771 | 0.6263 | 4.9100 | 0.0000 | 1.8495 | 4.3047 |
| $\gamma_{bike\ anxiety}$ | -0.3237 | 0.1762 | -1.8400 | 0.0660 | -0.6691 | 0.0218 |
| $\beta_{time}$ | -0.0601 | 0.0069 | -8.7500 | 0.0000 | -0.0735 | -0.0466 |
| $\beta_{slope}$ | -0.1675 | 0.0106 | -15.8600 | 0.0000 | -0.1882 | -0.1468 |
| $\rho_{slope \times phys.cond}$ | 0.0259 | 0.0063 | 4.1300 | 0.0000 | 0.0136 | 0.0383 |
| $\beta_{bike\ lane}$ | 0.4338 | 0.0585 | 7.4200 | 0.0000 | 0.3192 | 0.5484 |
| $\beta_{traffic}$ | -0.8675 | 0.0569 | -15.2600 | 0.0000 | -0.9790 | -0.7561 |
| $\beta_{rain}$ | -0.8492 | 0.3243 | -2.6200 | 0.0090 | -1.4848 | -0.2136 |
| $\beta_{snow}$ | -0.9785 | 0.1669 | -5.8600 | 0.0000 | -1.3056 | -0.6513 |
| $\beta_{T<75F}$ | 0.0479 | 0.0234 | 2.0500 | 0.0410 | 0.0020 | 0.0937 |
| $\beta_{T\geq75F}$ | 0.0090 | 0.0102 | 0.8900 | 0.3740 | -0.0109 | 0.0290 |
| Class 2 | | | | | | |
| $\beta_{bike}$ | 2.2629 | 0.3236 | 6.9900 | 0.0000 | 1.6286 | 2.8972 |
| $\gamma_{bike\ anxiety}$ | -0.4922 | 0.0944 | -5.2100 | 0.0000 | -0.6772 | -0.3071 |
| $\beta_{time}$ | -0.0337 | 0.0131 | -2.5600 | 0.0100 | -0.0594 | -0.0079 |
| $\beta_{slope}$ | -0.2683 | 0.0170 | -15.7800 | 0.0000 | -0.3017 | -0.2350 |
| $\rho_{slope \times phys.cond}$ | 0.0251 | 0.0085 | 2.9600 | 0.0030 | 0.0085 | 0.0417 |
| $\beta_{bike\ lane}$ | 0.3973 | 0.0982 | 4.0400 | 0.0000 | 0.2047 | 0.5898 |
| $\beta_{traffic}$ | -0.9745 | 0.0885 | -11.0200 | 0.0000 | -1.1480 | -0.8011 |
| $\beta_{rain}$ | -1.1527 | 0.1659 | -6.9500 | 0.0000 | -1.4778 | -0.8276 |
| $\beta_{snow}$ | -2.1573 | 0.2465 | -8.7500 | 0.0000 | -2.6403 | -1.6742 |
| $\beta_{T<75F}$ | -0.0096 | 0.0059 | -1.6500 | 0.0990 | -0.0211 | 0.0018 |
| $\beta_{T\geq75F}$ | -0.0215 | 0.0038 | -5.6700 | 0.0000 | -0.0289 | -0.0141 |
| Class Assignment | | | | | | |
| $\theta_{constant}$ | -0.2870 | 0.1556 | -1.8400 | 0.0650 | -0.5920 | 0.0180 |
| $\theta_{bicycle\ status}$ | 0.2447 | 0.0883 | 2.7700 | 0.0060 | 0.0717 | 0.4177 |
| Loglikelihood | -4081.6003 | | | | | |
| Adjusted ρ2 | 0.307 | | | | | |

The class assignment model uses class 2 as baseline, meaning that the positive $\theta_{bicycle\ status}$ can be interpreted as class 1 being the segment of more experienced cyclists. In effect, the higher the latent bicycle status is, the higher the probability of an individual being an experience cyclist. As a result, all the choice parameters for class 1 correspond to the valuations of more experienced cyclists, whereas those of class 2 represent the valuation of non-cyclists.

The effects in both classes are similar, but the sensitivities are different. Longer travel times, steeper slopes, the presence of heavy traffic, and rain and snow still bother individuals and make cycling a less attractive alternative. Bike lanes are still appreciated. The effect of colder temperatures is significant for class 1, while the effect of warmer temperatures is significant for class 2. For both classes, a higher latent physical condition results in a lower sensitivity to slopes. Finally, the effect of the latent bike anxiety is not significant for class 1, but is significant and negative for class 2. This last result makes sense, as class 2 is the segment of less experienced cyclists who are more concerned about the problems that may arise when cycling.

To analyze the taste differences between the two classes, we calculated the ratio of the marginal rates of substitution of each variable with respect to time. These ratios are presented in Table 3-7.

**Table 3-7: ratio of the marginal rate of substitution with respect to travel time for class 2 versus class 1**

| Variable | MRS ratio |
|---|---|
| Slope | 2.859 |
| Bike Lane | 1.634 |
| Traffic | 2.004 |
| Rain | 2.422 |
| Snow | 3.934 |

These results show that for less experienced cyclists, the degree of steepness of a slope affects almost 3 times more than for experienced cyclists, the presence of traffic affects twice as much, rain 2.4 times as much as snow almost 4 times more. An interesting result is that bike lanes are more appreciated by non-cyclists. (The benefit of the presence of bike lanes is valued 1.6 times higher by non-cyclists with respect to the valuation of cyclists.) This bike-lane appreciation is consistent with the hypothesis that bicyclists care less about the availability of bike lanes because they have higher skills than non-bicyclists. (e.g., Taylor & Mahmassani, 1996; Hunt & Abraham 2007).

# 4. TIME-SERIES ANLYSIS OF BICYCLE COUNT DATA

## 4.1 Introduction

There has been considerable growth in bicycle ridership over the past few decades – the total number of bike trips in the U.S. more than tripled between 1977 and 2009, while the bike share of total trips almost doubled, rising from 0.6% to 1.0% (Pucher, et al., 2011). The rising bicycle ridership suggests that careful planning and appropriate investment in bicycle infrastructure are necessary in order to accommodate increasing bicycle travel. Since accurate ridership prediction will be indispensable for those planning and investment decisions, it is important to understand various ridership prediction models which are capable of identifying significant factors related to the motivation for people to bicycle. This chapter examines the forecasting capability of various econometric models for bicycle counts.

In past studies, bicycle counts data have been primarily treated as cross-sectional. As is often the case with traffic count analysis, Poisson and Negative Binomial (NB) models have been used to analyze cross-sectional and time-series count data of bicycle riders. For example, Nasal and Miranda-Moreno (2011) estimated both a Poisson regression model and a Negative Binomial regression model using hourly ridership count and found that recreational facilities are more affected by weather conditions than utilitarian facilities in general, and non-recreational facilities are more sensitive to weather conditions on weekends. However, in their analysis, an independent and identically distributed error term was assumed – an assumption unlikely to hold in time series data. In addition, the data was treated as cross-sectional without checking for stationarity, ignoring thus the dynamics of cycling ridership, trends, seasonality, or within-week cycles. Modeling time series count data using Poisson regression or NB regression may result in inefficient parameter estimates as the time series data are often serially correlated. An appropriate prediction model is necessary in order to account for the serial correlation that exists in bicycle ridership. The key objective of this study therefore is to examine the effects of weather on bicycle ridership using cross-sectional count models and time-series models, and to examine their relative performances. We will discuss the strengths and weaknesses of these models and then examine the performance of alternative models, namely the so-called "state space" models. The rest of this chapter is organized as follows. In the next section, we will describe the econometric models we used for our analysis. We will then explore in detail the bicycle count data collected using an inductive loop bicycle counter in Portland, Oregon. We will present our results and discuss the limitations of this study and the directions for the future research.

## 4.2 Methodology

### *Models for Count Data*

The models for integer event count data are well developed and applied in various fields. Cross-sectional count data are often modeled using a Poisson regression model or a Negative Binomial regression if over-dispersion is present. When a Poisson model is appropriate for an outcome Y, the probabilities of observing any specific count, y, are given by the formula:

$$P(Y = y) = \frac{\lambda^y e^{-y}}{y!}$$

where λ is known as the population rate parameter (which usually needs to be estimated). A

Poisson random variable Y has expectation E(Y) = λ, and variance var(Y) = λ. The fact that the expectation and variance coincide provides a quick check on whether a Poisson model might be appropriate for a sample of observations. The Poisson regression is simply an extension of the Poisson model, where parameters for covariates are estimated to describe the relationship between covariates (e.g., weather variables) and responses covariates (e.g., ridership counts). This relationship can be parameterized by a log-linear model,

$$\log(\lambda(X_1, X_2, \ldots, X_n) = \mathrm{a} + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n$$

$$E(\lambda) = var(\lambda) = \mu$$

The Poisson regression has the severe limitation that the variance of the counts as well as that of explanatory variables is equal to the mean. If this fails to be true (i.e., count data shows "over-dispersion"), the estimates of the coefficients can still be consistent, but the standard errors can be biased. The Negative Binomial regression (NB) is an extension of the Poisson regression that can account for over-dispersion among variables. More specifically, the NB model will estimate an additional parameter, θ (dispersion parameter), for the variance:

$$var(\lambda) = \mu + \theta\mu^2$$

The Negative Binomial regression is often more appropriate than the Poisson regression when the data is observational as we would not expect that every variable that contributes to the rates of events is measured, and so there will always be residual variation. We will therefore use a Negative Binomial regression for our analysis, which is specified as follows:

$$\log(\mu_{m,d}) = \alpha + \beta X_{m,d} + \gamma + \epsilon_{m,d},$$

where

m, d = indexes representing the month and day of the week, respectively;
$\mu_{m,d}$ = mean number of bicycle counts during a specific month m, and day of the week d;
$X_{m,d}$ = weather conditions (precipitation, temperature, and dew point depression)
ϒ = a dummy variable for weekend days
$\epsilon_{m,\,d}$ = independent error term

***Models for Time Series Data***

The models for continuous autoregressive time series data were introduced by Box and Jenkins (1976), and have been applied in different fields such as finance and economics. The Box and Jenkins model such as the seasonal autoregressive integrated moving average (SARIMA) model is capable of taking into account the trend and seasonality (and hence the serial correlation) normally present in time series data. An extension to this model was proposed by Box and Tiao (1975) by adding the ability to examine the effects of various regressors (or/and interventions) as explanatory variables along with the trend and seasonal components. This model, called Auto Regressive Integrated Moving Average with Exogenous Input (ARIMAX), can be expressed as follows:

$$Y_t = \beta X_t + E_t$$

where $Y_t$ = the dependent variable for a particular time t;
$X_t$ = the deterministic effects of independent variables;
$E_t$ = the stochastic variation;

The stochastic variation can be represented by an ARIMA model as ARIMA (p,d,q) (for a non-seasonal time series), or a SARIMA model (for a seasonal time series) denoted as SARIMA $(p,d,q)\times(P,D,Q)_S$. In these models, p is the order of the non-seasonal autoregressive (AR) process, P is the order of the seasonal AR process, d is the order of the non-seasonal difference, D is the order of the seasonal difference, q is the order of the non-seasonal moving average (MA) process, Q is the order of the seasonal MA process and the subscripts is the length of seasonality (Box et al., 2008).

***State Space Models***

Recently there has been an increasing interest in the application of state space models in time series analysis. One major drawback of ARIMA models is the requirement of stationarity. The analysis on nonstationary time series requires a preliminary transformation of the data to get stationarity. The stationarity requirement becomes problematic in two ways: (1) stationarity may be difficult to attain through differencing; and (2) the coefficient estimates of regressors, if introduced in the model, are difficult to interpret. Multiple differencing is sometimes required for the series to exhibit stationarity, and additionally, the criterion for the series to be stationary is arbitrarily based on statistical tests such as Dickey-Fuller test. Inclusion of explanatory variables can provide a greater explanatory power to the model; however, since most series require differencing, the interpretation of the physical meanings of coefficient estimates of differenced variables can be arbitrary. The state space models solve all of those problems.

State space models allow a direct analysis on data that exhibits non-stationarity. In the state space model, the development of the system, $y_t$, is determined by unobserved series of states, $\theta_t$, whose relation with $y_t$ is specified by model. In general, a dynamic linear state space model is written as:

$$\begin{array}{lll} Observation\ Equation: & Y_t = F_t\theta_t + v_t, & v_t \sim N_m(0, V_t) \\ State\ Equation: & \theta_t = G_t\theta_{t-1} + w_t, & w_t \sim N_p(0, W_t) \end{array}$$

The state vector, $\theta_t$ is specified by a prior distribution. For example, a Normal prior distribution for the k-dimensional state vector at t = 0 is $\theta_t \sim N_m(0, C_t)$. In the classical approach, the estimation of a vector of unknown parameters is done by maximum likelihood. However, we will apply the Bayesian approach as it offers a more consistent formulation of the problem (Petris and Campagnoli, 2009). Estimation of unknown parameters is solved by computing conditional distributions of the quantities of interest given the most recent data using filtering. Filtering is the recursive steps needed to compute the densities p(θt|Yt) in the state space model. In the filtering problem, it is assumed that the data arrives sequentially in time and the object of filtering is to update our knowledge of the system each time new data arrives. Filtering involves the following steps:

- One step ahead predictive distribution for $\theta_t$ given $Y_{t-1}$, based on the filtering density $p(\theta_{t-1}|Y_{t-1})$:

$$p(\theta_t|Y_{t-1}) = \int p(\theta_t|\theta_{t-1})p(\theta_t|Y_{t-1})\, d\theta_{t-1}$$

- One step ahead predictive distribution for the next observation:

$$p(y_t|Y_{t-1}) = \int p(y_t|\theta_t)p(\theta_t|Y_{t-1})\, d\theta_t$$

- The posterior distribution $\pi(\theta_t|y_{1:t})$ using the prior distribution $p(\theta_t|Y_{t-1})$ and the likelihood $p(y_t|\theta_t)$:

$$p(\theta_t|Y_t) = \frac{p(y_t|\theta_t)p(\theta_t|Y_{t-1})}{p(y_t|Y_{t-1})}$$

The Kalman filter allows us to compute the predictive and filtering distributions recursively, using $\theta_0$ ~N($u_0$, $v_0$) to compute $p(\theta_1|y_1)$, and proceeding recursively as new data become available.

In this study, we will apply the random walk plus noise model:

$$\begin{aligned} Y_t &= \alpha_t + v_t, \qquad & v_t \sim N_m(0, {\sigma_v}^2) \\ \mu_t &= \alpha_{t-1} + w_t, \qquad & w_t \sim N_p(0, {\sigma_w}^2) \end{aligned}$$

where $v_t$ and $w_t$ are all mutually independent and independent of $\alpha_t$. The prior distribution of $\alpha_t$ is assumed to be d-inverse Gamma, and the Bayes estimator is based on Markov Chain Monte Carlo (MCMC) sampling. In theory, Bayes estimators are derived from minimizing the posterior Bayes risk function. However, in practice only few problems have a closed-form expression fo the Bayes risk. MCMC methods exploit a computational approach to approximate the posterior of interest.

**4.3 Cyclist Count Data Analysis**

The bicycle count data was collected in the Hawthorne Bridge in Portland, Oregon using an inductive loop bicycle counter. The counter detects bicycles by monitoring changes in an electric current in sub-pavement loops of cable and is capable of distinguishing cyclists from other traffic (Nasal and Miranda-Moreno, 2011). The data was acquired between 5 am and 7 pm each day from April to November in 2010. As shown in Figure 4-1, the Hawthorne Bridge is equipped with a cycle track and carries commuters from the east side of the city into the downtown area. The City of Portland has a population of about 1.6 million with a relatively high percentage of bicycle commuting (2.6 %) compared to other large US cities (Dill & Carr, 2003).

Figure 4-2 shows the average hourly counts of bicycle ridership observed at the Hawthorne Bridge. The ridership peaks twice a day, in the morning (8am) and in the evening (5pm), which indicates that the facility is used primarily for utilitarian purposes. Figure 4-2 shows the average daily counts and the average monthly counts of bicycle ridership. The ridership stays relatively flat on weekdays reaching its peak on Wednesday and dramatically descends on weekend, which further indicates that the bridge is used as a utilitarian facility. Figure 4-3 shows the average daily counts over months. The ridership gradually increases toward warmer months, achieving its peak in August and gradually decreases afterward. A drastic drop in the ridership is observed between October and November. To illustrate the cause of this drop in the ridership, Figure 4-4 shows the average temperature and precipitation in Portland between April and November of 2010. As the figure shows, the temperature and precipitation change dramatically in November. In particular, the temperature decreases rapidly from October to November.

Based on the data described above, we will estimate the following four models: SARIMA, Negative Binomial, SARIMAX, and the State Space model with a random walk plus noise to compare their fits and forecasting accuracy. As described above, the month of November in Portland is characterized by much colder weather with a high level of precipitation and much lower ridership (see Figures 4-5 and 4-6); therefore we will first use data from April to September to make out-of-sample predictions for October. The issue associated with the presence of inclement weather in November will be further discussed later on.

