Transportation Research Part A 75 (2015) 217–230

Contents lists available at ScienceDirect

Transportation Research Part A

journal homepage: www.elsevier.com/locate/tra

A hybrid-choice latent-class model for the analysis of the effects of weather on cycling demand

Yutaka Motoaki [1], Ricardo A. Daziano *

School of Civil and Environmental Engineering, Cornell University, Ithaca, NY 14853, United States

ARTICLE INFO

Article history:
Received 11 July 2014
Received in revised form 3 February 2015
Accepted 18 March 2015
Available online 11 April 2015

Keywords:
Discrete choice models
Discrete heterogeneity
Latent attributes

ABSTRACT

In this paper we analyze demand for cycling using a discrete choice model with latent variables and a discrete heterogeneity distribution for the taste parameters. More specifically, we use a hybrid choice model where latent variables not only enter into utility but also inform assignment to latent classes. Using a discrete choice experiment we analyze the effects of weather (temperature, rain, and snow), cycling time, slope, cycling facilities (bike lanes), and traffic on cycling decisions by members of Cornell University (in an area with cold and snowy winters and hilly topography). We show that cyclists can be separated into two segments based on a latent factor that summarizes cycling skills and experience. Specifically, cyclists with more skills and experience are less affected by adverse weather conditions. By deriving the median of the ratio of the marginal rate of substitution for the two classes, we show that rain deters cyclists with lower skills from bicycling 2.5 times more strongly than those with better cycling skills. The median effects also show that snow is almost 4 times more deterrent to the class of less experienced cyclists. We also model the effect of external restrictions (accidents, crime, mechanical problems) and physical condition as latent factors affecting cycling choices.

1. Introduction

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 and Laube, 2002). One of the solutions to the degradation in livability provoked by automobile dependency is the adoption of non-motorized alternatives. In particular, there are several benefits associated with the use of cycling (Hillman, 1993; Sallis et al., 2004), including better air quality, no fossil fuel dependency, less noise, more efficient use of space, increased levels of physical activity, competitive speed on middle range distances, low purchase price and virtually zero operating costs (Heinen et al., 2010; Rabl and de Nazelle, 2012; Akar and Clifton, 2009).

To encourage the use of non-motorized alternatives we need to better understand the motives underlying demand. Econometric travel demand models are highly valuable for assessing the effect of policies and incentives seeking to reduce the indiscriminate use of cars. In fact, forecasting demand using discrete choice models has proved to be successful in the case of modal split among motorized alternatives. Excellent literature reviews of modeling the disaggregate demand for cycling are provided in Sener et al. (2009), Heinen et al. (2010), Li et al. (2013), Fernandez-Heredia et al. (2014) and Maldonado-Hinarejos (2014). In particular, Fernandez-Heredia et al. (2014), Habib et al. (2013) and

* Corresponding author. Tel.: +1 607 255 2018; fax: +1 607 255 9004.
E-mail addresses: ym323@cornell.edu (Y. Motoaki), daziano@cornell.edu (R.A. Daziano).
[1] Tel.: +1 607 255 2018; fax: +1 607 255 9004.

http://dx.doi.org/10.1016/j.tra.2015.03.017
0965-8564/

Maldonado-Hinarejos (2014) discuss the integration of discrete choice models of cycling decisions with subjective (latent) factors, using hybrid choice models (Walker and Ben-Akiva, 2002; Ben-Akiva et al., 2002; see also Bhat et al., 2015; Kamargianni et al., 2015).

The review papers cited above discuss some of the challenges in the application of discrete choice analysis to non-motorized options. Users of the transportation system may be motivated to cycle or walk not because of the tradeoff between cost and time (the main determinant of motorized mode choice), but because of health and environmental benefits of these alternatives. Improvements in health and in environmental footprint, for example, are positive externalities that are difficult to quantify. Additionally, several factors may discourage choice of non-motorized transportation, such as poor accessibility, safety concerns (Wilkinson, 1994; Pucher and Dijkstra, 2000), and unfavorable route and weather conditions. For instance, it is often argued that the American North East has a poor climate for cycling (see the discussion in Pucher et al., 2011).

In this paper we analyze demand for cycling using a discrete choice model with latent variables and a discrete heterogeneity distribution for the taste parameters. Our technical contribution is to use the estimator of a hybrid choice model where latent variables not only enter utility but also inform assignment to latent classes. Using a discrete choice experiment we analyze the effects of weather (temperature, rain, and snow), cycling time, slope, cycling facilities ("bike paths"), and traffic on cycling decisions by members of the community of Cornell University. We note that analyzing commuting patterns in university campuses has become a relevant case study for better understanding adoption of sustainable transportation (Shannon et al., 2006; Akar et al., 2012; Akar et al., 2013; Whalen et al., 2013; Danaf et al., 2014; Rotaris and Danielis, 2014; Erdogan et al., 2015).

In the hybrid choice modeling literature, most empirical applications consider a conditional logit kernel (for example, Kamargianni and Polydoropoulou, 2013) because of the problems associated with maximizing the complex likelihood function of the model (cf. Kamargianni et al., 2015). Habib et al. (2013), for example, work with a combination of binary logit and bivariate probit models. Other researchers have been working on the incorporation of random parameters with continuous heterogeneity distributions. For example, Maldonado-Hinarejos (2014) consider a mixed logit kernel with latent variables. Nevertheless, these authors use a limited information estimator that has poorer statistical properties compared to the joint estimator. In this work we use a full information estimator for the combination of a hybrid choice model with a latent class module. Our methodological approach differs from the work of Hurtubia et al. (2014) in that we model the effect of the latent variables on the class assignment probabilities. The remainder of the paper proceeds as follows. In Section 2, we describe our data collection method and statistics of the data. In Sections 3 and 4, we describe our results from Structural Equation Modeling and Discrete Choice Modeling, respectively. In Section 5, we conclude by discussing the results and policy implications of our study.

2. The data: cycling choices in a university environment

2.1. Motivation

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. In the last decades, however, numerous colleges have been adopting transportation demand management plans that aim at reducing motorized trips to and within campus. Bicycling improvements are particularly appropriate and effective for transportation management on university campuses for a number of reasons. University communities consist of many young and physically active commuters. If student commuters acquire environmental transportation habits it is likely that they will retain these habits after their graduation (Balsas, 2004). Nevertheless, if one expects 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 is indispensable for those planning investment decisions, it is important to identify the significant factors related to the motivation for people to bicycle.

Since the institution of a Transportation Demand Management Plan in the 1990s, Cornell University has continued to promote sustainable land use and environmentally friendly transportation plans, such as encouraging mixed land use, limiting growth within core campus boundaries, and ensuring a walkable and a cycling-friendly campus environment. Despite these efforts and despite the fact that a majority of students – 84% of graduate students and 97% of undergraduates – live within 5 miles of campus, the share of bicycle as a commuting transportation mode is very small. The 2006 Cornell University Travel Survey showed that the share of bicycle as the primary commuting mode was 1.4% for undergraduate students, 4.0% for graduate students, and 3.1% for employees (the US average share of cycling was roughly 1% in 2009; Pucher et al., 2011). There are two main factors that discourage the use of bicycles as a commuting mode at Cornell. The first is topography. The Ithaca campus of Cornell is located on a hill about 400 ft. (122 m.) above downtown of the city of Ithaca. Buffalo street, which connects downtown Ithaca and the area adjacent to campus called Collegetown, has a 15% grade (slope). The second factor is weather. The climate of the area where Cornell is located – upstate New York – is characterized by hot and wet weather in summer (in July, the average maximum temperature is 80.1 °F/26.7 °C and the average precipitation is 3.54 in./8.99 cm), and cold and snowy weather in winter (in January, the average minimum temperature is 13.9 °F/ −10.1 °C, and the average snowfall is 17.9 in./45.47 cm).