**Figure 4-1** Hawthorne Bridge in Portland, OR


**Figure 4-2** Average cyclists per hour

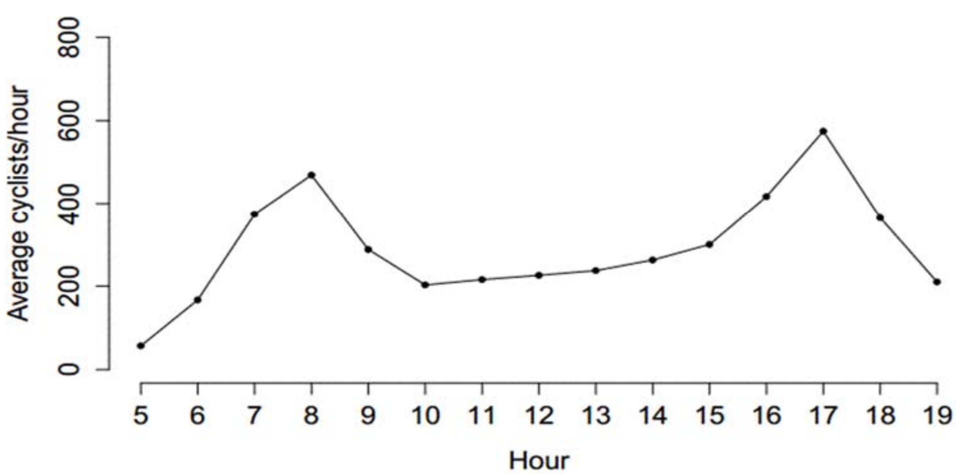


**Figure 4-3** Average cyclists per hour per day, per day of the week (left) and month (right)

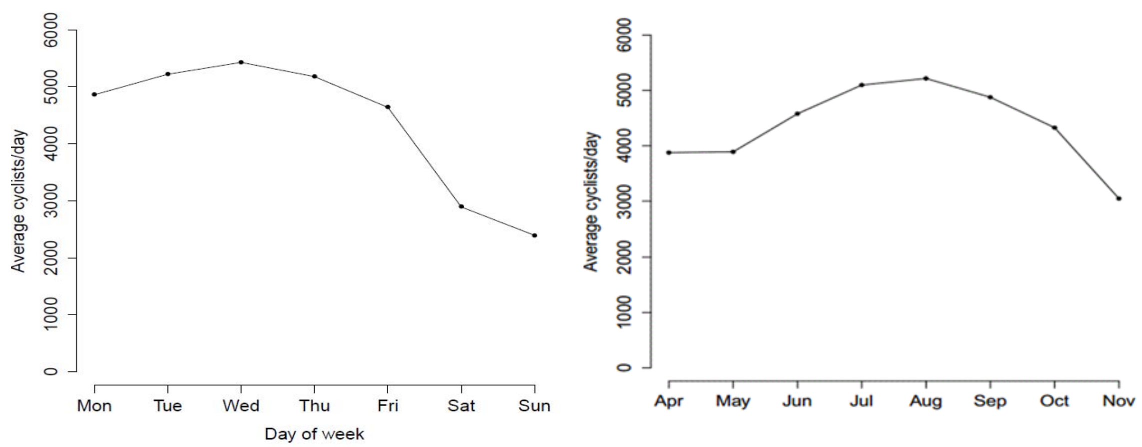



**Figure 4-4** Average monthly precipitation (left) and temperature (right)

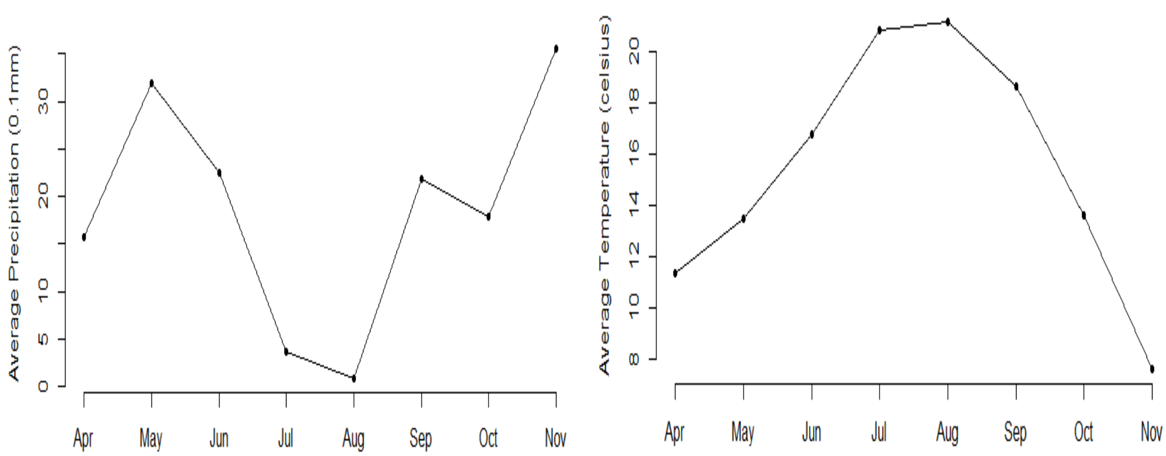



**Figure 4-5** Average ridership per day, during the month of November

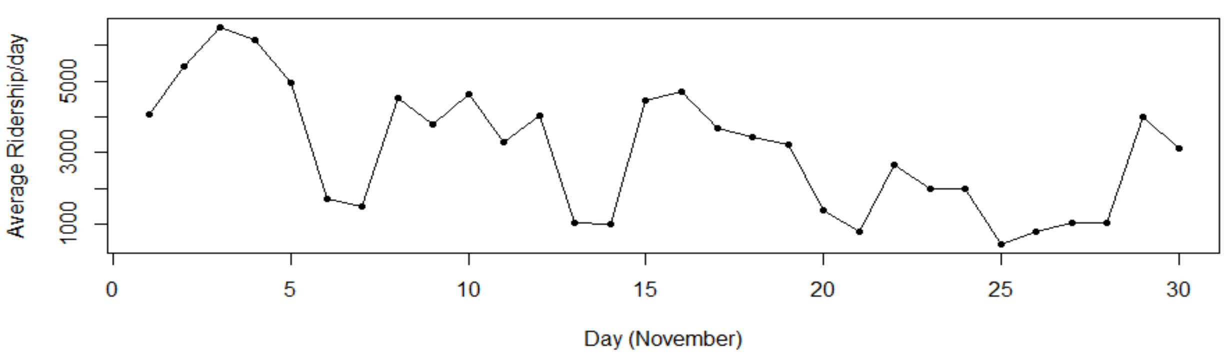
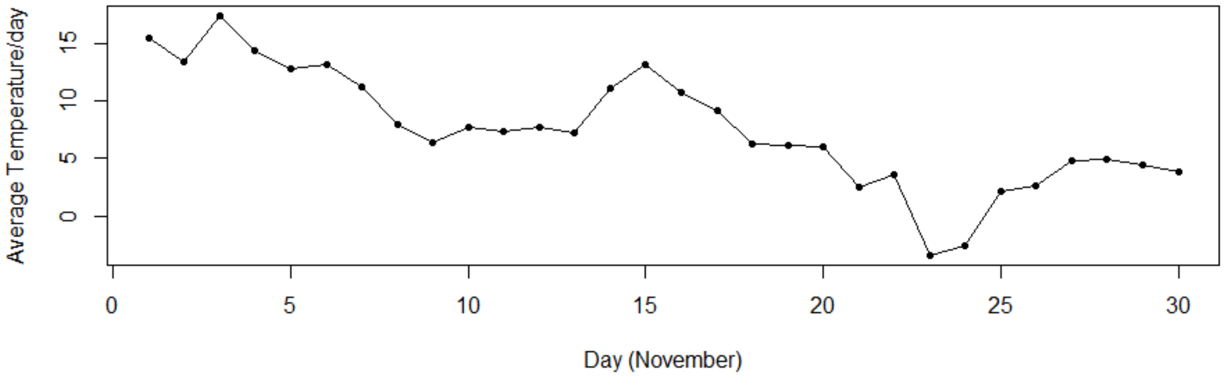


**Figure 4-6** Average temperature per day, during the month of November




***Stationarity***

We next examine the ridership as a time-series object. Figure 4-7 shows a plot of the average cyclist counts per day and Figure 4-8 shows a plot of autocorrelation function (ACF). The ACF plot exhibits a significant auto-correlation at lag 1 and a seasonal auto-correlation with neighboring effects at lag 7, i.e., a group of autocorrelations at the lag 6, 7, and 8 appear repeatedly every 7 periods. An augmented Dickey-Fuller Test of lag order 10 shows the presence of a unit root, i.e. the series is non-stationary. Figure 4-9 show plots of the series and ACF of the average ridership after differencing. After differencing, the Augmented Dickey-Fuller Test rejected the presence of a unit-root at the 1% level; meaning that the series is now stationary. There still appears a strong seasonal autocorrelation at period 7 as shown in the ACF plot in Figure 4-9.

**Figure 4-7** Time series of daily ridership

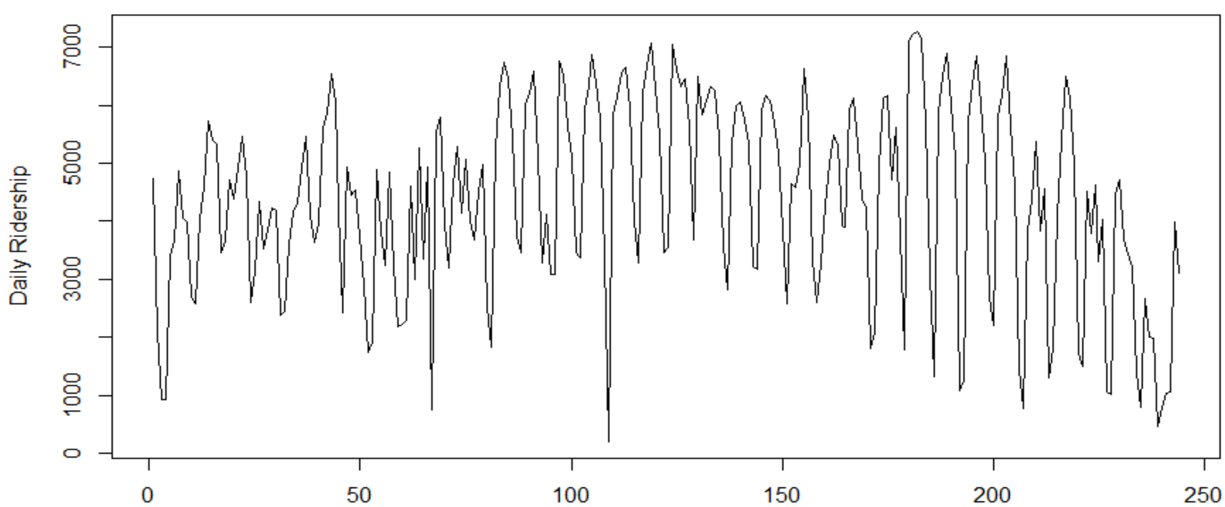


**Figure 4-8** Autocorrelation function (ACF)

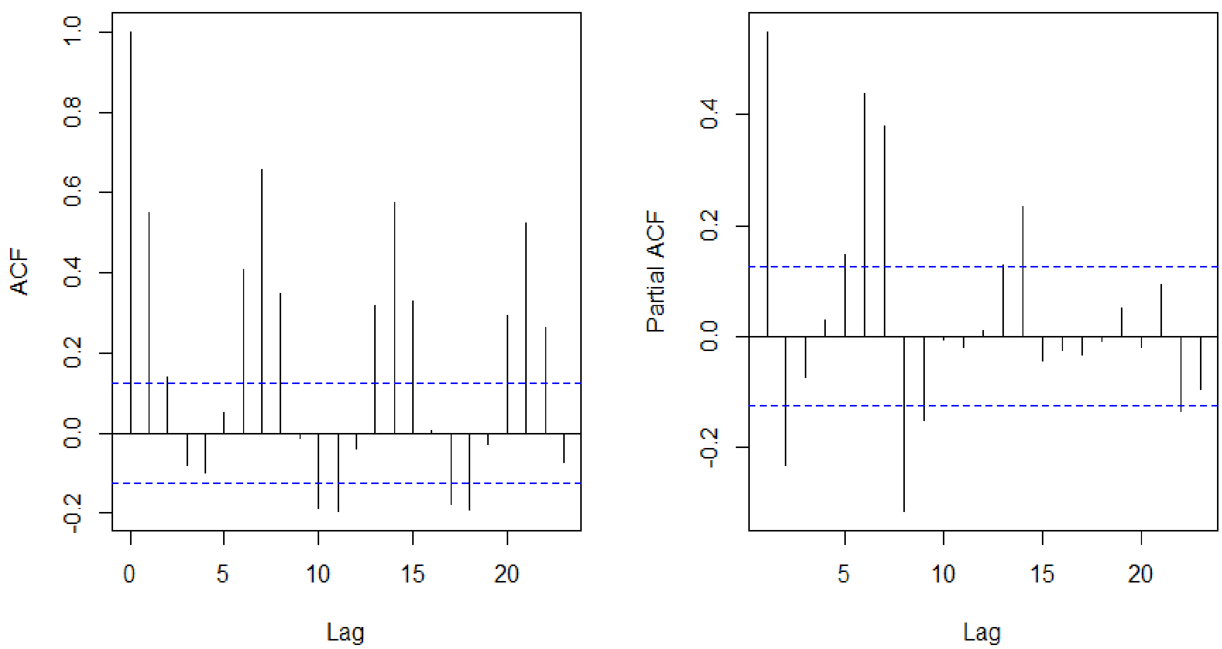
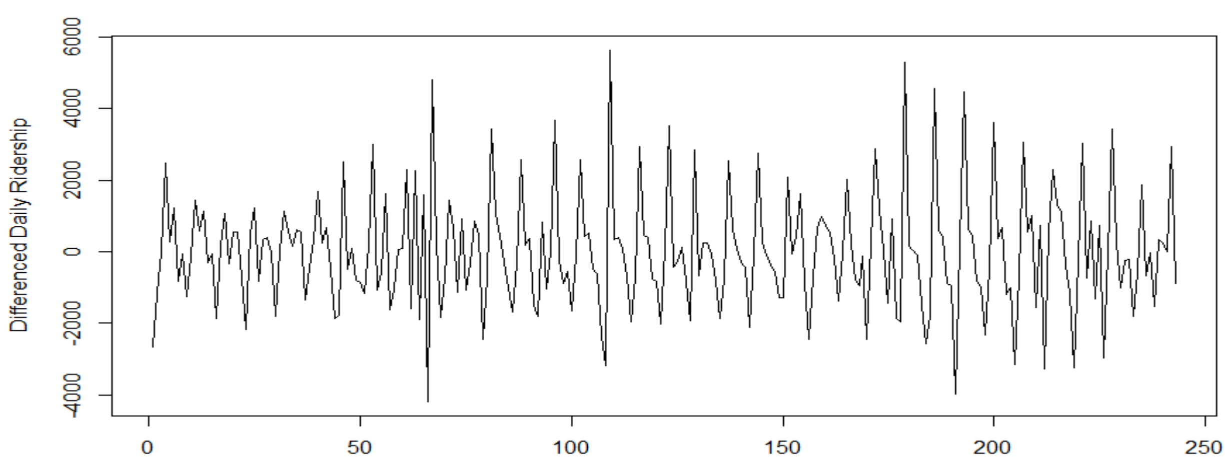
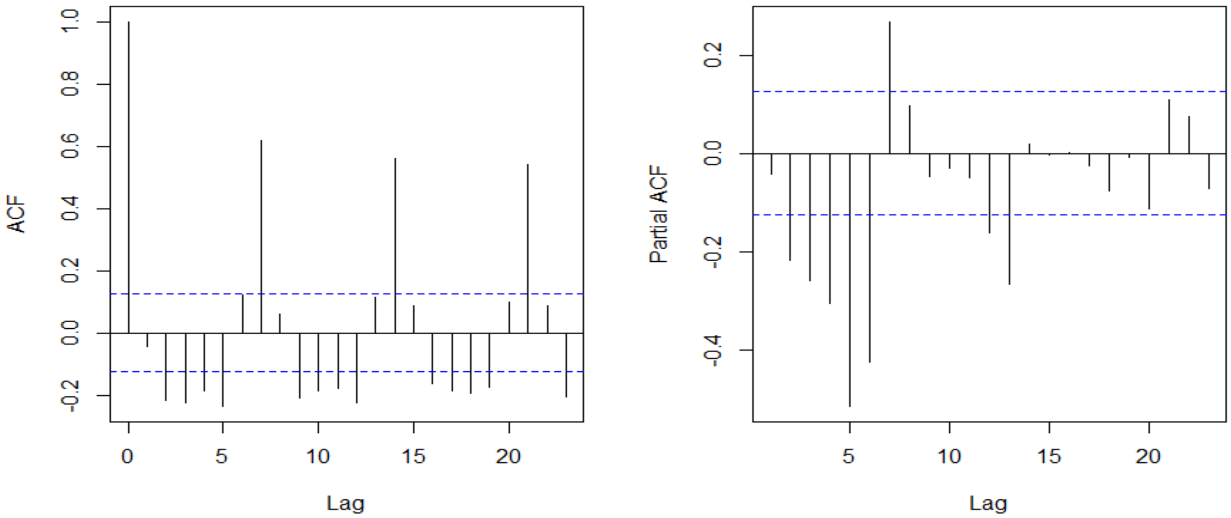



**Figure 4-9** Average ridership and ACF after differencing






### 4.4 Results

***SARIMA***

We first estimated SARIMA model and produced predicted values for the ridership in October. After study of ACF and PACF of the series and with the AIC criterion, we identify the appropriate model as ARIMA (5,1,2)(0,0,1)$_7$. The model was successful in removing auto-correlation as the Ljung-Box test gives a chi-square value of 10.3 with a p-value of 0.42. Figure 4-10 shows the observed daily ridership between April and September and the fitted values from the model. Figure 5-11 shows the observed ridership in October and the forecast from the SARIMA model. As can be seen in Figure 4-11, the forecast from SARIMA tend to overpredict the ridership, especially for weekends. The root mean square error is 1048.56, and the root mean forecast error is 2543.5.

**Figure 4-10** Observed and fitted daily ridership between April and September, SARIMA

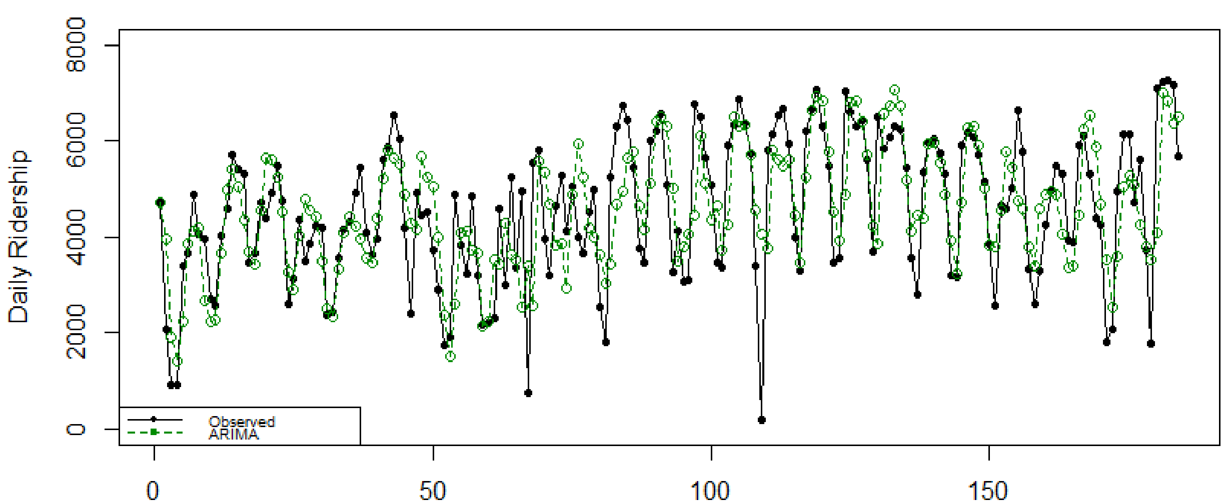
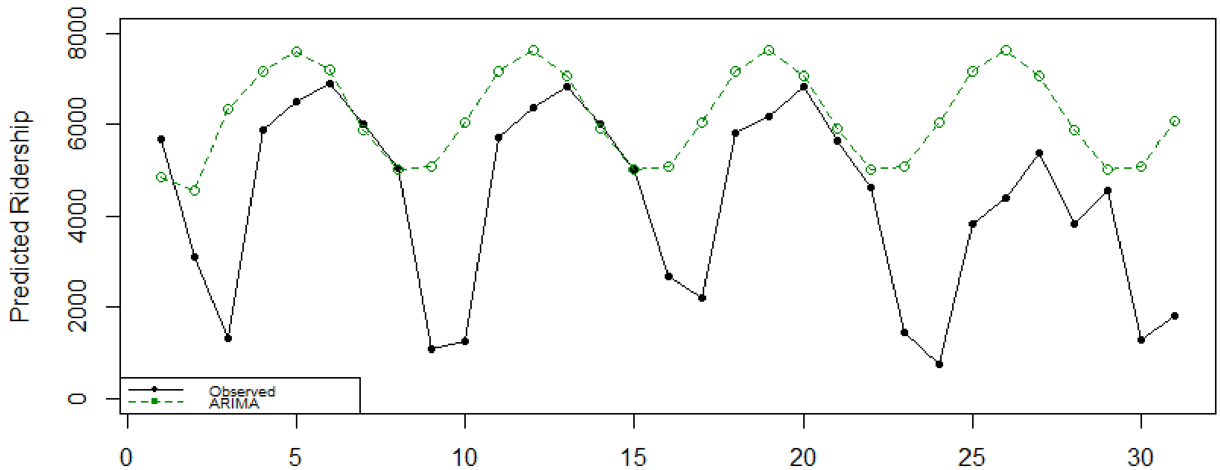


**Figure 4-11** Predicted ridership for the month of October, SARIMA




***Negative Binomial***
We next estimated a NB model with temperature, squared temperature, precipitation, weekend dummy, and dew point depression as explanatory variables. Figure 4-12 shows the fitted values, and Figure 4-13 shows the forecast from the NB model. The weather variables seem to explain the ridership very well. The prediction from NB model is much better than that from the ARIMA model, though the forecast tends to underpredict the ridership for weekdays and overpredict for weekends. The root mean square error is 805.21, and the root mean forecast error is 993.44.

**Figure 4-12** Observed and fitted daily ridership between April and September, NB

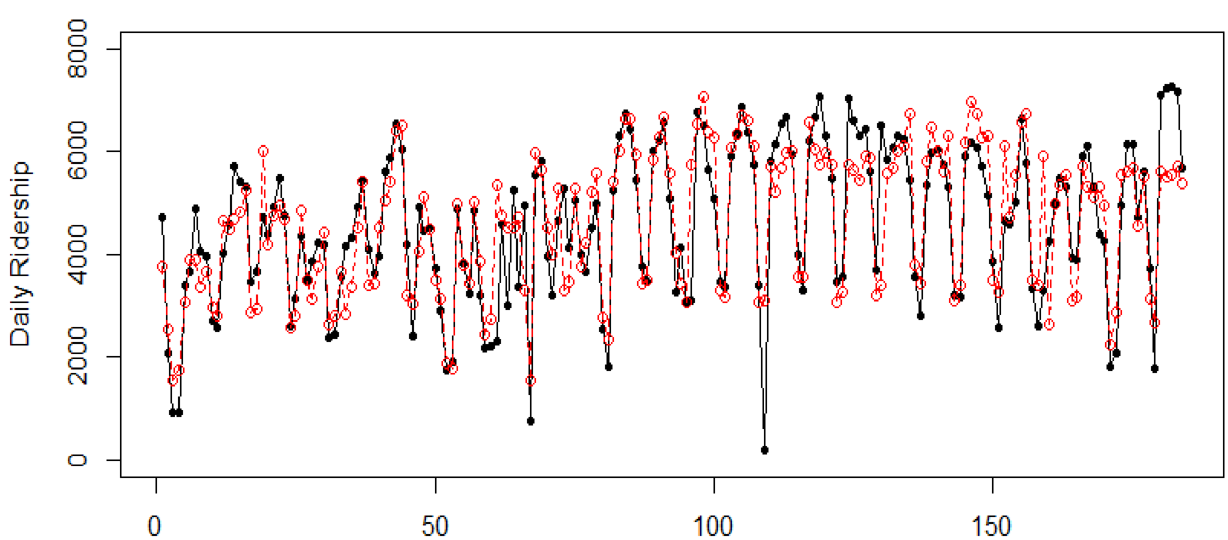
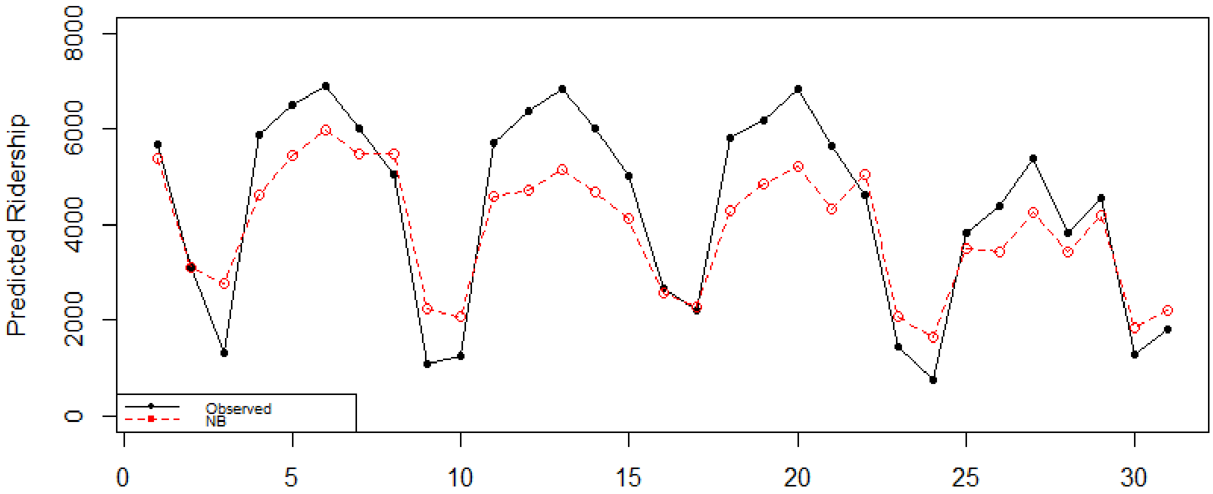


**Figure 4-13** Predicted ridership for the month of October, NB




***SARIMAX***

After the study of ACF and PACF, we specified the model as a SARIMAX $(0,1,1)(0,0,1)_7$. The model was successful in removing auto-correlation as the Ljung-Box test gives a chi-square value of 7.22 with a p-value of 0.70. Figure 4-14 shows the fitted values, and Figure 4-15 shows predicted ridership in October. As is obvious from Figure 4-15 and Table 4-1, SARIMAX outperforms both the SARIMA model and the NB model in terms of both model fit and forecasting accuracy. The root mean square error is 760.2, and the root mean forecast error is 842.29.