2.2. Data collection

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 travel time and existence of bike lanes, 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 of the Cornell community. 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 (What keeps/would keep you from riding a bicycle more to/ on campus?). The third category, ‘‘Improvement’’, was concerned with factors that encourage the use of bicycle (What encourages/would encourage you to ride a bike to/from campus?). 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 (Indicate whether you agree or disagree with the following statements). Categories 2–4 used 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 or 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 of 2013.[2]

The link to the dedicated Qualtrics webpage was sent to the mailing lists of the following colleges: Agriculture and Life Sciences; Architecture, Art, and Planning; Arts and Sciences; and Engineering (the email was sent by each college; we did not have information about how many students received the invitation). An iPad was the incentive for respondents to participate. Although we could not determine actual response rates, we considered response to the survey to be 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. The first sample was larger as the original motivation of the study was to assess the effect of cold weather on cycling demand. Partly as a control group, and also to add variability that would allow us to explore nonlinearity in the temperature effects, we decided to collect the second, smaller sample focused on hot temperatures. Table 1 presents the descriptive statistics of the sample after validation of the responses. Table 2 summarizes the responses to the attitudinal questions of the survey.

2.3. Discrete choice experiment

The discrete choice experiment is the key component of the survey. The experiments were based on hypothetical binary route choices for bicycling. The 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 (Fig. 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 (Fig. 2). The pictures, which were discussed in the two focus groups, were chosen to reflect the combination of the actual experimental levels of the route-specific attributes of slope (percent grade), presence of a bike lane, and traffic. The pictures also reflected general weather conditions (sun, rain, snow), but not actual experimental levels of these conditions. Using pictures to visualize the attribute levels has benefits and costs. For example, the angle and perspective of the picture may affect how slope is perceived; this is why we presented the triangle with the percent grade (which was an idea that emerged in the first focus group and tested in the second one). We also chose pictures of roads that are well known by Cornellians, so that even if the image was somewhat distorted, respondents would recognize the general conditions being depicted.

[2] Composition of the Cornell community is as follows (males/females): 7082/7371 undergraduates; 4148/3249 graduates; 2844/3735 non-academic staff; 1157/495 academic professionals.

Table 1
Descriptive statistics of the sample.

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: 18–22	351	59	332	61	19	35
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: less than once a month	34	6	31	6	3	5
Exercise frequency: once a month	19	3	19	3	0	0
Exercise frequency: 2–3 times a month	66	11	59	11	7	13
Exercise frequency: once a week	90	15	79	15	11	20
Exercise frequency: 2–3 times a week	225	38	198	36	27	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

Table 2
Descriptive statistics of the attitudinal questions.[a]

Question	Mean	Stdv
What encourages/would encourage you to ride a bike to/from campus?		
a. Dedicated bike lanes on roads	3.96	0.90
b. Bike pathways physically separated from the roadway	3.43	0.93
c. Regulating car traffic on roads	3.25	0.99
d. A campus map showing bicycle routes	3.29	0.96
e. More convenient bike parking	3.46	1.01
f. More secure or covered bike parking	3.23	1.02
g. A convenient place to shower/change clothes	3.39	1.00
h. A bicycle station on campus providing repairs/supplies	3.25	0.97
i. Bike racks on buses	3.30	1.00
j. Priority given by law to use road over vehicular traffic	3.99	1.33
Indicate whether you agree or disagree with the following statements		
a. Motor vehicle drivers seem to care little about bikers on road	3.89	1.36
b. Bicyclists seem to care little about vehicular traffic on road	3.73	1.40
c. Bicyclists seem to care little about pedestrians on street	3.90	1.55
d. I do not like to share road with bikers when I am driving	4.39	1.38
f. I am confident about my physical fitness	4.81	1.20
g. I enjoy outdoor activities (camping, fishing, jogging, etc.)	4.88	1.16
What keeps/would keep you from riding a bicycle more to/ on campus?		
a. I am not interested in biking	2.11	1.17
b. I live too far	2.86	1.23
c. I need to change clothes	2.38	1.05
d. Lack of adequate bicycle parking	2.96	1.26
e. I am worried about accidents	2.21	1.00
g. Lack of bike lanes on road	2.31	1.04
h. I am worried about possible mechanical problems that may occur, such as a flat tire	3.81	1.15
i. Severe weather conditions	3.27	1.18
j. Poor road conditions	2.81	1.19

[a] A main source for the definition of the attitudinal questions is the work of Akar and Clifton (2009).

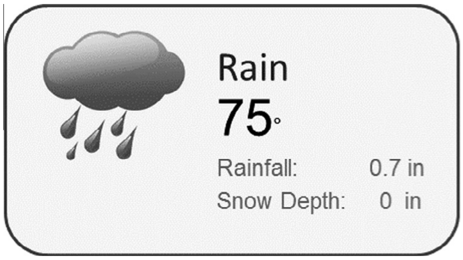

Fig. 1. Weather screenshot for rain as presented in the choice experiment.

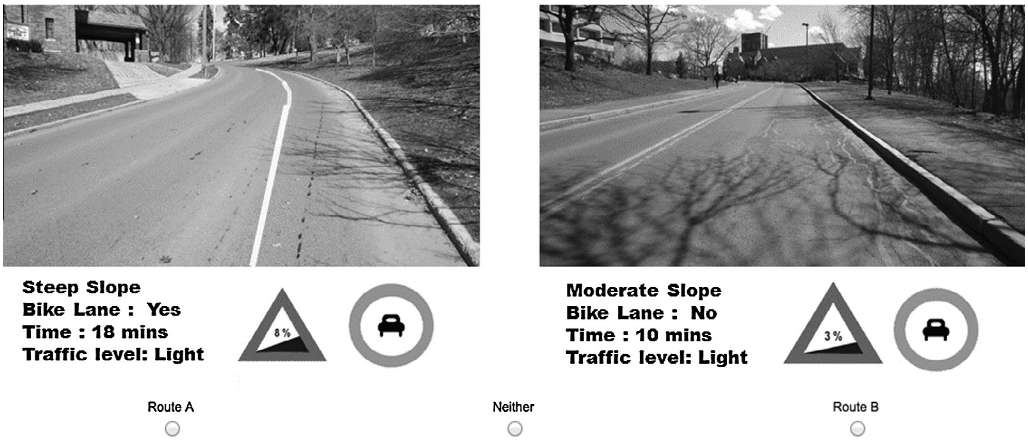

Fig. 2. Sample of route descriptions in the choice experiment.

Table 3
Attributes and attribute levels for the route choice experiment.

Attribute	Number of levels	Levels
Slope	3	0% (Flat), 3% (Moderate), 5% (Moderate), 8% (Steep), 10% (Steep)
Travel time	3	10, 15, 20 min
Bile lane	2	Yes, no
Traffic	5	Heavy, light
Weather	3	Sun, rain, snow
Precipitation: rain[a]	3	0, 0.3, 1 in. (0, 0.76, 2.54 cm)
Precipitation: snow[a,b]	3	0, 0.5, 2 in. (0, 1.27, 5.08 cm)
Traffic	5	Heavy, Light
Temperature	3/2	25 °F, 35 °F, 50 °F (Sample 1)/75 °F, 90 °F (Sample 2) −4 °C, 2 °C, 10 °C (Sample 1)/24 °C, 32 °C (Sample 2)

[a] Conditional on weather conditions (sun: zero chance of precipitation).
[b] Only for Sample 1.

Table 3 summarizes the experimental attribute and attribute levels. These levels were combined using a D-efficient design with 9 choice situations.[3]

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.

3. Psychometric modeling

In this section we discuss the main elements and results of psychometric modeling with the data collected in the attitudinal sections of the survey. First, we overview the use of structural equation modeling. Then, we identify the underlying latent variables that summarize attitudes toward bicycling.

[3] The design was based on a conditional logit model with interactions and on parameter values that we obtained in a pretest of the experiment among members of bicycling groups in the Ithaca area, as well as participants of the second focus group that included non-cyclists.

3.1. Structural equation modeling

Structural equation modeling (SEM) is a dimension-reduction 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. 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).

SEM has been applied in many research fields such as psychology, sociology, educational research, political science, and market research. Several SEM applications in travel behavior research have been conducted in the past (examples of seminal work include Tardiff, 1977; Allaman et al., 1982).

A bicycle trip has unique characteristics in that while factors such as time and cost constitute an objective appraisal, there are psychological motivations 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 factors. For a review of SEM applied to cycling demand in general, see Li et al. (2013), Fernández-Heredia et al. (2014), and Maldonado-Hinarejos (2014).

In a university context, Akar and Clifton (2009) identified three underlying variables that affect commuters' mode choice by using principal component analysis. The first factor is associated with people who see walking and cycling as an opportunity for exercising. 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 that the car parking costs on campus are high and think that they do not have many options to travel to campus. Heinen et al. (2010), conducted exploratory factor analyses to identify attitudes regarding cycling. The main attitudes these authors identified were labeled "direct trip-based benefit", "awareness", and "safety". Kamargianni and Polydoropoulou (2013) collected both revealed preference and stated preference data to study the effect of teenagers' attitudes toward walking or cycling to school in Cyprus. From the survey responses the authors defined a latent variable called "willingness to walk or cycle" as an explanatory variable. The authors found that this willingness to walk or cycle has a positive impact on the bicycle choice and walk choice, and a negative impact on car choice. Li et al. (2013) first identify eight latent factors ("need for flexibility", "sensitivity to time", "need for fixed schedule", "desire for comfort", "desire for economy", "environmental awareness", "perception toward cycling", and "willingness to use bicycle") and then apply attitudinal market segmentation based on a subset of four factors ("need for fixed schedule", "desire for comfort", "environmental consciousness", and "willingness to use bicycle"). Another example is the work of Fernandez-Heredia et al. (2014), where four latent variables were identified, namely "convenience", "pro-bike", "external restrictions" and "physical determinants". Finally, Maldonado-Hinarejos (2014) worked with the following four factors: "pro-bike", "context", "image" and "stress".

3.2. Identification of latent variables

Many SEM specifications were tested, but a Multiple Indicator and Multiple Causes (MIMIC) model with three latent variables was selected. The MIMIC model is a confirmatory factor analysis model with explanatory variables (causal indicators of the structural model). In the literature that integrates discrete choice models with SEM specifications, MIMIC models are preferred because the structural equations of the latent variables can be used for forecasting.

The underlying concepts that the data revealed in this study were labeled as: (1) bicyclist status; (2) external restrictions; and (3) physical condition. The path diagram of the selected MIMIC model is presented in Fig. 3.

The latent bicyclist 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, self-evaluation as a cyclist, and stated interest in cycling. The latent external restrictions 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. Different sociodemographic variables help to explain differences in both bicyclist status and external restrictions. For example, gender and access to a bike enter the structural equation of the two latent concepts. Gender also helps to explain differences in the latent physical condition, together with the frequency the individual exercises. Finally, note that in our model the latent external restrictions and physical condition also help to explain differences in the latent bicyclist status.

Actual parameter estimates of the MIMIC model are presented and discussed in the next section, where a discrete choice model is integrated into the SEM specification (Table 5).

4. Choice modeling

4.1. Specification and estimation of a discrete choice model with latent attributes and discrete heterogeneity

Discrete choice models describe the process of decision-making by individual agents among mutually exclusive alternatives under uncertainty (McFadden, 2001). This family of models are usually derived under an assumption of utility-

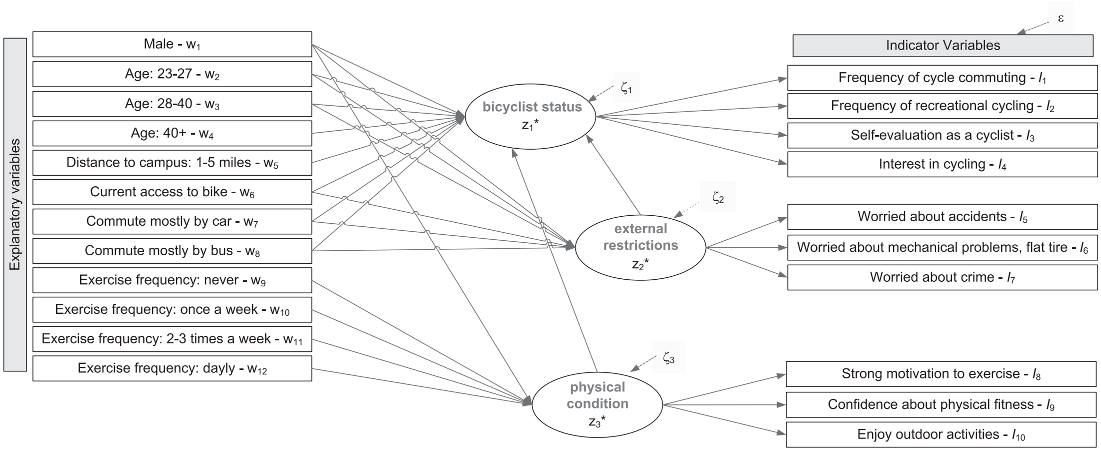

Fig. 3. Path diagram of the MIMIC model.

maximizing behavior by the decision maker. In random utility maximization (RUM) models, the researcher has limited information about the decision making process and hence an error term is added to the utility function. Hybrid choice models (HCMs) are a generalization of standard discrete choice models where different expanded models are considered simultaneously (Ben-Akiva et al., 2002). HCMs expand on standard discrete choice modeling by considering the integration of latent (unobserved) constructs, which may be abstract hypothetical constructs, such as attitudes as well as qualitative attributes that do not have a natural order. By using HCM, it is possible to incorporate an SEM, such as the one estimated in the previous section, into the responses to a discrete choice experiment, such as the route choice experiment.

In the past few years there has been an exponential increase in the use of HCM. In fact, because of the importance of non-monetary and indirect tradeoffs that characterize active transportation choices, recent studies have used HCM to analyze cycling usage (Fernandez-Heredia et al., 2014; Habib et al., 2013; Maldonado-Hinarejos, 2014). However, 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 work we propose to explore a discrete segmentation approach, where latent variables are used to identify the discrete segments of the population.