**Figure 4-14** Observed and fitted daily ridership between April and September, SARIMAX

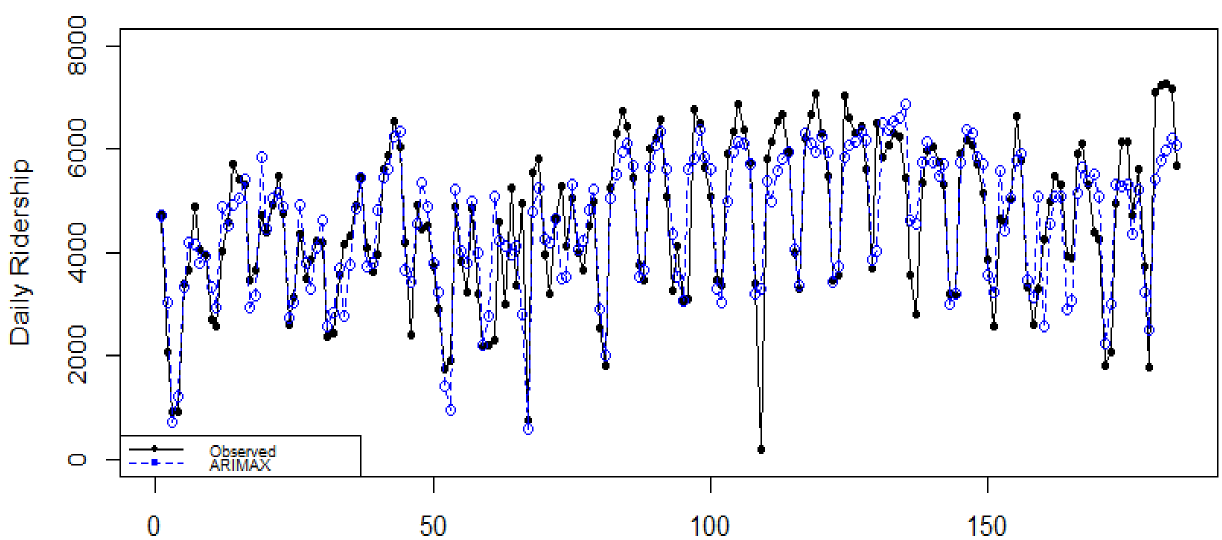


**Figure 4-15** Predicted ridership for the month of October, SARIMAX

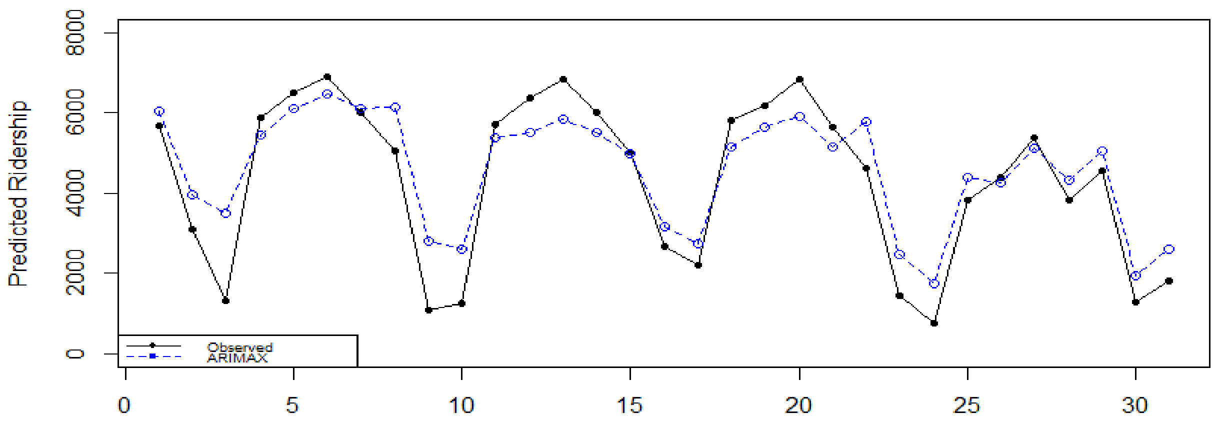



**Figure 4-16** Summary of predictions (green: SARIMA; blue: SARIMAX; red: NB; black: observed values)

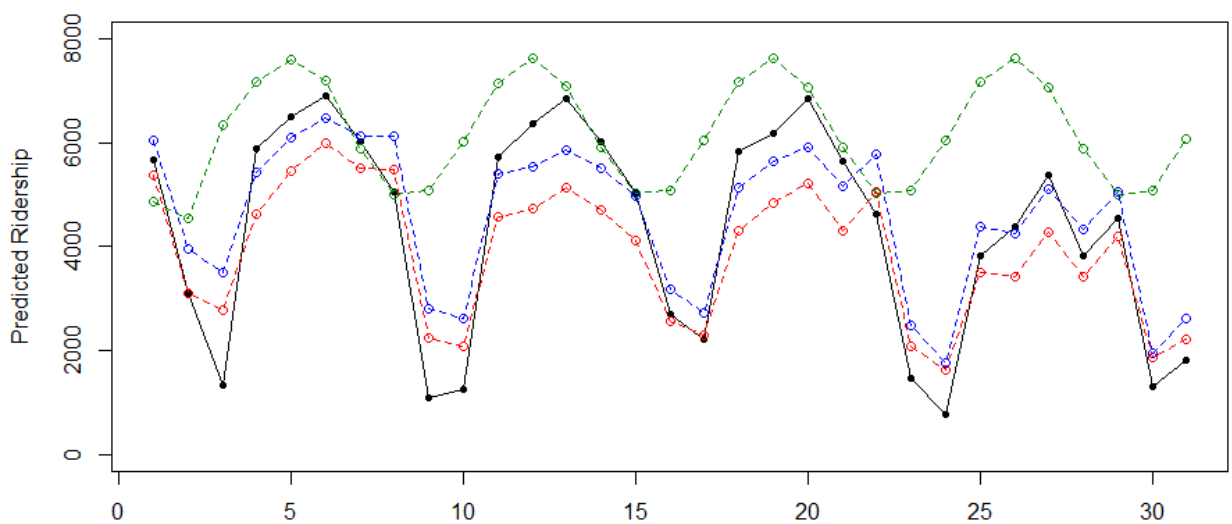


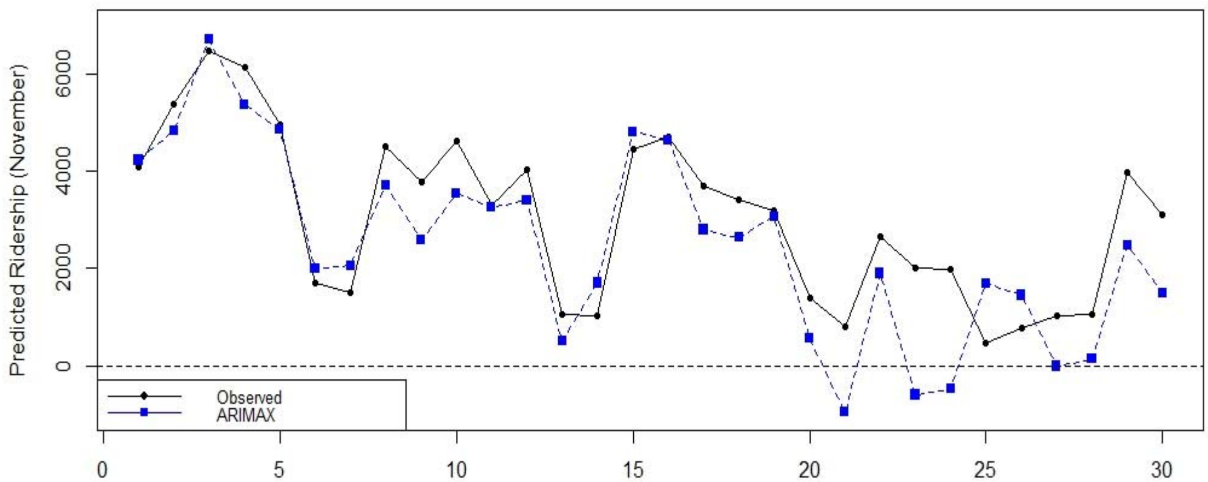
**Table 4-1** Time series estimates

| | ARIMA(5,1,2)(0,0,1)$_7$ | | ARIMAX(0,1,1)(0,0,1)$_7$ | | Negative Binomial | |
|---|---|---|---|---|---|---|
| | Coef | t - stat | Coef | t - stat | Coef | t - stat |
| MA1 | -1.219 | -44.663 | -0.869 | -14.480 | - | - |
| MA2 | 0.978 | 24.037 | - | - | - | - |
| Seasonal MA1 | 0.206 | 2.772 | 0.208 | 2.893 | - | - |
| AR1 | 0.648 | 8.503 | - | - | - | - |
| AR2 | -0.739 | -8.530 | - | - | - | - |
| AR3 | -0.231 | -2.285 | - | - | - | - |
| AR4 | -0.191 | -2.176 | - | - | - | - |
| AR5 | -0.233 | -3.140 | - | - | - | - |
| Precipitation | - | - | -4.19 | -3.051 | -0.002 | -3.54 |
| Temperature | - | - | 477.31 | 6.101 | 0.12 | 6.37 |
| Temperature square | - | - | -10.81 | -5.482 | -0.003 | -5.15 |
| Weekend | - | - | -2285.77 | -16.09 | -0.60 | -15.50 |
| dpd | - | - | 121.57 | 3.826 | 0.032 | 3.83 |
| Log-likelihood | -1536.41 | - | -1474.90 | - | -1534.55 | - |
| AIC | 3090.82 | - | 2965.80 | - | 3083.10 | - |
| Root mean square error | 1048.56 | - | 760.20 | - | 805.21 | - |
| Root mean square forecast error | 2543.5 | - | 842.29 | - | 993.44 | - |

The performance of the fitted models was investigated using the relative forecast error. The result of this study suggests that the SARIMAX model has the best prediction power. This is due to the fact that SARIMAX is able to take into account both serial correlation and exogenous factors influencing the ridership. However, our finding has some limitations – when modeling nonnegative integer-valued count data such as traffic count, Box and Jenkins models may be inappropriate because of the normality assumption on which the ARIMA model is based.

To illustrate this problem, we forecasted the ridership in November using the same models. As discussed above, the weather in the latter part of November is characterized by much colder temperatures and higher precipitation than any other month included in the data. As a result, some of the predicted ridership for November from the SARIMAX model becomes negative as shown in Figure 4-17. This suggests that although SARIMAX model has the highest forecasting power as well as the best fit, it is less appropriate as ridership count becomes smaller. In those cases, the NB model performs reasonably well without the problem of negative forecast values.

**Figure 4-17** Prediction of the SARIMAX model for the month of November




Our conclusion above suggests a doubt on the conventional statistical analysis on time-series count data. Poisson regression and Negative Binomial regression guarantees the integer forecast values, but ignores autocorrelation in the series. Although ARIMA-type models have been popular method to analyze time-series data, we found that using ARIMA-type models has at least three major problems when applied on count data:

1. Stationality requirement (pre-transformation of data)
2. Difficulty in making inferences on explanatory variables when data is transformed
3. Negativity of forecast values resulting from the normality assumption

For example, in this study we needed to take a first difference in order to get stationarity in the data. As a result, the coefficient estimates of weather variables became difficult to interpret. The negative sign of precipitation in our estimated model means that a positive difference of precipitation between days has a negative impact on the ridership; however, the positive difference can mean either increase in precipitation in the current period or a decrease in the last period. State Space models (or more specifically, dynamic regressions) solve all of the problems listed above as they do not require stationarity in data.

***State Space Model***
The estimation of State Space model

$$
\begin{aligned}
Y_t &= \alpha_t + v_t, \qquad & v_t \sim N_m(0, {\sigma_v}^2) \\
\mu_t &= \alpha_{t-1} + w_t, \qquad & w_t \sim N_p(0, {\sigma_w}^2)
\end{aligned}
$$

involved estimation of the two unknown variances, dV and dW. Figure 4-18 displays the MCMC output obtained using the R statistical software. The ergodic means seem to be stable in the last part of the plot. Figure 4-19 shows the observed daily ridership between April and September and the fitted values from the State Space model, and Figure 4-20 shows the predicted ridership in October. Since the model follows a random walk with noise, the forecast values are simply the filtered estimate from the last day in September.

**Figure 4-18** MCMC plots

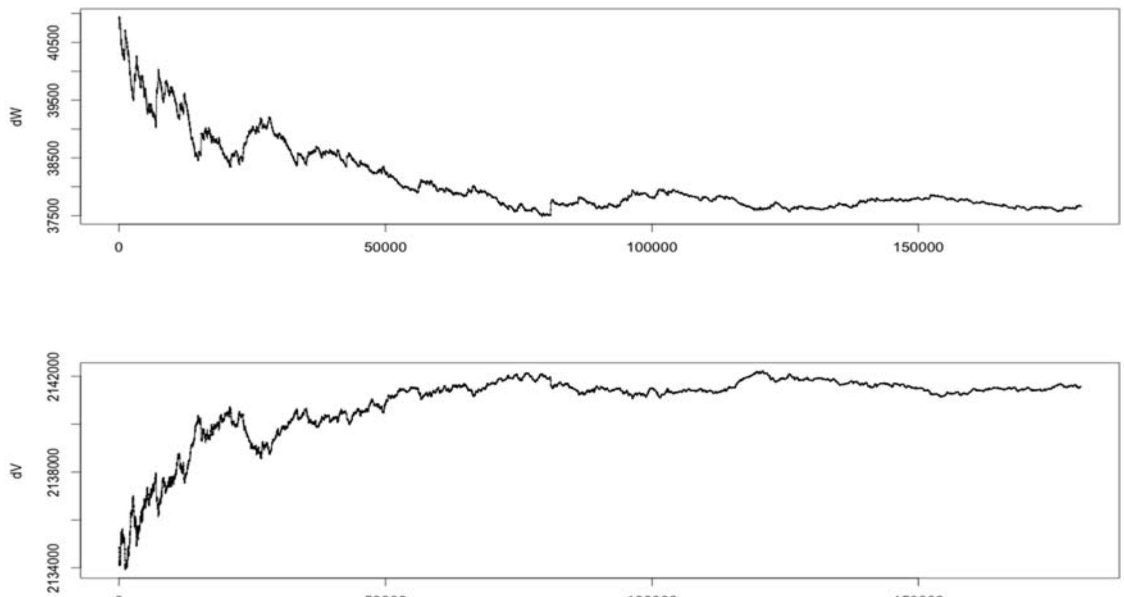



**Figure 4-19** Observed and fitted daily ridership between April and September, State Space Model

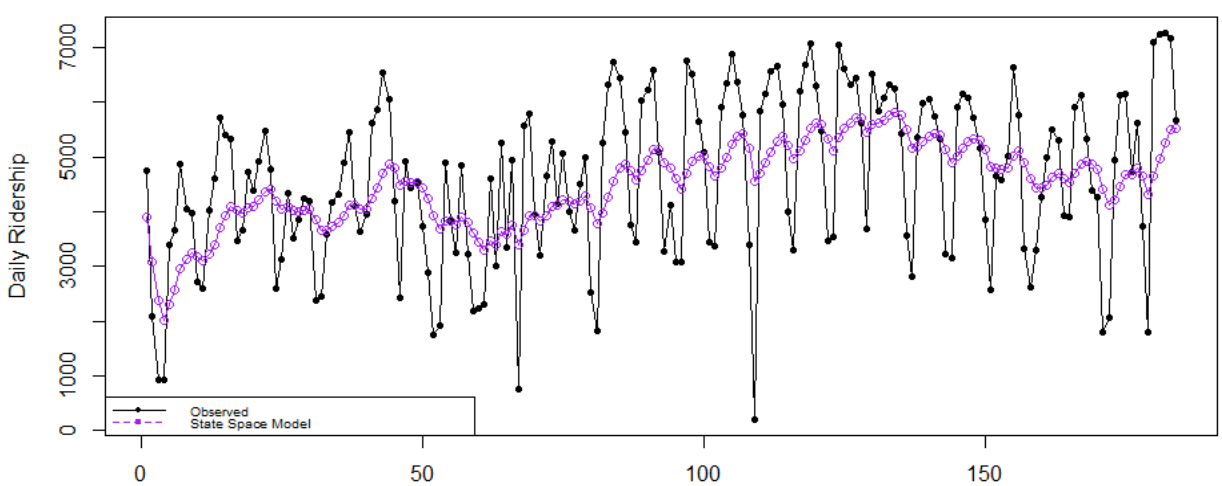
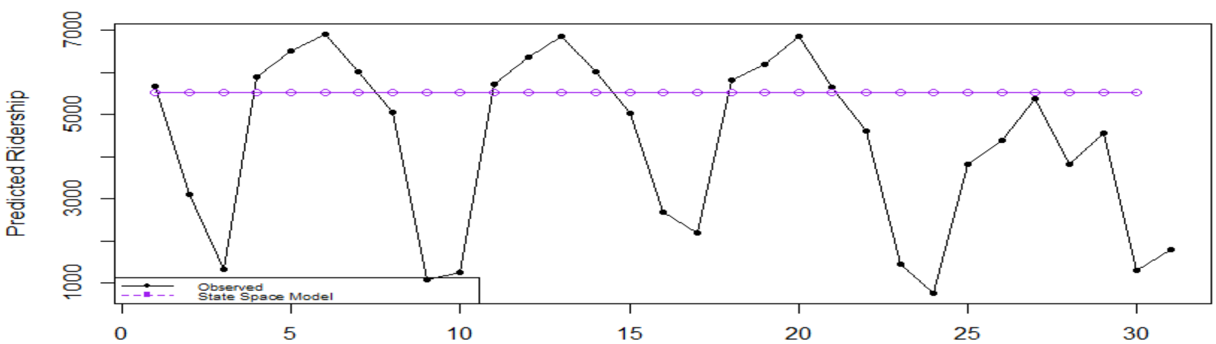


**Figure 4-20** Predicted ridership for the month of October, State Space Model




State space models may include explanatory variables as in the following specification.

$$Y_t = X\theta_t + v_t \qquad v_t \sim \tau(0, \sigma_v{}^2)$$
$$\theta_t = Z\theta_{t-1} + w_t \qquad w_t \sim \tau(0, \sigma_v{}^2)$$
$$where\ Y_t = \left(Y_{1,t}, \cdots, Y_{n,t}\right)',\ \theta_t = (\beta_1, \cdots, \beta_k)' \sim \varphi(l, m),$$
$$X = \begin{bmatrix} 1 & f_2(x_1) \cdots & f_k(x_1) \\ & \vdots & \\ 1 & f_2(x_n) \cdots & f_k(x_n) \end{bmatrix}$$
$$Z = diag(\delta_1, \cdots, \delta_k)$$
$$\sigma_v{}^2 = diag(\emptyset_1{}^{-1}, \cdots, \emptyset_n{}^{-1})$$
$$\sigma_w{}^2 = diag(\omega_1{}^{-1}, \cdots, \omega_n{}^{-1})$$

We expect to extend the model in order to accommodate count data. With alternative distributional assumptions for the random terms $v_t$, $w_t$, and $\theta_t$, we will be able to estimate a dynamic Poisson regression and dynamic Negative Binomial regressions, which (1) does not require stationarity, (2) produces integer-forecast values, and (3) produces coefficient estimates that are much more intuitive to interpret. Implementation of a Negative Binomial regression in a state space model will require derivation of the Gibbs sampling steps (i.e. finding the full conditional distributions of the parameter space).

# 5. DISCUSSION

In this project we have analyzed different econometric methods for making statistical inference on the demand for nonmotorized transportation, with a focus on cycling. In particular, the methodological framework applied in this project looks at investigating the factors that influence people's decision to use bicycles as a means of transportation.

For this project we designed a web-based survey aimed at collecting attitudinal data about cycling in general. The design was based on responses we obtained in two focus groups. In the second focus group we also tested a preliminary version of the survey. The survey also incorporated a discrete choice experiment for hypothetical scenarios of cycling route decisions. The experimental design has some new elements when compared with previous empirical applications for cycling route choice. Besides the common attributes (time, slope-topography, cycling infrastructure), we included variables related to weather conditions (temperature, amount of expected rain, amount of expected snow). The weather information was presented mimicking weather conditions forecasts for the day as depicted in smartphone applications. In addition, for each choice situation we considered the possibility of opting out (i.e. choose not to bike). The survey was applied to members of the Cornell community. More than 600 individuals responded to the survey, for a total of 599 valid responses. Since each responded answered 9 experimental choice situations, there were a total of 5391 pseudo-individuals for estimation purposes.

For analyzing the stated route choices and for determining the impact of cycling determinants (including weather and topography), we used discrete choice theory. We tested several models, including logit-based models integrated with a structural equation model for three latent variables, summarizing bicycle status, bike anxiety, and physical condition. In particular we derived a latent class model with a class assignment mechanism based on the latent bicycle status of the respondent. Two segments were identified: more-skilled and experienced cyclists, versus less-skilled- and non-cyclists. The two segments have different sensitivities to the factors that may encourage or discourage riding a bike. For instance, slope inclination is considered almost 3 times as bad by less-skilled cyclists. Heavy traffic affects twice as much to less-skilled cyclists, who also consider rain to be 2.4 times more bothersome (and snow almost 4 times more bothersome) than more-skilled cyclists. Because in cycling route decisions there is no direct monetary cost involved, to analyze differences in the taste parameters we have proposed to use the ratio of the marginal rate of substitution with respect to travel time. In addition, we measured the diminishing negative effect of a hilly topography (slope inclination) as a function of the physical condition of the cyclist (i.e. the more fit the cyclist, the less bothersome a steeper route.)

In terms of cycling infrastructure, our results are in line with previous findings from past research on bicycling: having more bicycle facilities result in a higher share of cycling (Akar & Clifton, 2009; Barnes & Thompson, 2006; Pucher & Buehler, 2006; Klobucar & Fricker, 2007; Dill & Voros, 2007). Our analysis showed that the presence of bike lanes is appreciated not only by individuals with higher skills and experience in bicycling but also by individuals who have less skills and experience. In fact, the estimates of the latent class model shows that less-skilled cyclists appreciate the presence of bike lanes 1.6 times more than more-skilled cyclists. It seems that people with less skills and experiences in bicycling value bike lanes as safety measures; thus the availability of bike lanes increases their likelihood to choose a bike route. In terms of policy recommendations, the provision of bike lanes may encourage an increase in the modal share of cycling, especially among those individuals with a lower cycling status (i.e. using a bike infrequently, or mostly for recreational purposes).

As a second project, in this study we have used a time series bicycle ridership count dataset was used to examine the performance of several ridership prediction models, including the Negative Binomial regression and time-series models such as ARIMA and ARIMAX (e.g., Nihan & Holmesland, 1980 and Houston & Richardson, 2002).

The result of this study suggests that the SARIMAX model has the best prediction as this model is able to take into account both serial correlation and exogenous factors influencing the ridership. However, our finding is that Box and Jenkins models may be inappropriate for count data because of the normality assumption on which the general ARIMA model is based, especially when ridership counts become smaller. This problem is especially relevant for months with weather conditions that discourage the use of cycling – and a conclusion that is shared by both the discrete choice and time series analyses is that poor weather conditions are indeed a main determinant of cycling decisions. The special family of State Space models has the potential to addresses the non-normality issues. Although estimation of State Space models is computationally intensive, software implementations have increased with the progress in computing capabilities and the modernization of computer software. Many well-known statistical and econometric software packages currently have options for the use of state space methods (Durbin & Koopman, 2012). However, most of them are currently only capable of estimating the basic state space model – that predicts averages – identical to the model we used in this study. The estimation of more generalized state space models requires much more intensive computer programming, and therefore we leave this task to future research.

Finally, we would like to mention that in New York City a project just started, gathering data via Eco-Counter's ZELT inductive loops that count using electromagnetic signature analysis of bicycle wheels (see Figure 5-1). In the near future we expect to apply, test, and validate the methodology that we outline in this project – including generalizations of State Space models – to cycling count data for New York City.