Consider the following system of latent variables for individual n:

$$\mathbf{z}_n^* = (\mathbf{I}_L - \mathbf{\Pi})^{-1}\mathbf{B}\mathbf{w}_n + (\mathbf{I}_L - \mathbf{\Pi})^{-1}\zeta_n, \zeta_n \sim \mathcal{N}(0, [(\mathbf{I}_L - \mathbf{\Pi})^{-1}]\mathbf{\Psi}[(\mathbf{I}_L - \mathbf{\Pi})^{-1}]') \tag{1}$$

$$\mathbf{U}_{tn}^{*(q)} = \mathbf{X}_{tn}\boldsymbol{\beta}_q + \mathbf{Y}_{tn}^*(\mathbf{X}_{tn}, \mathbf{z}_{1n}^*)\boldsymbol{\varrho}_q + \mathbf{\Gamma}_q\mathbf{z}_{1n}^* + \boldsymbol{\nu}_{tn}, \nu_{itn} \overset{iid}{\sim} \mathrm{EV1}(0,1) \tag{2}$$

$$\mathbf{I}_n^* = \mathbf{\Lambda}\mathbf{z}_n^* + \boldsymbol{\varepsilon}_n, \boldsymbol{\varepsilon}_n \sim \mathcal{N}(0, \mathbf{\Theta}) \tag{3}$$

$$I_{rn} = \begin{cases} 1 & \text{if } \mu_{0r} < I_{rn}^* \leqslant \mu_{1r} \\ 2 & \text{if } \mu_{1r} < I_{rn}^* \leqslant \mu_{2r} \\ \vdots & \\ M_r & \text{if } \mu_{M_r-1} < I_{rn}^* \leqslant \mu_{M_r}, \end{cases} \tag{4}$$

where $\mathbf{z}_n^*$ in Eq. (1) is a vector of L latent variables representing factors that are either hard or impossible to measure by direct observation (such as attitudes); $\mathbf{w}_n$ is a vector of M characteristics of the individual; $\mathbf{B}$ is a matrix of K unknown parameters; $\mathbf{I}_L$ represents the identity matrix of size L; the matrix $\mathbf{\Pi}$ allows for interactions among the latent variables; and ζ_n is an error term that has a multivariate normal distribution with a full covariance matrix $\mathbf{\Psi}$.

Eq. (2) represent the random utility vector of individual n in segment q and choice situation $t \in \{1, \ldots, T\}$ for J alternatives. This utility depends on the design matrix $\mathbf{X}_{tn}$ with elements $\mathbf{x}'_{tin}$ for alternative i; the vector of unknown taste parameters $\boldsymbol{\beta}_q$ for segment q; a subset of the latent variables $\mathbf{z}_{1n}^*$ the matrix of interactions $\mathbf{Y}_{tn}^*(\mathbf{X}_{tn}, \mathbf{z}_{1n}^*)$ between the observable attributes and the latent variables; the vector of unknown parameters associated with these interactions $\boldsymbol{\varrho}_q$; the matrix of unknown parameters associated with the latent variables $\mathbf{\Gamma}_q$; and an error term $\boldsymbol{\nu}_{tn}$ with elements that are iid $\mathrm{EV1}(0,1)$.

Eqs. (3) and (4) represent a system of ordered probit models for measurement of $\mathbf{z}_n^*$, where $\mathbf{I}_n^*$ is a latent vector of continuous measurement indicators of the latent variables, with elements $I_{rn}^* = \mathbf{\Lambda}_r'\mathbf{z}_n^* + \varepsilon_{rn}$; $\mathbf{\Lambda}$ is a matrix of G unknown factor loadings; $\boldsymbol{\varepsilon}_n$ is a normally distributed error term with a diagonal covariance matrix $\mathbf{\Theta}$; I_{rn} is an categorical indicator (assuming $r \in \{1, \ldots, R\}$ measurement elements) with M_r categories; and $\boldsymbol{\mu}_r = (\mu_{0r}, \ldots, \mu_{M_r})'$ is a vector of threshold parameters.

Because the structural parameters of utility are assumed to have a discrete heterogeneity distribution, with $q \in \{1, \ldots, Q\}$ possible values (segments), we need to consider a class assignment probability. In the current literature, class assignment is usually modeled as a function of characteristics of the individuals. We construct the class assignment probability as a function of a subset of the latent variables $\mathbf{z}_{2n}^*$ (cf. Hurtubia et al., 2014), following a multinomial distribution with probabilities

$$\pi_{nq} = \frac{\exp(\mathbf{z}_{2n}^{*\prime}\boldsymbol{\theta}_q)}{\sum_{q=1}^{Q}\exp(\mathbf{z}_{2n}^{*\prime}\boldsymbol{\theta}_q)}, \tag{5}$$

where π_{nq} represents the probability of individual n belonging to class q, and $\boldsymbol{\theta}_q$ is a vector of allocation parameters.

The likelihood function of the whole system is:

$$\ell(\mathbf{y},\mathbf{I};\boldsymbol{\delta}) = \prod_{n=1}^{N}\int_{z_n^*}\sum_{q=1}^{Q}\pi_{nq}\left[\prod_{t=1}^{T}P_{tn}(i_{tn}|q,\mathbf{z}_n^*)\right]\prod_{r=1}^{R}f(I_{rn})g(\mathbf{z}_n^*)d\mathbf{z}_n^*, \tag{6}$$

where $\boldsymbol{\delta}$ the complete set of unknown parameters of the system; $\mathbf{y}_n$ is $(y_{1n},\ldots,y_{Tn})'$; $\mathbf{I}_n$ is $(I_{1n},\ldots,I_{Rn})'$; $P_{tn}(i_{tn}|q,\mathbf{z}_n^*)$ is the choice probability of the chosen alternative conditional on both class q and the latent variables, we note that this choice kernel is given by the choice probability of a conditional logit model; $g(\mathbf{z}_n^*) \sim \mathcal{N}((\mathrm{I}_L-\boldsymbol{\Pi})^{-1}\mathbf{B}\mathbf{w}_n, [(\mathrm{I}_L-\boldsymbol{\Pi})^{-1}]\boldsymbol{\Psi}[(\mathrm{I}_L-\boldsymbol{\Pi})^{-1}]')$; and $f(I_{rn})$ is given by

$$f(I_{rn}=m) = \Phi\left(\frac{\mu_{mr}-\boldsymbol{\lambda}_r'\mathbf{z}_n^*}{[\boldsymbol{\Theta}]_{rr}}\right) - \Phi\left(\frac{\mu_{m-1\,r}-\boldsymbol{\lambda}_r'\mathbf{z}_n^*}{[\boldsymbol{\Theta}]_{rr}}\right) \tag{7}$$

with Φ being the CDF of a standard normal distribution, and $[\boldsymbol{\Theta}]_{rr}$ being the r-th element of the diagonal of $\boldsymbol{\Theta}$.

To propose a value for $\boldsymbol{\delta}$ we use the maximum simulated likelihood estimator $\hat{\boldsymbol{\delta}}_{MSL} = \arg\max \hat{\ell}(\boldsymbol{\delta};\mathbf{y},\mathbf{I})$, where $\hat{\ell}$ is the Monte Carlo approximation of the original likelihood.