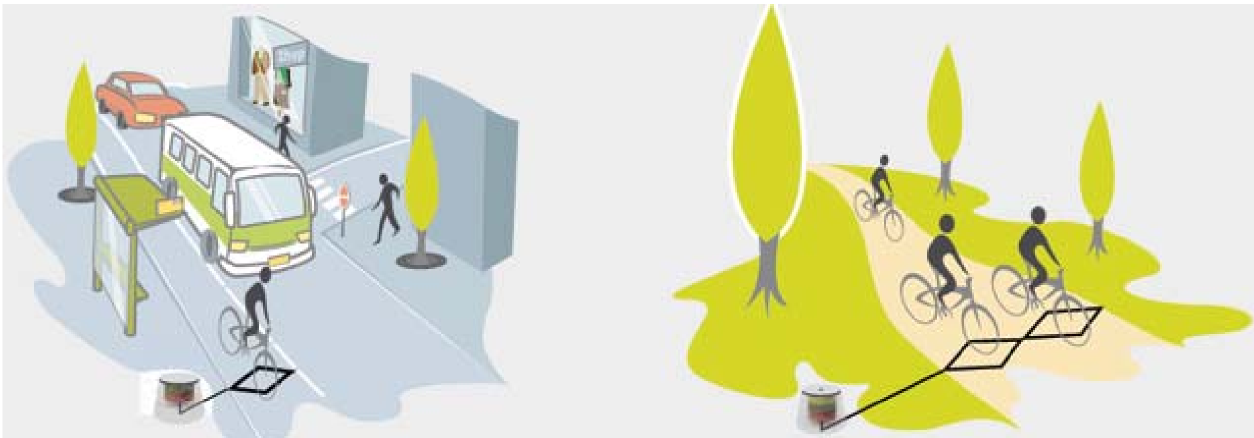


**Figure 6-1: Inductive loops for cycling counts in mixed traffic (left) and dedicated cycling paths (right)**
Source: Eco-Counter Inc.

# 6. REFERENCES

Agresti, A. (2013). Categorical data analysis. New York: Wiley.

Akar, G., & Clifton, K. J. (2009). Influence of Individual Perceptions and Bicycle Infrastructure on Decision to Bike. *Transportation Research Record,* 2140, 165.

Allaman, P. M., Tardiff, T. J., Dunbar, F. C., & National Cooperative Highway Research Program. (1982). *New approaches to understanding travel behavior*. Washington, D.C: Transportation Research Board, National Research Council.

Balsas, C. J. L. (2002). New Directions for Bicycle and Pedestrian Planning Education in the US. *Planning Practice and Research, 17,* 1, 91-105.

Balsas, C. (2004). Sustainable transportation planning on college campuses. *Sage Urban Studies Abstracts, 32,* 2.)

Baltes, M. R. (1996). Factors Influencing Nondiscretionary Work Trips by Bicycle Determined from 1990 U S Census Metropolitan Statistical Area Data.Transportation Research Record, 1538, 96.

Barnes, G. and Thompson, K. (2006) A Longitudinal Analysis of the Effect of Bicycle Facilities on Commute Mode Share Unpublished paper presented at presented at the 85th TRB Annual Meeting, Washington, DC.

Barrett, P. (2007). Structural equation modelling: Adjudging model fit. *Personality and Individual Differences, 42,* 5, 815-824.

Ben-Akiva, Mosh, et al. (2002) Integration of choice and latent variable models. *In: Perpetual motion: Travel behaviour research opportunities and application challenges*, 431–470.

Box, G. E. P., & Jenkins, G. M. (1976). Time series analysis: Forecasting and control. San Francisco: Holden-Day.

Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2008). Time series analysis: Forecasting and control. Hoboken, N.J: John Wiley.

Box, G. E. P., & Tiao, G. C. (1975). Intervention Analysis with Applications to Economic and Environmental Problems. Journal of the American Statistical Association, 70, 349

Brandenberg, C., A. Matzarakis, and A. Arnberger. (2007). Weather and cycling – a first approach to the effects of weather on cycling. Meteorological Applications, Vol. 14, 61-67.

Buehler, R. (2010). Transport Policies, Automobile Use, and Sustainable Transport: A Comparison of Germany and the United States. Journal of Planning Education and Research, 30, 1, 76-93.

Buehler, R. (2012). Determinants of bicycle commuting in the Washington, DC region: The role of bicycle parking, cyclist showers, and free car parking at work. *Transportation Research Part D: Transport and Environment, 17,* 7, 525-531.

Cornell University (2008). Transportation, Cornell, and the Community. http://www.tgeisproject.org/TGEIS_Documents/tGEISInsert-8-27-08final.pdf (accessed March 6, 2013).

Cornell University (2008) Transportation Impact Mitigation Strategies: Draft Report. http://www.tgeisproject.org/TGEIS_Documents/Final-Draft-TIMS-060308.pdf (accessed March 6, 2013).

Cornell University (2012). *Cornell Facts 2011-2012*. http://www.cornell.edu/about/facts/cornell-facts-2011-12.pdf (accessed March 6, 2013).

Cornell University Department of Transportation and Mail Services. Campus Bike Map. http://transportation.fs.cornell.edu/default.cfm (accessed March 6, 2013).

Campus Planning Office (2008). *Cornell Master Plan*. http://www.masterplan.cornell.edu (accessed March 13, 2013).

Dang, M. (2007). University Changes Free Bus Pass Policy. *The Cornell Daily Sun*. April 25, 2007. http://cornellsun.com/node/23149 (accessed March 6, 2013).

Dill, J. (2009). Bicycling for transportation and health: the role of infrastructure. *Journal of Public Health Policy, 30,* 95-110.

Dill, J., & Carr, T. (2003). Bicycle commuting and facilities in major U.S. cities: If you build them, commuters will use them. *Transportation Research Record,* 1828, 116-123.

Dill, J., Voros, K., (2007). Factors affecting bicycling demand: Initial survey findings from the Portland, Oregon, Region. Transp. Res. Rec. 2031, 9–17.

Durbin, J., & Koopman, S. J. (2012). *Time series analysis by state space methods*. Oxford: Oxford University Press.

Gatersleben, B., & Appleton, K. M. (2007). Contemplating cycling to work: Attitudes and perceptions in different stages of change. *Transportation Research Part A, 41,* 4, 302-312.

Garrard, J., Rose, G., & Lo, S. K. (2008). Promoting transportation cycling for women: The role of bicycle infrastructure. *Preventive Medicine, 46,* 1, 55-59.

Golob, T. F. (2003). Structural equation modeling for travel behavior research.*Transportation Research Part B: Methodological, 37,* 1, 1-25.

Greene, W. H., & Hensher, D. A. (2003). A latent class model for discrete choice analysis: contrasts with mixed logit. *Transportation Research Part B: Methodological, 37,* 8, 681-698.

Habib, M. A., Shaw, N., Peterlin, M., (2013) "Examining the Anticipated Integration of

Bikeshare with Travel Modes: A Latent Class Model Application", Unpublished paper presented at presented at the 92nd TRB Annual Meeting, Washington, DC.

Heinen, E., van, W. B., & Maat, K. (2010). Commuting by bicycle: An overview of the literature. Transport Reviews, 30, 1, 59-96.

Heinen, E., Maat, K., & Wee, B. (2011). The role of attitudes toward characteristics of bicycle commuting on the choice to cycle to work over various distances. Transportation Research Part D: Transport and Environment, 16, 2, 102-109.

Heinen, E., van, W. B., & Maat, K. (2010). Commuting by Bicycle: An Overview of the Literature. Transport Reviews, 30, 1, 59-96.

Houston, D. J., & Richardson, L. E. (2002). Traffic safety and the switch to a primary seat belt law: the California experience. *Accident Analysis and Prevention, 34,* 6, 743-751.

Hunt, J., & Abraham, J. (2007). Influences on bicycle use. *Transportation, 34,* 4, 453-470.

Jong, de P. (2012). The health impact of mandatory bicycle helmet laws. *Risk Analysis : an Official Publication of the Society for Risk Analysis, 32,* 5, 782-90.

Kamargianni, M., Polydoropoulou, A. (2013) "Development of a Hybrid Choice Model to Investigate the Effects of Teenagers' Attitudes Towards Walking and Cycling on Mode Choice Behavior", Unpublished paper presented at presented at the 92nd TRB Annual Meeting, Washington, DC.

Keijer, M. J. N., & Rietveld, P. (2000). How do people get to the railway station? The dutch experience. Transportation Planning and Technology, 23, 3, 215-235.

Klobucar, M. S., & Fricker, J. D. (2007). Network evaluation tool to improve real and perceived bicycle safety. *Transportation Research Record,* 2031, 25-33.

Lewin, A. (2011) Temporal and Weather Impacts on Bicycle Volumes, In Transportation Research Record: Journal of the Transportation Research Board, No. 2536, Transportation Research Board of the National Academies, Washington, D.C., 18p

Litman, T., Laube, F. (2002). Automobile dependency and economic development. Victoria Transport Policy Institute, Canada

Markland, D., & Special issue on Structural Equation Modeling. (2007). The golden rule is that there are no golden rules: A commentary on Paul Barrett's recommendations for reporting model fit in structural equation modelling. *Personality and Individual Differences, 42,* 5, 851-858.

McFadden, Daniel. (1986) "The Choice Theory Approach to Market Research." Marketing Science 5.4: 275-297.

McIntosh, C. N., & Special issue on Structural Equation Modeling. (2007). Rethinking fit assessment in structural equation modelling: A commentary and elaboration on Barrett (2007). *Personality and Individual Differences, 42,* 5, 859-867.

McMillan, T. E. (2007). The relative influence of urban form on a child's travel mode to school. *Transportation Research Part A, 41,* 1, 69-79.

Miranda-Moreno, L., & Nosal, T. (2011). Weather or not to cycle: Temporal trends and impact of weather on cycling in an urban environment. *Transportation Research Record,* 2247, 42-52.

Morikawa, T., Ben-Akiva, M., McFadden, D. (2002). Discrete choice models incorporating revealed preferences and psychometric data. Econometric Models in Marketing 16, 27–53.

Moudon, A. V., Lee, C., Cheadle, A. D., Collier, C. W., Johnson, D., Schmid, T. L. and Weather, R. D. (2005) Cycling and the built environment: a US perspective, Transportation Research Part D, 10, pp. 245–261.

Northeast Regional Climate Center. Ithaca Climate Normals. http://www.nrcc.cornell.edu/index.html (accessed March 20, 2013)

Nankervis, M. (2000). The effect of weather and climate on bicycle commuting. *Sage Urban Studies Abstracts, 28,* 2.)

Nelson, A. C., & Allen, D. (1997). If you build them, commuters will use them: Association between bicycle facilities and bicycle commuting. *Transportation Research Record,* 1578.)

NHTSA and Bureau of Transportation Statistics, U.S. Department of Transportation. (2003). National Survey of Pedestrian and Bicyclist Attitudes and Behaviors: Highlights Report.

Nihan, N. L., & Holmesland, K. O. (1980). Use of the box and Jenkins time series technique in traffic forecasting. *Transportation, 9,* 2.

Noland, R. (1995). Short-run and long-run policies for increasing bicycle transportation for daily commuter trips. *Transport Policy, 2,* 1, 67-79.

Olaru, D., Smith, B., & Taplin, J. H. E. (2011). Residential location and transit-oriented development in a new rail corridor. *Transportation Research Part A, 45,* 3, 219-237.

Petris, G., Petrone, S., & Campagnoli, P. (2009). *Dynamic linear models with R*. Dordrecht: Springer-Verlag.

Prochaska, J.O., DiClemente, C.C., (1984). The Transtheoretical Approach: Crossing Traditional Boundaries of Change. Dow Jones/Irwin, Homewood IL.

Pucher, J. (1998). Urban transport in Germany: providing feasible alternatives to the car. Transport Reviews, 18, 4, 285-310.

Pucher, J., & Buehler, R. (2006). Why Canadians cycle more than Americans: A comparative analysis of bicycling trends and policies. Transport Policy, 13, 3, 265-279.

Pucher, J., Buehler, R., & Seinen, M. (2011). Bicycling renaissance in North America? An update and re-appraisal of cycling trends and policies. *Transportation Research Part A: Policy and Practice, 45,* 6, 451-475.

Pucher, J., Dijkstra, L., (2000). Making walking and cycling safer: lessons from Europe. Transportation Quarterly 54 (3), 25–50.

Pucher, J., Komanoff, C., & Schimek, P. (1999). Bicycling renaissance in North America?. Transportation Research. Part A, Policy and Practice, 33, 625-654.

Richardson, A.J., (2000), Seasonal and Weather Impacts on Urban Cycling Trips . TUTI Report1-2000, The Urban Transport Institute, Victoria.

Rose, G., Ahmed, F., Filiozzi, M., Jakob, C. (2011), Quantifying and comparing the effects of weather on bicycle demand in Melbourne (Australia) and Portland (USA). Unpublished paper presented at presented at the 90th TRB Annual Meeting, Washington, DC.

Saneinejad, S., Roorda, M. J., & Kennedy, C. (2012). Modelling the impact of weather conditions on active transportation travel behaviour. *Transportation Research Part D, 17,* 2, 129-137.

Sener, I. N., Bhat, C. R., & Bhat, C. R. (2009). An analysis of bicycle route choice preferences in Texas, US. *Transportation, 36,* 5, 511-539.

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

Stinson, M. A. and Bhat, C. R. (2005). A Comparison of the Route Preferences of Experienced and Inexperienced Bicycle Commuters (Washington, DC: Transportation Research Board).

Tardiff, T. J. (1977). Causal inferences involving transportation attitudes and behavior.*Transportation Research, 11,* 6, 397-404.

Taylor, D., & Mahmassani, H. (1996). Analysis of stated preferences for intermodal bicycle-transit interfaces. *Transportation Research Record, 1996,* 1556.)

Thomas, T., R Jaarsma, B. Tutert. (2012). Exploring temporal fluctuations of daily cycling demand on Dutch cycle paths: the influence of weather on cycling. Transportation DOI 10.1007/s11116-012-9398-5.

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, 41,* 4, 287-301.

Train, K. (2003). *Discrete choice methods with simulation*. New York: Cambridge University Press.

United States. (2000). *Bicycle and pedestrian data: Sources, needs, & gaps*. Washington, DC: U.S. Dept. of Transportation, Bureau of Transportation Statistics.

U.S. Census Bureau. (2010). State & county Quickfacts, N.Y. http://quickfacts.census.gov. (accessed March 25, 2013).

Wang, C., Akar, G., Guldmann, J. (2013) "Do Your Neighbors Affect Your Mode Choice? A Spatial Probit Model for Commuting to The Ohio State University ", Unpublished paper presented at presented at the 92nd TRB Annual Meeting, Washington, DC.

Wardman, M., Tight, M., & Page, M. (2007). Factors influencing the propensity to cycle to work. *Transportation Research Part A, 41,* 4, 339-350.

Wee, B. van, Rietveld, P., & Meurs, H. (2006). Is average daily travel time expenditure constant? In search of explanations for an increase in average travel time. (Journal of Transport Geography, 14, 109-122.)

Weiss, B. D. (1996). Helmet use among university bicyclists. *Journal of American College Health : J of Ach, 44,* 6, 298-300.

The Weather Channel. Monthly Weather for Ithaca, NY. http://www.weather.com/ (accessed March 20, 2013)

# Appendix A

## SURVEY INSTRUMENT & FOCUS GROUP MEETINGS

This chapter is dedicated to the description of the survey development process, including the focus group meetings we conducted prior to the data collection effort.

### Focus Group Meeting (1)

A focus group session with students of Cornell University was organized in January 2013 to investigate attitudes, knowledge, and behaviors of people in the Cornell community related to bicycle. The purpose of the forum was for participants to discuss the potential factors that contributed to their travel mode decision so that attitudes and beliefs that would not be revealed otherwise could be identified.

The forum opened up with direct questioning and then went on to an open format to make sure any concerns and opinions the participants held were voiced. Special emphasis was placed to elucidate attitudes toward cycling, benefits and problems of using a bike for commuting to campus, and attitudes toward the potential use of cycling as a mode to use for trips within campus. A focus group guide was developed for use by the focus group moderator from the Cornell Survey Research Institute (SRI), and included questions and prompts designed to address four topics:

1. Decision making process to use bicycles or not use bicycle as a commuting transportation mode
2. Students' current attitudes toward biking in general from the points of view of both bikers and non-bikers
3. The extent to which environmental factors affect students' decision to bike or not bike
4. The role of social effects on biking patterns (friends' habits)

Focus group participants were recruited by distributing emails with invitations to students at Cornell. The focus group session of two and a half hours was held at the School of Civil and Environmental Engineering at Cornell University. A professional focus group facilitator and an assistant managed the session, while the researchers observed the discussion in the same room. The recorded focus group discussions were transcribed and then the researcher followed a deductive process of assessing the transcripts for patterns of responses and developing a quantitative data to link responses with the study research questions.

***Participants***
The group discussion began with open ended introductions, where each participant was asked, in addition to their name, affiliation, and area of employment or degree, to mention how he or she travels between home and campus. The individuals who participated in the focus group consisted of four undergraduate students and 10 graduate students; some of them were bicycle and others were not. Table A-1 presents a summary of their personal characteristics such as socio-demographic and travel behavior. Eight of them stated that they bicycle regularly and the remaining six of the participants stated they do not bicycle at all. Out of eight bikers, six participants stated that they commute to the campus by bicycle.