4.2. Empirical results: point estimates

There are three components in the joint model: (1) the discrete choice kernel, (2) the class assignment model, and (3) the MIMIC model. The path diagram of the joint model is depicted in Fig. 4. As a reminder, through the MIMIC model we identified three latent concepts, one that identifies in a single index their bicyclist status, one that summarizes the physical condition of the respondents, and a third one summarizing problems that may be encountered when cycling, labeled as external restrictions. In our model, the latent bicyclist status helps in assigning individuals to one of two classes of cyclists. Whereas the latent external restrictions entered additively to the bike constant in the discrete choice kernel, for the latent physical condition we considered an interaction with the slope of the route.

Even though estimation is performed using the full information maximum likelihood estimator, we discuss results of the class assignment model first (estimates at the bottom of Table 4). We tried models with differing numbers of classes, but the only specification with satisfactory results was the one with 2 discrete segments.

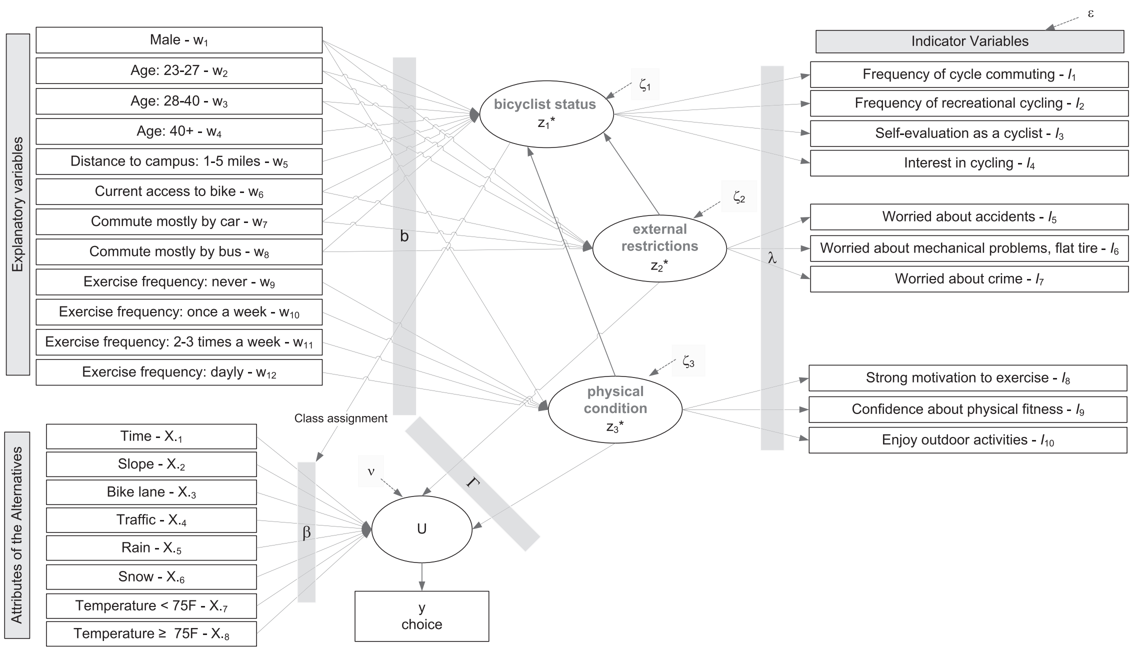

Fig. 4. Path diagram of the joint hybrid choice model.

Table 4
Latent class model with latent variables.

Parameter	Estimate	(s.e.)	*t*-Stat	*p*-Value	LB 95% CI	UB 95% CI
Class 1						
Constant	3.0771	0.6263	4.9100	0.0000	1.8495	4.3047
External restrictions	−0.3237	0.1762	−1.8400	0.0660	−0.6691	0.0218
Travel time	−0.0601	0.0069	−8.7500	0.0000	−0.0735	−0.0466
Slope	−0.1675	0.0106	−15.8600	0.0000	−0.1882	−0.1468
Slope × Physical condition	0.0259	0.0063	4.1300	0.0000	0.0136	0.0383
Bike lane	0.4338	0.0585	7.4200	0.0000	0.3192	0.5484
Heavy traffic	−0.8675	0.0569	−15.2600	0.0000	−0.9790	−0.7561
Rain	−0.8492	0.3243	−2.6200	0.0090	−1.4848	−0.2136
Snow	−0.9785	0.1669	−5.8600	0.0000	−1.3056	−0.6513
Temperature < 75 °F	0.0479	0.0234	−2.0500	0.0410	0.0020	0.0937
Temperature ⩾ 75 °F	0.0090	0.0102	−0.8900	0.3740	−0.0109	0.0290
Class 2						
Constant	2.2629	0.3236	6.9900	0.0000	1.6286	2.8972
External restrictions	−0.4922	0.0944	−5.2100	0.0000	−0.6772	−0.3071
Travel time	−0.0337	0.0131	−2.5600	0.0100	−0.0594	−0.0079
Slope	−0.2683	0.0170	−15.7800	0.0000	−0.3017	−0.2350
Slope × Physical condition	0.0251	0.0085	2.9600	0.0030	0.0085	0.0417
Bike lane	0.3973	0.0982	4.0400	0.0000	0.2047	0.5898
Heavy traffic	−0.9745	0.0885	−11.0200	0.0000	−1.1480	−0.8011
Rain	−1.1527	0.1659	−6.9500	0.0000	−1.4778	−0.8276
Snow	−2.1573	0.2465	−8.7500	0.0000	−2.6403	−1.6742
Temperature < 75 °F	−0.0096	0.0059	−1.6500	0.0990	−0.0211	0.0018
Temperature ⩾ 75 °F	−0.0215	0.0038	−5.6700	0.0000	−0.0289	−0.0141
Class assignment						
Constant class 1	−0.2870	0.1556	−1.8400	0.0650	−0.5920	0.0180
Bicycle status	0.2447	0.0883	2.7700	0.0060	0.0717	0.4177
Simulated loglikelihood	−4081.60					
Pseudo ρ^2	0.307					

The class assignment model uses class 2 as baseline. We tried several specification for the class assignment probability. The selection of the class assignment model was based on statistical significance of the latent variables and on the Bayesian Information Criterion. The resulting dependence on just the latent bicyclist status makes sense in terms of the utility estimates for each class and the policy implications. Because of the positive (and significant) bicyclist status parameter of the class assignment model, class 1 can be interpreted as 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 experienced 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 less experienced or non-cyclists (Table 4). For example, for individual n in class q and choice situation t, the structural equation of choice is

$$\begin{aligned} U_{atn}^{(q)} &= \beta_{bike}^{(q)} + \beta_{TT}^{(q)}\text{Travel.Time}_{at} + \beta_{Slope}^{(q)}\text{Slope}_{at} + \varrho_{SPC}^{(q)}\text{Slope}_{at} \times \text{Phys.Cond}_n + \beta_{BL}^{(q)}\text{Bike.Lane}_{at} + \beta_{HT}^{(q)}\text{Heavy.Traffic}_{at} \\ &\quad + \beta_{rain}^{(q)}\text{rain}_t + \beta_{snow}^{(q)}\text{snow}_t + \beta_{T<75F}^{(q)}\text{Temp} < 75\text{F}_t + \beta_{T\geqslant 75F}^{(q)}\text{Temp} \geqslant 75\text{F}_t + \gamma_{ER}^{(q)}\text{Ext.Rest}_n + \nu_{atn} \end{aligned} \tag{8}$$

$$\begin{aligned} U_{btn}^{(q)} &= \beta_{bike}^{(q)} + \beta_{TT}^{(q)}\text{Travel.Time}_{bt} + \beta_{Slope}^{(q)}\text{Slope}_{bt} + \varrho_{SPC}^{(q)}\text{Slope}_{bt} \times \text{Phys.Cond}_n + \beta_{BL}^{(q)}\text{Bike.Lane}_{bt} + \beta_{HT}^{(q)}\text{Heavy.Traffic}_{bt} \\ &\quad + \beta_{rain}^{(q)}\text{rain}_t + \beta_{snow}^{(q)}\text{snow}_t + \beta_{T<75F}^{(q)}\text{Temp} < 75\text{F}_t + \beta_{T\geqslant 75F}^{(q)}\text{Temp} \geqslant 75\text{F}_t + \gamma_{ER}^{(q)}\text{Ext.Rest}_n + \nu_{btn} \end{aligned} \tag{9}$$

$$U_{Otn}^{(q)} = \nu_{Otn} \tag{10}$$

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 have a negative impact on the likelihood of bicycling. For example, if a particular route has a steeper slope, then individuals are less likely to choose that route, either by taking a different route or by choosing not to bike. The positive parameter for the interaction Slope × Physical condition indicates that the more fit the individual is, the lower the effect of the slope is on his or her cycling decisions. Bike lanes are appreciated: the presence of a bike lane increases the probability of riding a bike.

Regarding weather impacts, if the amount of precipitation increases then individuals are more likely to choose not to cycle (as rain or snow affect all routes). The effect of colder temperatures is significant for class 1, while the effect of warmer temperatures is significant for class 2.

The effect of the latent external restrictions 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. The latent external restrictions can be thus interpreted as a variable measuring cycling anxiety for individuals that are not frequent cyclists.

Table 5
MIMIC component estimates.

Measurement equation: Λ (R squared)	Estimate	s.e.	t-Stat	p-Value	LB 95% CI	UB 95% CI
Bicycle status	Reliability: 0.755 (α), 0.765 (ω)					
Frequency of cycle commuting (0.803)	1.000					
Frequency of recreational cycling (0.449)	0.661	0.051	12.893	0.000	0.561	0.761
Self-evaluation as a cyclist (0.661)	0.861	0.056	15.329	0.000	0.751	0.971
Interest in cycling (0.501)	0.711	0.050	14.277	0.000	0.613	0.809
External restrictions	Reliability: 0.728 (α), 0.733 (ω)					
Worried about accidents (0.544)	1.000					
Worried about mechanical problems (0.518)	0.975	0.057	17.071	0.000	0.863	1.087
Worried about crime (0.453)	0.908	0.055	16.486	0.000	0.800	1.016
Physical condition	Reliability: 0.778 (α), 0.780 (ω)					
Strong motivation to exercise (0.834)	1.000					
Confidence about physical fitness (0.722)	0.898	0.005	17.880	0.000	0.800	0.996
Enjoy outdoor activities (0.436)	0.644	0.036	17.685	0.000	0.573	0.715
Bicycle status (0.578)						
Male	0.355	0.090	3.927	0.000	0.179	0.531
Age: 23–27	0.341	0.120	2.842	0.004	0.106	0.576
Age: 28–40	0.488	0.147	3.318	0.001	0.200	0.776
Age: 40+	0.455	0.164	2.771	0.006	0.134	0.776
Distance to campus: 1–5 miles	0.278	0.121	2.306	0.021	0.041	0.515
Latent physical condition	0.211	0.046	4.609	0.000	0.121	0.301
Latent external restrictions	−0.427	0.062	−6.919	0.000	−0.549	−0.305
Access to bike	0.989	0.102	9.692	0.000	0.789	1.189
Commute mostly by car	−0.389	0.135	−2.872	0.004	−0.654	−0.124
Commute mostly by bus	−0.239	0.098	−2.435	0.015	−0.431	−0.047
External restrictions (0.105)						
Male	−0.306	0.074	−4.122	0.000	−0.451	−0.161
Age: 23–27	−0.204	0.105	−1.950	0.051	−0.410	0.002
Age: 28–40	−0.295	0.123	−2.404	0.016	−0.536	−0.054
Access to bike	−0.213	0.074	−2.884	0.004	−0.358	−0.068
Commute mostly by car	0.173	0.118	1.471	0.141	−0.058	0.404
Commute mostly by bus	0.174	0.084	2.068	0.039	0.009	0.339
Physical condition (0.426)						
Male	0.195	0.088	2.222	0.026	0.023	0.367
Exercise frequency: never	−0.535	0.185	−2.888	0.004	−0.898	−0.172
Exercise frequency: once a week	0.568	0.147	3.863	0.000	0.280	0.856
Exercise frequency: 2–3 times a week	1.155	0.126	9.137	0.000	0.908	1.402
Exercise frequency: daily	1.889	0.145	13.015	0.000	1.605	2.173
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 $\leqslant$ 0.05	0.049					

The measurement equations of the embedded MIMIC model 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. Bicyclist 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 bicyclist 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 bicyclist 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.

4.3. Empirical results: inference

To analyze the taste differences between the two identified classes, we calculated the ratio of the marginal rates of substitution of each variable with respect to time. Both the median and mean of these ratios are presented in Table 6. We also present the lower and upper bounds of the 95% Krinsky–Robb confidence intervals using 100,000 repetitions for the required simulation.

The median effects show that the degree of steepness of a slope deters less experienced cyclists almost 3 times more strongly relative to experienced cyclists, the presence of traffic discourages bicycling twice as much, rain 2.5 times as much,

Table 6
Ratio of the marginal rate of substitution with respect to travel time for class 2 vs. class 1.

Variable	Median	Mean	LB 95% CI	UB 95% CI
Bike lane	1.6387	2.4489	0.5015	4.3964
Heavy traffic	1.9921	3.2415	0.7741	5.7088
Slope	2.8360	4.6637	1.0203	8.3072
Rain	2.5480	6.4246	0.4239	12.4253
Snow	3.9785	6.6600	1.2050	12.1151

and snow almost 4 times more. An interesting result is that bike lanes are more appreciated by less experienced cyclists and non-cyclists. (The benefit of the presence of bike lanes is valued 1.6 times higher by less experienced cyclists.) This bike-lane appreciation is consistent with the hypothesis that experienced cyclists care less about the availability of bike lanes because they have higher skills than non-cyclists (e.g., Taylor and Mahmassani, 1996; Hunt and Abraham, 2007).

In order to provide a visual representation of the nonlinear effects of the deterrents on route choice, we plotted the changes in choice probabilities against the level of the deterrents. We chose the first respondent in choice situation 1 in the data as the representative agent in this analysis. Route A is characterized by 20 min of travel time, slope gradient of 0%, and light traffic. Route B is characterized by 10 min of travel time, slope gradient of 10%, and heavy traffic. Fig. 2 shows the choice probability for route A as a function of slope for the two classes. At the original conditions, the probability of choosing route A is close to 0.80 for both classes (route A takes longer, but is much more attractive than route B in terms of slope and traffic). When slope is considered to be the same for both routes (10%), the class of more experienced cyclist now choose route A with a probability of 0.5, while the choice probability for less experienced cyclists is slightly above 0.2.