**Table A-1:** Focus group participants' socio-demographic and travel behavior characteristics

| Personal Characteristics | Number of participants | Personal Characteristics | Number of participants |
|---|---|---|---|
| Student Status | | Biker Status | |
| Graduate | 10 | Bikers | 8 |
| Undergraduate | 4 | Non bikers | 6 |
| Gender | | Commute mode | |
| Male | 5 | Bicycle | 6 |
| Female | 9 | Others (walk, bus, car) | 8 |
| Car ownership | | Biking during winter | |
| Car owners | 4 | Yes | 1 |
| Non-car owners | 10 | No | 13 |

***Results***

After a brief introduction, the focus group session preceded to a discussion of the travel mode choice of the participants, in particular why they bike to school and why they don't. As expected, the participants identified weather and hilly topography as the primary factors affecting their decision to commute to campus by bicycle. The cyclists enumerated several reasons for bike commuting to campus. For example, when the facilitator asked the participants why they bike or not bike, some of the cyclist participants stated that convenience was the primary reason for them to commute by bicycle. One of their responses was, "I like having the flexibility, because the bus that goes closer to my house is only once an hour, so if I ride my bike I can be on my own time more". Similarly, another biker participant said, "it (bicycling) saves a lot of time, and it's more fun than walking". Meanwhile, non-biker participants seemed to consider bicycling as an inconvenient mode of travel. For example, a non-biker said, "…bus transportation (near campus) is very good. Whenever I get out of the house there's a bus waiting there, so I don't think it's necessary for me to bike." Another non-biker stated, "If I bike, I feel like my paths would become extremely inefficient… there are a lot of trails that I can walk, but there aren't necessarily bike routes." From this discussion, we found bicycling can be considered either convenient or inconvenient form of transportation depending on the location of the residence, availability of other transportation modes in the area of residence, and existence of bike routes.

The facilitator pursued the role of environmental concerns by specifically asking participants whether the hilly topography and extreme weather play any role in the decision. All participants, both bikers and non-bikers, responded ''yes.'' The extent the hilly topography affects a student's decision to bike or not bike vary across the participants depending on the steepness of the hill they need to overcome and the availability of bike rack on buses, as can be seen in the following exchange:

> **Participant #8:**
> "I ride a bike, but usually not to campus, because I live at the bottom of a hill. Occasionally the bus has been full—the bike rack—and I've been out of luck, so that sort of turned me off of it."
> **Facilitator:**
> "Ah. Okay, so you would have biked to campus…."

**Participant #8**
: "Yeah, as in putting my bike on the bus."
**Facilitator:**
"So you bike to a bus stop and then…?"
**Participant #8**:
"Yeah, you're really out of luck if the bike rack is full."
**Facilitator**:
"So what prevents you from biking the entire distance?"
**Participant #8**:
"It's too steep."

Since the focus group session was held in January, most cyclist participants did not commute to campus as frequently as they wished to. The following exchange summarizes how most of the participants felt about biking to campus in the winter.

**Participant #4:**
"I bicycle to campus because if I drive it takes me approximately six minutes to get to the B (Parking) Lot and then waiting for the bus, it could easily take me another 15 minutes to be walking into Ives. If I bicycle, it takes me 17 minutes. So, it's basically the same amount of time."
**Facilitator:**
"So, the time constraint isn't different and you bike…how often do you bike to campus?"
**Participant #4:**
"Today is the first day I've done it this semester, but fall semester I did it about every time I had to come to campus, so probably about four times per week."
**Facilitator:**
"Okay, and why did you not bike so frequently so far?"
**Participant #4:**
"It's been cold."

Interestingly, some cyclists seemed to perceive effects of weather somewhat differently. While for non-bikers, cold weather completely keeps them from biking, some bikers stated they did not care much about the weather itself but they are more concerned about the road conditions that are invoked by bad weather. For example, one biker participant said, "If the ground is like a lot of puddles, then I don't really want to ride because I'll get really dirty, I guess. The cold weather is fine." We also found bikers care about the protection of their bicycles from harmful environment as can be seen in the following exchange:

**Facilitator:**
"Anything else discouraging you from biking?
**Participant #5:**
"Sometimes I just feel that I shouldn't bring my bike out during the winter just because the salt through the streets. It's a good bike, so I don't want to ruin it with the salt."
**Participant #4:**
"Actually that also reminded me…. Yesterday I was thinking about it, but they were calling for rain and there's no covered bike parking, at least none that I know of on

campus. I also just recently got a new bike and I was like, "Well, I don't have a tarp for it." So I was thinking about making or finding some sort of tarp so I could protect it somewhat, but it'd be nice if there was some place that we could put our bikes that was—at least sleet wouldn't get on them or rain wouldn't get on them."

Our speculation prior to the focus group session was that the bicycle habit of friends' may potentially have a positive effect on an individual's decision to bike: students would feel more comfortable to bike with their friends, especially for travels within campus. The participants agreed on the former but not the latter. Most participants, both cyclists and non-cyclists, said that seeing other people frequently bicycling would be an incentive for them to bicycle. As for the latter part of our speculation, the participants felt the opposite way. Cyclists consider bike commuting as a solitary process as opposed to walking, which is more social. Non-cyclists agreed on it that they felt a similar way but more strongly. Responses included these comments:

**Participant #2:**
"If I know I'll be walking with someone else, then I don't want to ride my bike alongside them walking. That would be awkward… (, and) it just takes up space on the sidewalk, so I wouldn't do that."
**Participant #1:**
"Bicycling limits your options, basically, in terms of socializing and enjoying the environment."
**Participant #3:**
"I don't own a bike, but I think it's kind of weird to bike with another person for transportation. I would bike with someone if we're, say, going to one of the parks and go on a trail and bike around, but I wouldn't bike around campus with someone else just because we could. That seems really weird to me. "
**Participant #10:**
"…when I'm biking on campus, this (bicycling) is like for efficiency, because you can just do it by yourself and you go, you grab your bike, and you just go right away. When you start doing this (bicycling with friends), it complicates the logistics a little bit. So, it's a little bit weird, if you have a friend and you're going to the same class, just agree to meet to go to class."

The facilitator pursued further the role of the social effects of bicycling by asking the participants, "So does anyone bike socially but not for transportation?" Several participants said they have biked with their friends for recreational purpose. For example, Participant #2 and Participant #6 said:

"There's a Cornell Cycling club and they have group rides, but those are like at a set time, they're not like impromptu rides like from one place to another. Or you can do like a group ride—just a ride with someone else, but off campus, and you go to the mall or something, but it wouldn't be like an on-campus-thing."

"Social biking is very interesting. I enjoy it very much, especially when I was not quite familiar with the areas outside of campus. It is a good way, you know, riding with a friend, and exploring different areas"

### *Discussion*

Findings from the focus group session raised several challenges for bicycle to be considered as a sustainable mode of transportation by students at Cornell University. All participants in the session perceived the hilly topography of Cornell campus and severe weather as the primary factors that discourage them to use bicycle as a transportation mode to travel to, from, and on campus. Non-cyclists particularly showed their distaste to consider bicycles as an alternative transportation mode because of those negative characteristics of the environment surrounding campus. While cyclists feel less aversion to those environmental factors, they had their own concerns, specifically regarding road conditions. Because of heavy snow and snow clearing maintenance, the road conditions in Ithaca tend to be rough. Snow shoveling creates bumps and depression on roads causing paddles when wet, and anti-icing agents on roads damage bicycles. We found that bikers feel that they are much more prone to suffer from these negative road conditions than cars.

One of the most interesting findings was how cyclists perceive (do not perceive) additional safety from the existence of bike lanes. When the focus group facilitator specifically asked about whether bicycle lanes would encourage them to bicycle more, the participants only moderately agreed on it. Cyclists in particular were used to ride on roads without bike lanes and they do not think that bike lanes would provide them any additional safety on the road. One cyclist said, "I mean, it (a road by the gym on campus) has a nice, wide, marked bike lane. I think it's just the drivers – they don't tend to signal, they just pull over quickly to the side of the road to drop people off…" This suggests that building bicycle lanes may not be a solution to providing safe road for the bikers, but rather drivers' awareness of bikers need to be promoted.

## Questionnaire development

We designed the survey instrument in multiple stages to reflect specific needs to the Cornell community. We first reviewed several survey instruments used in the past studies and selected questions from them to target desired topics, such as travel patterns, environmental factors, and perceptions associated with bicycling (Akar & Clifton, 2009; Stinson & Bhat, 2003). Our focus, however, is on analyzing bicycle route choice by evaluating (1) the trade-offs among the route facility attributes such as existence of bike lane and travel time, and (2) the effects on weather conditions on route choice decisions. We chose an online survey instead of a paper survey because of the high rate of Internet usage in the target population and because of logistical issues related to conducting other types of survey. We administered the survey through the Qualtrics software, Version 2013 of the Qualtrics Research Suite. We tested the survey instrument for feedback on phrasing and comprehension in a focus group session with students and staff of Cornell.

The survey instrument consisted of 23 survey items, which fell into six categories. The first category, "Travel characteristics", asked about basic travel characteristics, including bicycle use. The second category, "Obstacles", was concerned with factors that discourage respondents to commute to campus by bicycle. The third category, "Improvement", was concerned with factors that encourage to bicycle. Included in the second and third category were factors primarily related to the natural environmental factors and the lack of bicycle facilities. The fourth category, "Behavior and perception", included seven statements about subjective perceptions related to bicycling and self-estimate of physical ability using a five-point Likert Scale ("Strongly

Disagree", "Disagree", "Neither Disagree or Agree", "Agree", and "Strongly Agree"). In this category, the first question asked for the respondent's perception of motor traffic, and the following three questions asked for the respondent's perceptions of bicyclists from the points of view of drivers and pedestrians, respectively. We chose to include these questions with a speculation that a commuter's decision to bicycle is dependent on his/her perception of safety related to sharing the road with vehicular traffic and pedestrians, and how he/she perceives bicyclists' behavior. Since bicycling necessarily requires physical efforts, we also speculate that the decision to bicycle is also related to one's willingness and confidence to exercise; thus we included three questions asking about a respondent's willingness to exercise and his/her perception of own physical ability. The fifth category consisted of 6 binary route choice problems, which will be discussed in the next section. Finally, the sixth category consisted of questions regarding respondents' basic socio-demographic information.

**Table A-2:** Survey Items

| Characteristic category | Characteristic name | Characteristic type |
|---|---|---|
| Travel characteristics | Distance from campus | Categorical |
| | Frequency of commute to campus, days per week | Categorical |
| | Mode of transportation for travel on campus | Categorical |
| | Frequency of exercise | Categorical |
| | Self-assessment of bicyclist skill level | Categorical |
| | Availability of access to bicycle | Categorical |
| | Purpose of bicycling | Categorical |
| Obstacles | Deterrent factors | Likert |
| | Environmental factors | Likert |
| Improvements | Support factors | Likert |
| | Environmental factors | Likert |
| | Facility factors | Likert |
| Behavior and perception | Motor vehicle drivers seem to care little about bikers on road | Likert |
| | Bicyclists seem to care little about vehicular traffic on road | Likert |
| | Bicyclists seem to care little about pedestrians on street | Likert |
| | I do not like to share road with bikers when I am driving | Likert |
| | I have strong motivation to exercise | Likert |
| | I am confident about my physical fitness | Likert |
| | I enjoy outdoor activities (camping, fishing, jogging, etc.) | Likert |

<table>
<tr><td rowspan="3">Choice problems</td><td rowspan="3">If route A and route B are the only available options for your commute by bicycle, which route would you choose? If you would not bicycle on either route under the given conditions, choose “Neither”.</td><td>1 = Route A</td></tr>
<tr><td>2 = Neither</td></tr>
<tr><td>3 = Route B</td></tr>
<tr><td rowspan="4">Demographics</td><td>Cornell affiliation</td><td>Categorical</td></tr>
<tr><td>Gender</td><td>Categorical</td></tr>
<tr><td>Age</td><td>Categorical</td></tr>
<tr><td>City lived before moving to Ithaca</td><td>Categorical</td></tr>
</table>

**Table A-3: Attribute levels of discrete choice experiment**

| Attribute | Attribute levels | Attribute | Attribute levels |
|---|---|---|---|
| Weather | 1. Sunny<br>2. Rain<br>3. Snow | Bike lane | 1. Yes<br>2. No |
| Hilliness | 1. 0 %<br>2. 10 %<br>3. 20 % | Temperature | 1. Cold<br>2. Moderate<br>3. Warm |
| Time | 1. 10 minutes<br>2. 15 minutes<br>3. 20 minutes | | |

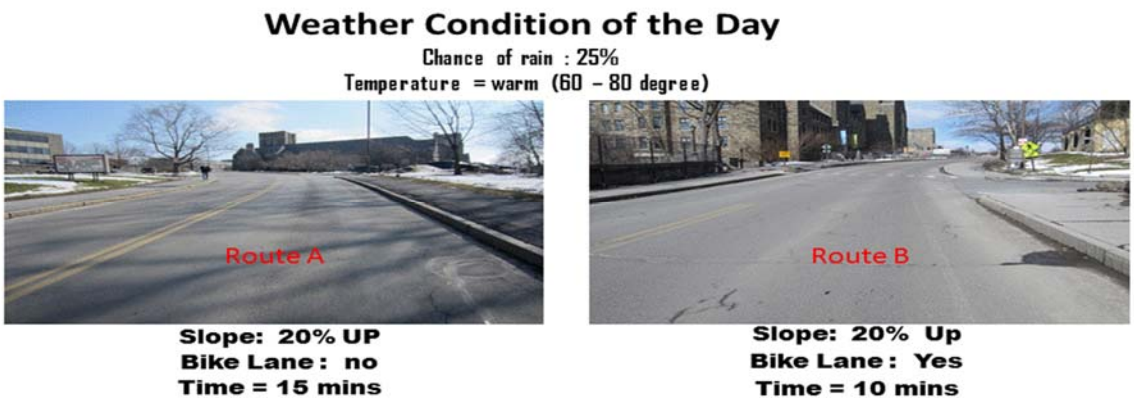
The discrete choice experiment is the key component of the survey. The experiments are based on binary route choice for bicycling, similar to that of Stinson & Bhat (2003), except that, in addition to worded descriptions of route characteristics, we also presented pictures of routes for each choice problem in the survey.

We believed that with pictures, the respondents would be able to visualize the hypothetical routes described in the problems much more easily. In order to examine effects of weather on a respondent’s choice of route, we specified the weather condition of the day, including temperature and chance of precipitation, for each route choice problem. Details of the attributes of the choice problems are summarized in Table A-3. Figure A-1 shows an example of the choice problem prepared for the draft of survey.

**Figure A-1.**

For each of the following questions, you will be given a hypothetical situation in which you decide whether you would be willing to ride a bicycle. You will be given **weather forecast** of the day and images of two routes from which you choose to ride a bicycle on. If you would not ride a bicycle on either route under the given conditions, choose "**Neither**".

Each route will be characterized by:
1. **Slope gradient** (Flat, Moderate slope, Steep slope)
2. **Bike lane or No bike lane**
3. **Trip length**




**Focus Group (2)**

During the development of the survey instrument, we organized a focus group forum to test the performance of the survey instrument. The purpose of the forum was to conduct a test run of the survey instrument and to get feedback from the participants on their comprehension of the survey questions and the choice experiments. The focus group forum of two hours was held at the School of Civil and Environmental Engineering at Cornell University in February of 2013.

***Methodology***

The methodology of organizing and conducting the focus group forum followed the same manner as the first focus group forum held in January. A group of the students and staff of Cornell University were invited to the session. We prepared an online version of the survey instrument described above for the participants to take as a test-run during the session. We were particularly interested in learning the effectiveness of visual aids in illustrating characteristics of routes. For example, we wanted to see if the participants could perceive hilliness of a landscape from two-dimensional pictures.

***Participants***

The group discussion began with introductions, where each participant was asked their name, affiliation to Cornell University, and whether he/she ever bikes. The individuals who participated in the focus group consisted of 5 undergraduate students, 7 graduate students, and 2 staff members. Table A-4 presents a summary of their affiliations with Cornell and their cycling status. Three of the participants stated that they bicycle regularly for recreation, another three used to bike for commuting but do not anymore, and the remaining six stated they do not bicycle at all. Only one of the participants stated that she commutes to the campus by bicycle. After a brief introduction, the participants were given a draft of survey instrument prepared for the session for evaluation.

**Table A-4:** Focus group participants' socio-demographic and travel behavior characteristics

| Personal Characteristics | Number of participants | Personal Characteristics | Number of participants |
|---|---|---|---|
| Student Status | | Biker Status | |
| Undergraduate | 5 | Bike regularly | 3 |
| Graduate | 7 | Bike for recreation | 3 |
| Staff | 2 | Used to bike | 3 |
| Gender | | Never bike | 5 |
| Male | 4 | | |
| Female | 10 | | |

***Results***

As expected, it took less than 10 minutes for all participants to finish the survey. When asked, almost all respondents stated that they felt the survey was shorter than their expectations. This finding suggested us that we could make the survey longer. The length of the survey will be further discussed later. After the survey, we proceeded to an open discussion on the design and questions on the survey.

Since we did not specify a particular purpose of hypothetical trips the respondents were asked to imagine, one of the initial responses from the participants was regarding the purpose of the hypothetical trip. Participant #10, who described herself as an avid biker, stated that she would choose a different route depending on the purpose of her bicycle trip, and thus had a hard time answering what she would do in the choice situations given in the survey. That is, she would choose a hilly route if she were biking for exercise and choose the shortest route if she was commuting. Because of this feedback, we decided to specify and describe the purpose of the hypothetical trip as commuting to Cornell campus in the actual survey.

Several participants stated that they had difficulty in imagining hypothetical weather conditions described in words. For example, Participant #9 said, "I had thought that the icons of the weather would be really helpful. I thought that the temperature sometimes is lost, because it's just a word and not an image. Maybe having a thermometer with kind of like a reading on it, then at least there's a visual representation of that concept as well so people don't just look at the pictures and answer the question without recognizing that temperature is also a consideration." As shown in Figure A-1, the weather condition was described in words for each choice situation and it can also be seen in the pictures of routes; however, the participants seemed to want an additional visual aid solely for describing weather. Furthermore, as a part of description of weather condition, we included a probability of precipitation; however, this also created some ambiguity in interpretation among the participants. For example, Participant #13 stated, "I bike while it's snowing like while it's starting, but then as it accumulates like on the side and they haven't yet salted it then I'm going to be like, 'No.' Just snow on the grass is not a huge a deal, but otherwise…" Because of these suggestions, we decided to represent weather with additional visual aids in the actual survey with more specific information about the amount of precipitation. The design of the visual aids for weather conditions will be discussed again later.

Illustrating the hilliness of routes in the discrete choice experiment was a challenging task in the development of the survey, and it was one of the main topics we wanted to be discussed in the focus group session. We anticipated that the visual aids (the real pictures of routes) will help the respondents visualize the route; however, since the pictures are still two-dimensional, we still wanted to test if the respondents would be able to perceive the inclination of the roads. For the focus group, we also described hilliness as a percentage grade, which is a commonly used unit for describing slopes in transportation (see, Figure A-1). Nevertheless, many of the participants stated that they had hard time in discerning the steepness of slopes from its description and corresponding pictures. The following exchanges among the participants summarize the feedback regarding the slope description.

**Participant #10:**
"Some of the—like this picture that says it's a steep slope—look more steep. It's kind of hard to tell the moderate versus steep."
**Participant #14:**
"I mean, if you just measured the degree of the slope and slap it on there, that'd be…"
**Participant #11:**
"I was going to say, would it help to put a triangle with the steepness of the slope?"
**Participant #8:**
"I think it'd be helpful to put a triangle. It's a little easier to see the slope, whereas here this kind of looks pretty flat to me."
**Participant #14:**
"Doing the triangle I think would be really, really good, because you've got some kind of visual reference."

The triangle they mentioned is a type of road sign that exist in reality, which serves as a warning sign for inclination of the road ahead. Since many of the participants suggested that "the triangle" would help them visualize hilliness, we decided to illustrate an image that resembles the actual "triangle" road sign and place it on the choice problems, which we will discuss later.

Several other suggestions were made regarding the pictures of routes. For example, some of the routes in the pictures had bike lanes drawn on the surface; however, the participants suggested adding visual enhancement on the bike lanes in the pictures.

**Participant #10**: "I actually didn't know that was a bike lane. I thought that was a shoulder. I honestly did not know that's bike lanes on campus. I just thought that was the shoulder." [Murmurs of agreement]
**Facilitator**: "I mean, taking this one as our example. How many people saw this (a picture of road with a bike lane) and knew that's the bike lane. So just by show of hands: nine."
**Participant #3**: "I know from personal experience, but if I just lived here I don't think I would know."

**Facilitator**: “So nine or ten. And I assume the complement…can I just do show of hands of how many people looked at that and said, “I’m not really clear that’s a bike lane.” So we’ve got four and a half. Okay.”
**Participant #1**: “Also, some of the paths, like the one that goes up past the law school is really, really faded.” [Murmurs of agreement]

As many of the participants pointed out, the bike lanes on roads were faded as it was the way the actual bike lanes were maintained near the university in reality, i.e., they were not maintained well. As will be shown later, because of the comments above, we superimposed the lines on the roads to clarify the presence of bike lanes.

Overall, the use of real pictures of actual roads that Cornell students and staff recognized helped the respondents in interpreting the route characteristics in one way, but seemed to also create a problem of giving “too much information”. For example, Participant #7 said, “For me the pictures help me, because I’ve walked these places before, so I know how steep the hills are. That’s the only way I could tell how steep the slope was.” On the other hand, Participant #4 said, “I found the photographs just to be distracting in general. I’m overthinking this. And the photos are not particularly clear. But, for me, it doesn’t really matter if there are cars. If there are cars, I’ll ride in the middle of the road and they’ll have to slow down. I just found the pictures to be distracting.” Participant #5, in particular, seemed to have had a problem with visual aids presented in the survey and pointed out that since she had experiences in bicycling on the actual routes that are in the pictures in the past, the pictures gave her much more information in addition to how the routes were described in words. She described her thought process as the following:

> “When you have photos you can’t help but lead the person on to think about their own personal experience, which is going to depend a lot from one to the other. So when you had the picture of the construction one (road with construction), that’s a path that I used to bike every day, but later on in that path, the moment you turn right from the junction, it becomes even more dangerous when the construction is all the way to the edge of the pavement. Then when I saw that, I thought about that and I was like, you know, immediately...even though that picture wasn’t asking me specifically about that, but that just kind of leads me to think about other things. I think in that sense it is a little bit of a distraction.”

Participant #6 made a similar comment regarding the gap between the description of a route and her experience in using the route. She said, “I just have a comment. I know that…that’s the road that I take, and I know it kind of becomes from steep to moderate, so just from personal knowledge I don’t know if by looking at that picture that I would feel that that’s a moderate steep.” As those two participants pointed out, the use of pictures of actual routes had some unintended consequences. While some participants were able to imagine hypothetical trips easily because of the pictures, others rather had difficulty in separating how each route is described in the survey and what they thought the route was like in reality. However, this is something that is nearly impossible to control for since each person may have different perceptions about the same routes depending on their experiences, and we anticipate respondents of the actual survey to go through the choice problems rather quickly and not overthink about much detail of the pictures.

Some of the other participants also pointed out that several other pictures contained elements that were not parts of characteristics of routes for the choice problems such as signs of construction, paddles and sludge on road, and parked cars on the side of road, and those elements needed to be removed. For example, one of the pictures contained parked cars on the side of road, and the following exchange was made regarding them.

> **Participant #3**: "If there are cars there instead of a bike lane, then that's like another separate issue besides just not having a bike lane. Because there's no parking on any of the other roads that you showed, you know?"
> **Facilitator**: "Why do the parked cars bother you? So if there are parked cars on the side of the road, what's your concern as a biker?"
> **Participant #3**: "Well that's another danger as a biker."
> **Participant #12**: "You could get 'doored'. There's less of a way to get around you if you're going uphill and going slowly, then everybody's trying to get around you. If there's more room, even though there's no bike lane per se, then it's more comfortable."
> **Participant #1**: "… in general, I think the pictures are kind of distracting and difficult to keep consistent. To keep the conditions consistent for what you're trying to say. There're cars in this one, there's not cars in the other one, but we're not looking at cars, we're looking at slope. Things like that."
> **Participant #10**: "I'd say for the most part I looked at the quality of the road rather than the slope…or rather than thinking about the slope itself. If there was a pothole, like a little covered divot or something in the road, that made me wonder did the other part of the road have problems. And like where there's water running in the bike lane, I don't want to be riding in that. I took more into consideration than road condition itself."
> **Participant #6**: I was just going to say having these real pictures kind of raises these other factors when otherwise I wouldn't think about, "Oh, there's going to be construction there."

From the above exchange, we learned that there needs to be much more consistency between the descriptions of routes and what the corresponding pictures actually show, and especially it is important that pictures do not contain elements that are not parts of attributes of routes. We, therefore, decided to replace troubled pictures with new ones that do not contain those elements.

In addition to route characteristics, we tried to see if there are any other important attributes of bicycle trip that can be integrated into the choice experiments. For instance, we speculated that since a bicycle can be a highly valuable property for the owner, he/she might want to park in a covered parking place to protect it from theft and hazardous environment, and lack of adequate covered parking place at the destination could affect his/her decision to bicycle. Therefore, we asked the participants regarding their preference on bicycle parking. Interestingly; however, the preference of bicycle parking turned out to vary largely. Most of the participants stated that they preferred covered parking though it also depends on convenience of locations of parking. For example, Participant #2 said**,** "I just want to qualify that it (where to park a bicycle) depends on convenience. A lot of times the uncovered parking can be right by the entrance. You just get off and get in." Similarly, Participant #11 said, "I was going to say the same thing. I don't know what the access to this space (covered parking space) is. Maybe you have to navigate a hallway or doors or something that is hard to open that it would be easier and more convenient to just be outside." Some also mentioned they were more concerned about the security for their bicycles.

For example, Participant #10 stated, “I usually bring my bike inside the building into the office. I know I’ve gotten yelled at, but I don’t feel safe. I only have the bike that I race on, so I don’t want to leave it outside… I know how to take my bike apart really easily, so I know how easy somebody could pull it all apart if they really wanted to so I don’t really want to leave it outside.”

Although parking availability seemed to somewhat matter to the participants in the decision to bicycle, we speculated from the discussion that (1) its effect on the decision to bike is much lower than other factors such as weather and hilliness of routes. and (2) it is very difficult to combine the type of parking, its convenience of access, and route characteristics into one alternative. Therefore, we decided not to include the parking facility as an additional attribute.

The facilitator proceeded to asking the participants about their suggestions on additional attributes that they thought are relevant in their decisions to bike.

> **Facilitator**: “So what would be the things that, you know—and this is again a personal choice—what are the things that you would consider universally when you’re trying to decide. When you’re trying to make a decision what do you need to know?”
> **Participant #3**: “I think of traffic. I think of the amount of parked cars on a street and everything else that you had mentioned, but those two things are my biggest ones. It’s that I don’t feel comfortable, as a bicyclist, biking on a very narrow road when there’s parked cars. It’s just not a path I would ever take.”
> **Participant #4**: “On top of those four parameters, I think traffic should be listed as another one..” [People agreeing]
> **Facilitator**: “Did I hear there was a second for that? So traffic levels, you’re saying, are a consideration?”
> **Participant #4**: “Are important.”

Other participants seemed to agree upon the suggestion and thus we decided to include traffic volume as an additional attribute for the routes.

As we approached to the end of the session, the facilitator asked the participants about the overall design of the survey. In particular, we went back to the discussion on the length of the survey.

> **Facilitator**: “Again, we can keep circling back to some of this material, but I do want to pick up on something we just asked what if we doubled it (the number of choice problems) and asked you twelve comparisons?”
> **Participant #1**: “I think ten is a good number. I mean, I know twelve isn’t that far, but I think twelve might be a little bit repetitive or you might just kind of at the end be like, ‘I feel like I’ve already done this.’ So I feel like a couple more would be good, but much more than that is a little bit…”
> **Participant #3**: “I agree that twelve is a little overwhelming.”
> **Facilitator**: “Okay. And what if it said we’re going to ask you to do ‘X’ of these and it said something like, ‘Here’s how far along you are in getting to (the end)’”
> **Participant #3**: “I think that’s good for any survey to tell you where you are in the survey. Like always do that, no matter how short or long.”

> **Participant #1**: "I've just taken psych surveys where they don't do that, and sometimes they end up being really long, and I've definitely had surveys that I just gave up because I got the really long pages of the bubbles and I'm just like, 'I'm not going to do this endlessly.'"

Subsequently, since the length of the survey the participants are willing to take seemed to depend on the incentive we give them, such as prize or entrance to lottery, we asked about what kind of prize would motivate them enough to take a similar survey of the same length in the future. The following exchange summarizes the conversation.

> **Facilitator**: "What if there were an incentive? What if there was some sort of prize or a lottery at the end of it (the survey)? Could the survey be longer (and would you take it)?"
> **Participant #3**: "I think if you offer a prize you can make it longer. One thing you could do to save you guys some money is like the one in ten people win a prize, and then you'd just have a chance of winning a prize."
> **Participant #13**: "I think that when you do just the one prize and your chances are so slim, you just go, 'Eh, not worth it.' I think that would greatly decrease the number."
> **Participant #3**: "I agree. I think even just seventy $10 gift cards are still…"
> **Participant #10**: "I usually only fill out the surveys that say that you will definitely receive something at the end."
> **Participant #14**: I'll get on this for $5. [Laughter]
> **Facilitator**: "So in its current incarnation, $5 would have definitely…"
> **Participant #14**: "A guaranteed $5. Yeah."

Although according to the conversation above, a guaranteed $5 prize would give a respondent more than enough incentive to take the survey, we did not consider this option as financially plausible for our project. We instead decided to give each respondent a free entrance to a lottery for prizes in the actual survey.

Finally, we went on to the last topic of the session. During our preliminary research on bicycling in the city of Ithaca, we learned that the city was considering about implementing a traffic policy called "Bicycle Boulevard" under which bicyclists are given a priority over motor traffic to use the road. So, as the final discussion of the session, we tried to see what the participants think of the policy in terms of incentivizing them to bicycle. We first showed a picture of Bicycle Boulevard as an example, and asked if the participants have seen something similar.

> **Facilitator**: "So let's start with…I'm curious, what are people thinking when they see this? For those who have not seen this before with certainty, what does this say to you?"
> **Participant #1**: "I just have a question. Is this like a notice 'there will be bikers on this road' sign?"
> **Participant #14**: "Bikes only." [People agreeing]
> **Facilitator**: "Bikes only. Okay."
> **Participant #9**: "I haven't seen something that large before. I've just seen the one… I forget what they're called. I would've assumed that this was a bike only except for there's cars on the road."
> **Facilitator**: "Okay. Any other…?"
> **Participant #12**: "Yeah, I would've thought it was bike only except for the people who

live in the homes. So it's like local traffic."

**Participant #5**: "With the words I guess it means boulevard. I thought that it was warning bikers that there was going to be like a big junction boulevard or hill. That's what I thought."

Although some of the participants seemed to be confident about their knowledge of Bicycle Boulevard, there appeared to be some confusion over what it actually is. They clearly did not know of "Bicycle Boulevard" after all, and moreover, many of them mistakenly interpreted the road sign as "a road dedicated solely to bicyclists". This finding is important in policy perspective because it shows that people who do not know of Bicycle Boulevard do not understand the signs for Bicycle Boulevard. Even after the concept of Bicycle Boulevard was explained to the participants, the participants seemed to be uncertain about how the policy could be effective in providing additional safety to bicyclists on roads as can be seen in the following exchange.

**Participant #5**: "I think I'm still a little bit fuzzy on what the bike boulevard actually entails. You know, you're not very clear on that. Are we asking for a reduction in speed? Are we asking for complete lanes?"

**Facilitator**: "I think the idea is like we're trying to draw out from you what the requirements are that would make you say, 'Yes. I feel comfortable with this bike boulevard idea. This is now preferable to me as a means of transit rather than just a bike lane option.' So what would have to happen to a typical road? Would it have to have the speed limit dropped and signs? Would just signs be enough? What needs to change about a road to turn it into a bike boulevard, because as you said, the definition is not entirely clear."

**Participant #6**: "The problem I think is that cars don't realize that and they try to go around you, which is when it gets really dangerous."

**Participant #12**: "Yeah, I agree. I just didn't know if it was the same here or not. Good to know."

**Facilitator**: "In general is that—I just want to make sure everyone else is sort of weighing in on this. Would this be more or less safe—the bike boulevard—than the bike lane?"

**Participant #3**: "I guess a bike boulevard with a low enough speed limit, I would be more inclined to take than a bike lane, just because bike lanes don't tend to be kept clean, especially in Ithaca with snow. There's snow and then there's puddles and the roads get pretty bad and I would be more inclined to take a road that that you share with a car with a low speed limit than I would be taking a bike lane."

**Facilitator**: "Okay. So the bike boulevard preferable, but again subject to the constraint – you said that attribute of low speed limit would be a key one for you."

As discussed in above exchange, we learned that the ambiguous definition of Bicycle Boulevard creates uncertainty in how it can provide safety to bicyclists and thus people do not understand if they would use it or not if provided. For example, we showed a picture of Bicycle Boulevard in Reno, Nevada, but like many of the few Bicycle Boulevards in the U.S., that Reno boulevard did not have any additional speed limit imposed upon for vehicular traffic; it merely had road signs and stencils that indicate that the road is a Bicycle Boulevard. Since none of the participants knew about Bicycle Boulevard and their approval of the policy was conditional on the type of traffic regulations imposed on it, we figured that it would be difficult to introduce it in the choice problems in the actual survey.