For analyzing the effects of precipitation, for each class we display the choice probability of bicycling (probability of choosing either route A or route B). Because of the binary nature of this choice probability, the right vertical axis can be read as the probability of opting out (probability of choosing not to bike). Fig. 3 shows the effect of rain on the choice probability of bicycling. Fig. 4 summarizes the effect of snow on the choice probability of bicycling. In the case of rain, if the individual belongs to the class of experienced cyclists then when there is no rain the probability of cycling is 0.9. With 2 in. of rain, the experienced cyclist is still more like to cycle (probability of 0.6) than not to cycle. With 1 inch of rain in the forecast, the class of less experienced cyclists are indifferent between cycling and not cycling (probability of 0.5).

From the three sets of graphs it is possible to see that, in general, the marginal probability effects are less pronounced (and almost constant) for the class of more experienced cyclists. In fact, a linear approximation of the marginal probability effects seems reasonable for class 1, whereas nonlinear effects are evident for the class of less experienced and non cyclists (especially for snow where a dramatic fall in the probability of biking is obtained). Because the effects of the cycling deterrents are lower for experienced cyclists, the preferences of this class can be seen as the benchmark when more cycling experience is gained. For example, snow precipitation is a big deterrent only for less experienced cyclists. Communities with weather conditions that are traditionally considered as poor for encouraging biking can make efforts for increasing the possibilities for gaining more experience in biking.

Another set of similar plots using a randomly chosen respondent in a different choice situation is presented in Figs. 5–7.

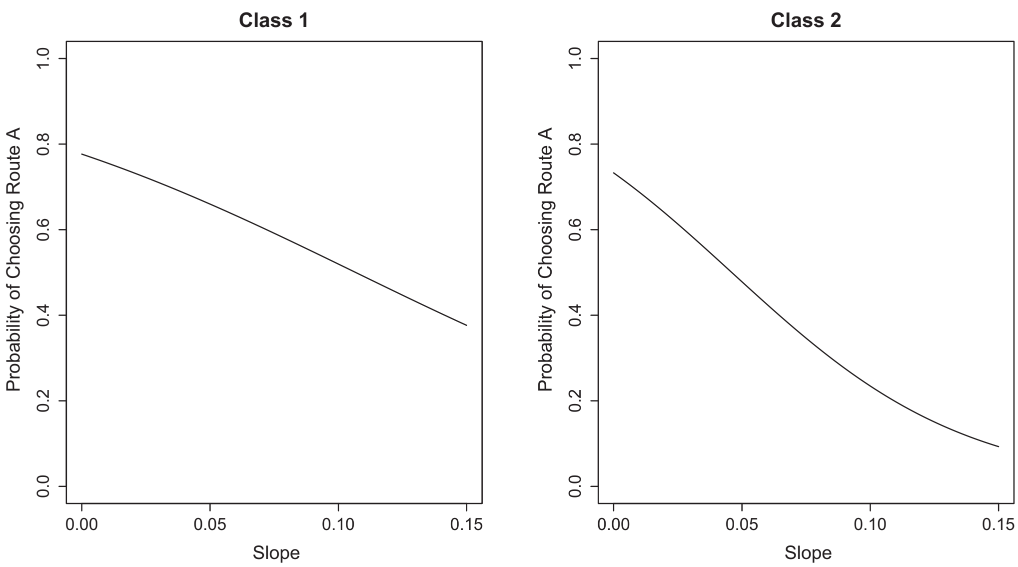

Fig. 5. Choice probability of choosing route A as a function of slope.

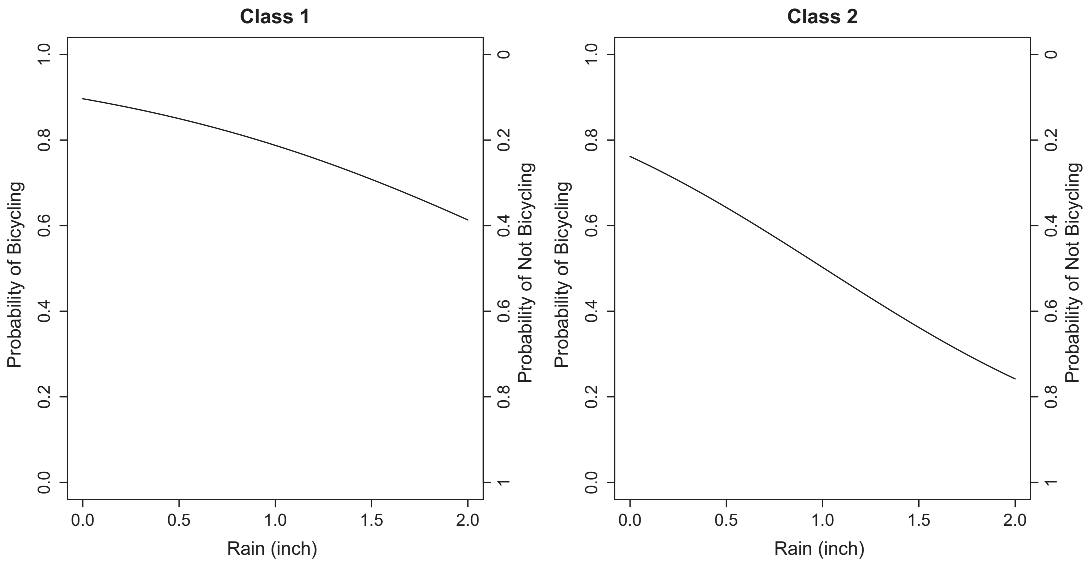

Fig. 6. Choice probability of bicycling vs. not bicycling as a function of rain.

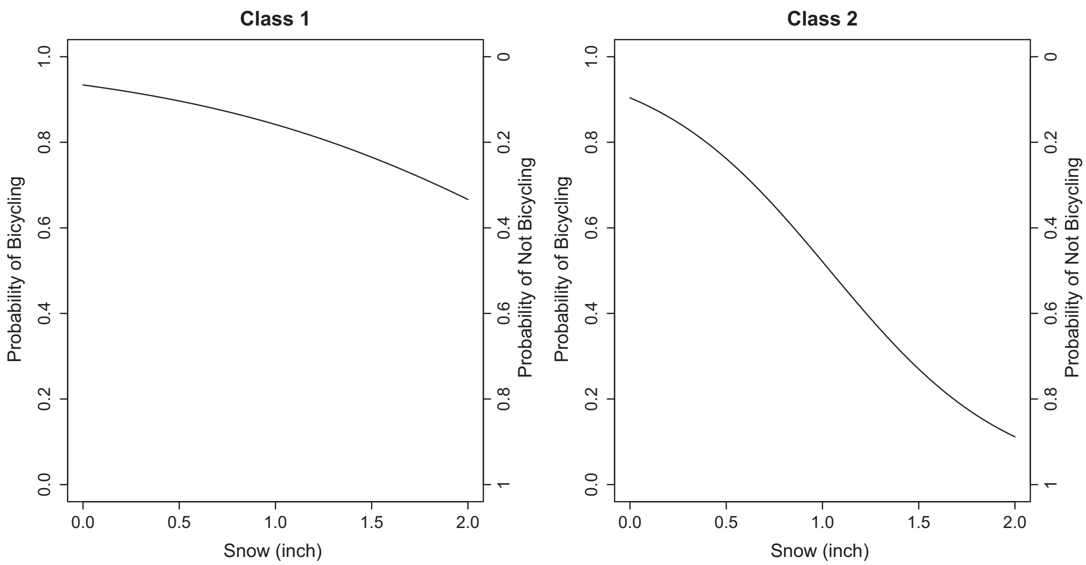

Fig. 7. Choice probability of bicycling vs. not bicycling as a function of snow.