**Final Design of the Survey Instrument**

This section will describe the revisions we made on the survey instrument based on the feedback we received in the focus group session. First of all, since many of the respondents had difficulty in understanding attributes of route (especially, hilliness) and weather conditions from worded descriptions, we added visual aids for those attributes in the form of illustration. The hilliness of the route is described with a triangle sign (Figure 3-3) as it was suggested in the focus group session. Since some of the participants mentioned that they would decide whether or not to bike depending on the amount rainfall and snowfall, the weather condition is described in more detail in the new survey. Instead of probability of precipitation, the new illustration of the weather (Figure 3-4) presents the amount of rain as "rainfall" and the amount of snow as "snow depth", both in inches. As for the problem of bike lanes fading in the pictures, we superimposed the white lines (Figure 3-5). Finally, we introduced traffic volume level as an additional attribute of route, and presented it in both words and in the form of road sign (Figure 3-3). Table 3-5 is the summary of the set of attributes included in the choice problems in the final survey. Figure 3-5 shows examples of route choice problems with different level of attributes and under different weather conditions (in the order of sunny, snow, and rain).

**Figure A-3**

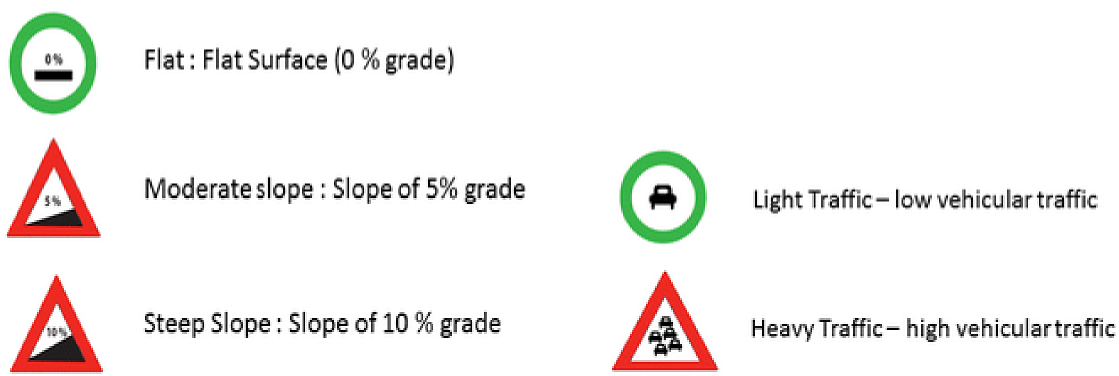
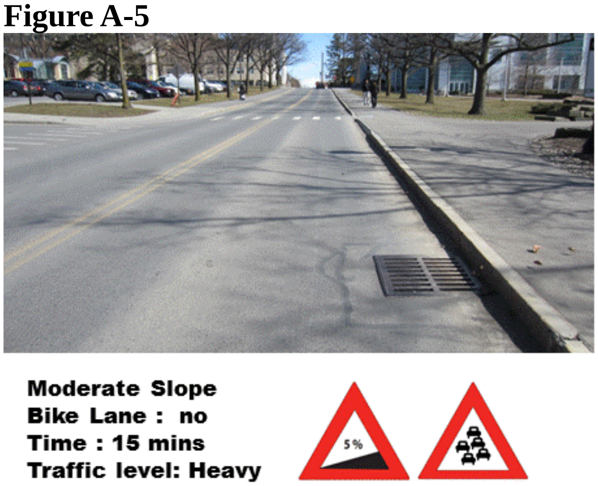


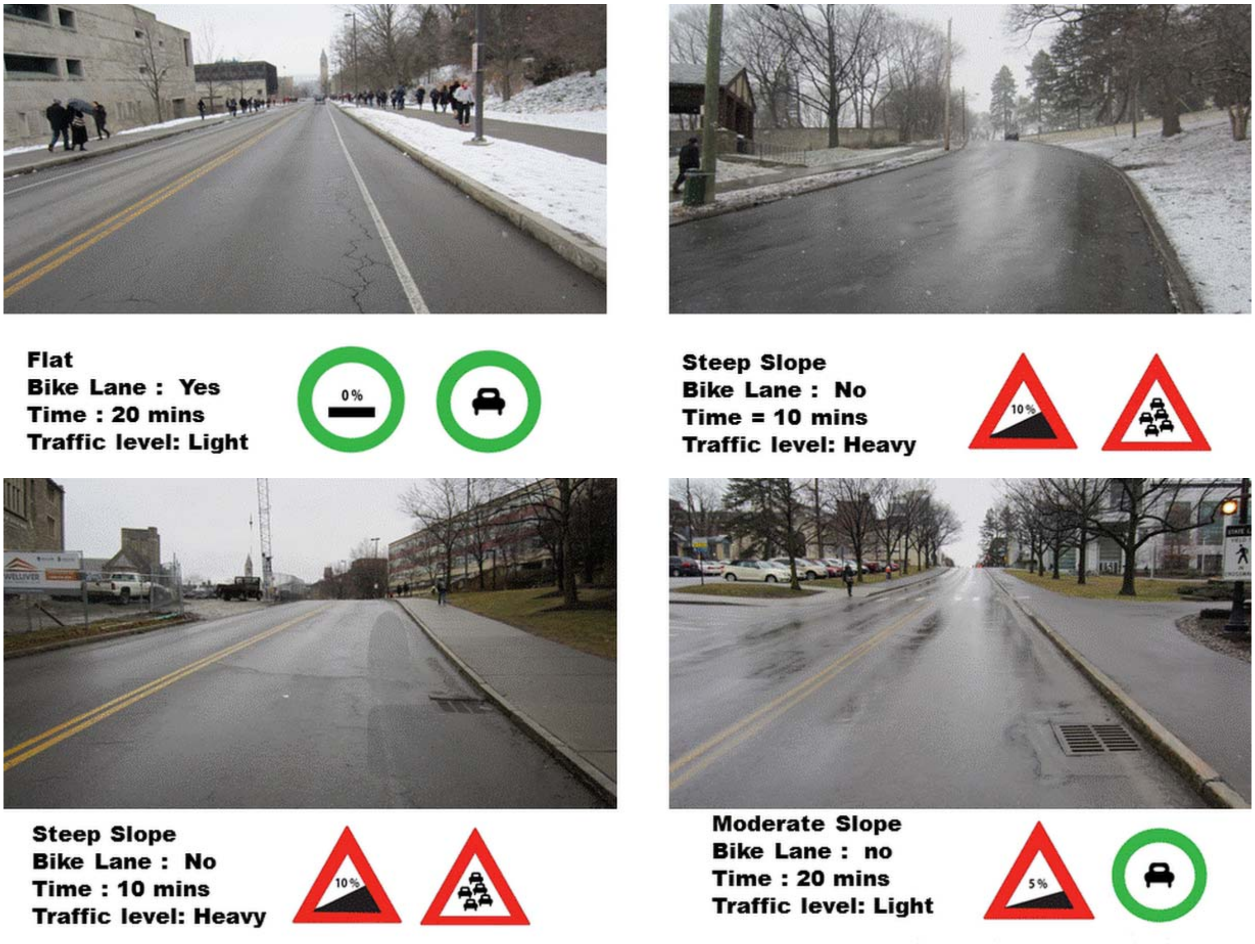

**Figure A-4**

**Table A-5:** Final attribute levels of the discrete choice experiment

| Attribute name | Attribute levels | Attribute | Attribute levels |
|---|---|---|---|
| Weather | 1. Mostly Sunny<br>2. Rain<br>3. Snow | Temperature | 1. 25 F<br>2. 35 F<br>3. 50 F<br>4. 75 F (2nd survey)<br>5. 95 F (2nd survey) |
| Hilliness | 1. 0 %<br>2. 5 %<br>3. 10 % | Traffic Volume | 1. Light<br>2. Heavy |
| Time | 1. 10 minutes<br>2. 15 minutes<br>3. 20 minutes | Rainfall | 1. 0.3 inch<br>2. 1 inch |
| Bike lane | 1. Yes<br>2. No | Snow depth | 1. 0.5 inch<br>2. 2 inch |

**Figure A-5**




## *Sample and recruitment*

We distributed the finalized survey to students, faculty, and staff of the Cornell University's Ithaca campus during the spring semester in 2013 by distributing emails with the survey link attached to several thousand students and employees[1]. The survey was quite successful and we received more than 600 responses within 2 weeks. Whereas the main survey presented a discrete choice experiment focused on colder temperatures (25F, 35F, 50F), a second (and smaller) survey had an almost identical design but with warmer temperatures (75F, 95F).

The following tables describe and summarize the data, for the main and second surveys.

[1] We do not know exactly how many students and employees were contacted due to privacy protection.

**Table A-6:** Results from the main survey (DCE with colder temperatures)

**Roughly how far do you live from campus?**

| # | Answer | | Response | % |
|---|---|---|---|---|
| 1 | Live on campus | | 183 | 27% |
| 2 | Within 1 mile | | 315 | 47% |
| 3 | 2 to 5 miles | | 126 | 19% |
| 4 | 5 to 10 miles | | 18 | 3% |
| 5 | More than 10 miles | | 28 | 4% |
| | Total | | 670 | 100% |

**DURING A TYPICAL WEEK in SPRING 2013, how many days do you travel to campus by the following forms of transportation?**

| # | Question | Never | Less than once a week | 1-2 days a week | 3-4 days a week | 5 or more days a week | Total Responses | Mean |
|---|---|---|---|---|---|---|---|---|
| 1 | Walk | 98 | 64 | 55 | 67 | 386 | 670 | 3.86 |
| 2 | Bus | 264 | 132 | 98 | 76 | 100 | 670 | 2.43 |
| 3 | Car | 363 | 142 | 89 | 31 | 45 | 670 | 1.89 |
| 4 | Carpool | 487 | 112 | 48 | 14 | 9 | 670 | 1.43 |
| 5 | Bike | 564 | 39 | 32 | 18 | 17 | 670 | 1.34 |
| 6 | Bus/Bike (bicycle on rack on bus) | 632 | 22 | 7 | 2 | 7 | 670 | 1.10 |

**When you are on campus, how do you get from one place to another?**

| # | Answer | | Response | % |
|---|---|---|---|---|
| 1 | Walk | | 621 | 93% |
| 2 | Bus | | 17 | 3% |
| 3 | Car | | 4 | 1% |
| 4 | Bike | | 28 | 4% |
| | Total | | 670 | 100% |

**How often do you exercise?**

| # | Answer | | Response | % |
|---|---|---|---|---|
| 1 | Never | | 46 | 7% |
| 2 | Less than Once a Month | | 41 | 6% |
| 3 | Once a Month | | 24 | 4% |
| 4 | 2-3 Times a Month | | 74 | 11% |
| 5 | Once a Week | | 98 | 15% |
| 6 | 2-3 Times a Week | | 246 | 37% |
| 7 | Daily | | 141 | 21% |
| | Total | | 670 | 100% |

| Which of these phrases best describes you as a cyclist? | | | | |
|---|---|---|---|---|
| # | Answer | | Response | % |
| 1 | An advanced, confident cyclist who is comfortable riding in most traffic situations | | 127 | 19% |
| 2 | An intermediate cyclist who is somewhat comfortable riding in some traffic situations | | 212 | 32% |
| 3 | A beginner cyclist who prefers to stick to the bike trails, paths and/or sidewalks | | 302 | 45% |
| 4 | I do not know how to bike | | 29 | 4% |
| | Total | | 670 | 100% |

| Do you own or have access to a bicycle? | | | | |
|---|---|---|---|---|
| # | Answer | | Response | % |
| 1 | Yes | | 308 | 53% |
| 2 | No | | 269 | 47% |
| | Total | | 577 | 100% |

| What is your age? | | | | |
|---|---|---|---|---|
| # | Answer | | Response | % |
| 1 | 17 or younger | | 0 | 0% |
| 2 | 18 - 22 | | 349 | 61% |
| 3 | 23 - 27 | | 97 | 17% |
| 4 | 28 - 32 | | 45 | 8% |
| 5 | 33 - 40 | | 29 | 5% |
| 6 | 40 - 50 | | 19 | 3% |
| 7 | 50 or older | | 34 | 6% |
| | Total | | 573 | 100% |

| What is your gender? | | | | |
|---|---|---|---|---|
| # | Answer | | Response | % |
| 1 | Male | | 232 | 40% |
| 2 | Female | | 341 | 60% |

| Which category best describes your Cornell affiliation? | | | | |
|---|---|---|---|---|
| # | Answer | | Response | % |
| 1 | Undergraduate student | | 349 | 61% |
| 2 | Graduate/Professional student | | 157 | 27% |
| 3 | Faculty | | 37 | 6% |
| 4 | Staff | | 30 | 5% |
| 5 | I am not affiliated with Cornell University | | 0 | 0% |
| | Total | | 573 | 100% |

**Table 3-7:** Results from the second survey (DCE with warmer temperatures)

| Roughly how far do you live from campus? | | | | |
|---|---|---|---|---|
| # | Answer | | Response | % |
| 1 | Live on campus | | 13 | 18% |
| 2 | Within 1 mile | | 32 | 44% |
| 3 | 2 to 5 miles | | 19 | 26% |
| 4 | 5 to 10 miles | | 5 | 7% |
| 5 | More than 10 miles | | 3 | 4% |
| | Total | | 72 | 100% |

| DURING A TYPICAL WEEK in SPRING 2013, how many days do you travel to campus by the following forms of transportation? | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| # | Question | Never | Less than once a week | 1-2 days a week | 3-4 days a week | 5 or more days a week | Total Responses | Mean |
| 1 | Walk | 16 | 5 | 12 | 6 | 33 | 72 | 3.49 |
| 2 | Bus | 30 | 11 | 7 | 10 | 14 | 72 | 2.54 |
| 3 | Car | 39 | 14 | 10 | 4 | 5 | 72 | 1.92 |
| 4 | Carpool | 60 | 9 | 2 | 1 | 0 | 72 | 1.22 |
| 5 | Bike | 61 | 4 | 3 | 3 | 1 | 72 | 1.32 |
| 6 | Bus/Bike (bicycle on rack on bus) | 67 | 4 | 0 | 1 | 0 | 72 | 1.10 |

| When you are on campus, how do you get from one place to another? | | | | |
|---|---|---|---|---|
| # | Answer | | Response | % |
| 1 | Walk | | 68 | 94% |
| 2 | Bus | | 1 | 1% |
| 3 | Car | | 1 | 1% |
| 4 | Bike | | 2 | 3% |
| | Total | | 72 | 100% |

| How often do you exercise? | | | | |
|---|---|---|---|---|
| # | Answer | | Response | % |
| 1 | Never | | 1 | 1% |
| 2 | Less than Once a Month | | 5 | 7% |
| 3 | Once a Month | | 2 | 3% |
| 4 | 2-3 Times a Month | | 10 | 14% |
| 5 | Once a Week | | 15 | 21% |
| 6 | 2-3 Times a Week | | 32 | 44% |
| 7 | Daily | | 7 | 10% |
| | Total | | 72 | 100% |

| Which of these phrases best describes you as a cyclist? | | | | |
|---|---|---|---|---|
| # | Answer | | Response | % |
| 1 | An advanced, confident cyclist who is comfortable riding in most traffic situations | | 20 | 28% |
| 2 | An intermediate cyclist who is somewhat comfortable riding in some traffic situations | | 24 | 33% |
| 3 | A beginner cyclist who prefers to stick to the bike trails, paths and/or sidewalks | | 24 | 33% |
| 4 | I do not know how to bike | | 4 | 6% |
| | Total | | 72 | 100% |

| Do you own or have access to a bicycle? | | | | |
|---|---|---|---|---|
| # | Answer | | Response | % |
| 1 | Yes | | 33 | 53% |
| 2 | No | | 29 | 47% |
| | Total | | 62 | 100% |

| What is your age? | | | | |
|---|---|---|---|---|
| # | Answer | | Response | % |
| 1 | 17 or younger | | 0 | 0% |
| 2 | 18 - 22 | | 22 | 37% |
| 3 | 23 - 27 | | 25 | 42% |
| 4 | 28 - 32 | | 7 | 12% |
| 5 | 33 - 40 | | 2 | 3% |
| 6 | 40 - 50 | | 1 | 2% |
| 7 | 50 or older | | 2 | 3% |
| | Total | | 59 | 100% |

| What is your gender? | | | | |
|---|---|---|---|---|
| # | Answer | | Response | % |
| 1 | Male | | 32 | 54% |
| 2 | Female | | 27 | 46% |
| | Total | | 59 | 100% |

| Which category best describes your Cornell affiliation? | | | | |
|---|---|---|---|---|
| # | Answer | | Response | % |
| 1 | Undergraduate student | | 18 | 31% |
| 2 | Graduate/Professional student | | 36 | 61% |
| 3 | Faculty | | 0 | 0% |
| 4 | Staff | | 5 | 8% |
| 5 | I am not affiliated with Cornell University | | 0 | 0% |
| | Total | | 59 | 100% |

The data was validated by checking consistency of the responses. After the process of data validation, 599 observations were kept for modeling purposes.

University Transportation Research Center - Region 2
Funded by the U.S. Department of Transportation

Region 2 - University Transportation Research Center
The City College of New York
Marshak Hall, Suite 910
160 Convent Avenue
New York, NY 10031
Tel: (212) 650-8050
Fax: (212) 650-8374
Website: www.utrc2.org