5. Conclusions

We designed a web-based survey aimed at collecting attitudinal data about cycling, including behavioral intentions of using a bike for commuting. The experimental design has some new elements when compared with previous empirical applications of cycling route choice modeling. Besides common attributes (time, slope/topography, cycling infrastructure) we included variables related to weather conditions (temperature, amount of expected rain, amount of expected snow). In addition, for each choice situation we considered the possibility of opting out (i.e. choosing not to cycle). The survey was applied to members of the Cornell University community in Ithaca, NY (an area characterized by its hills as well as by hot and wet weather in summer, and cold and snowy weather in winter). More than 600 individuals responded to the survey, for a total of 599 valid responses.

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 bicyclist status, external restrictions/cycling anxiety, and physical condition). In particular, we derived a latent class model with a class assignment mechanism based on the latent bicyclist status of the respondent.

Two segments were identified: (1) more-skilled and experienced cyclists, vs. (2) less-skilled and non-cyclists. The two segments have different sensitivities to the factors that may encourage or discourage riding a bike. Because cycling route decisions do not involve any direct monetary cost, to analyze differences in the taste parameters we proposed the ratio of the marginal rate of substitution with respect to travel time. The median of the ratios reveal that slope inclination is considered almost 3 times worse by less-skilled cyclists. Less-skilled cyclists are not only affected twice as much by heavy traffic, but also consider rain to be 2.5 times more bothersome and snow almost 4 times more bothersome relative to more-skilled cyclists.

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. In particular, having more bicycle facilities results in a higher share of cycling (Akar and Clifton, 2009; Barnes and Thompson, 2006; Pucher and Buehler (2006); Klobucar and Fricker, 2007; Dill and 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. 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. People with less skills and experience in bicycling apparently value bike lanes as safety measures so that 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 bicyclist status (i.e. using a bike infrequently, or mostly for recreational purposes). Because the marginal probability effects of cycling deterrents are much lower for more experienced cyclists, the creation of opportunities for the community to gain experience in cycling emerges as a promising policy. In this respect, bike lanes not only have a direct effect in increasing the likelihood of cycling but also an indirect effect on neutralizing factors that discourage the use of a bike (rain, snow, slope, "bike anxiety").

Acknowledgement

This research is based upon work supported by the University Transportation Research Center Region II, Grant NYSDOT ID 49997-35-24.

References

Akar, G., Clifton, K.J., 2009. Influence of individual perceptions and bicycle infrastructure on decision to bike. Transport. Res. Rec. 2140, 165–172.

Akar, G., Flynn, C., Namgung, M., 2012. Travel Choices and Links to Transportation Demand Management. Case Study: Ohio State University. Transportation Research Record, No. 2319, 2012, pp. 77–85.

Akar, G., Fischer, N., Namgung, M., 2013. Bicycling choice and gender, case study: the Ohio state university. Int. J. Sust. Transport. 7 (5), 347–365.

Allaman, P.M., Tardiff, T.J., Dunbar, F.C., 1982. New Approaches to Understanding Travel Behavior. NCHRP Report 250. Transportation Research Board, National Research Council, Washington, DC.

Balsas, C., 2004. Sustainable transportation planning on college campuses. Transport Policy 10 (1), 35–49.

Barnes, G., 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.

Ben-Akiva, M., McFadden, D., Train, K., Walker, J., Bhat, C., Bierlaire, M., Bolduc, D., Boersch-Supan, A., Brownstone, D., Bunch, D., Daly, A., de Palma, A., Gopinath, D., Karlstrom, A., Munizaga, M.A., 2002. Hybrid choice models: progress and challenges. Market. Lett. 13 (3), 163–175.

Bhat, C.R., Dubey, S., Nagel, K., 2015. Introducing non-normality of latent psychological constructs in choice modeling with an application to bicyclist route choice. Paper Presented at the 94th Annual Meeting of the Transportation Research Board, Washington DC.

Danaf, M., Abou-Zeid, M., Kaysi, I., 2014. Modeling travel choices of students at a private, urban university: insights and policy implications. Case Stud. Transport Policy 2 (3), 142–152.

Dill, J., Voros, K., 2007. Factors affecting bicycling demand: initial survey findings from the Portland, Oregon region. Transport. Res. Rec. 203, 9–17.

Erdogan, S., Cirillo, C., Tremblay, J.M., 2015. Ridesharing as a green commute alternative: a campus case study. Int. J. Sust. Transport. 9 (5), 377–388.

Fernandez-Heredia, A., Monzon, A., Jara-Diaz, S., 2014. Understanding cyclists' perceptions, keys for a successful bicycle promotion. Transport. Res. Part A 63, 1–11.

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 Rev. 30 (1), 59–96.

Hillman, M., 1993. Cycling and the promotion of health. Policy Stud. 14 (2), 49–58.

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

Hurtubia, R., Nguyen, M.H., Glerum, Bierlaire, M., 2014. Integrating psychometric indicators in latent class choice models. Transport. Res. Part A: Policy Pract. 64, 135–146.

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.

Kamargianni, M., Dubey, S., Polydoropoulou, A., Bhat, C.R., 2015. Investigating subjective and objective factors influencing teenagers' school travel mode choice: integrated choice and latent variable model. Paper Presented at the 94th Annual Meeting of the Transportation Research Board, Washington DC. Klobucar, M. S., & Fricker, J. D. (2007). Network evaluation tool to improve real and perceived bicycle safety. Transportation Research Record, 2031, 25-33.

Klobucar, M.S., Fricker, J.D., 2007. Network evaluation tool to improve real and perceived bicycle safety. Transport. Res. Rec. 2031, 25–33.

Li, Z., Wang, W., Yang, C., Ragland, D.R., 2013. Bicycle commuting market analysis using attitudinal market segmentation approach. Transport. Res. Part A 47, 56–68.

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

Maldonado-Hinarejos, R., 2014. Modeling with an attitude: models that consider attitudes and latent variables. Unpublished Paper Presented at Presented at the 85th TRB Annual Meeting, Washington, DC.

McFadden, D., 2001. Economic choices. Am. Econ. Rev. 91 (3), 351–378.
Pucher, J., Dijkstra, L., 2000. Making walking and cycling safer: lessons from Europe. Transport. Quart. 54, 25–50.
Pucher, J., Buehler, R., 2006. Why Canadians cycle more than Americans: a comparative analysis of bicycling trends and policies. Transp. 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. Transport. Res. Part A: Policy Pract. 45 (6), 451–475.
Rabl, A., de Nazelle, A., 2012. Benefits of shift from car to active transport. Transport Policy 19 (1), 121–131.
Rotaris, L., Danielis, R., 2014. The impact of transportation demand management policies on commuting to college facilities: a case study at the University of Trieste, Italy. Transport. Res. Part A 67, 127–140.
Sallis, J.F., Frank, L.D., Saelens, B.E., Kraft, M.K., 2004. Active transportation and physical activity: opportunities for collaboration on transportation and public health research. Transport. Res. Part A: Policy Practice 38 (4), 249–268.
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.
Shannon, T., Giles-Corti, B., Pikora, T., Bulsara, M., Shilton, T., Bull, F., 2006. Active commuting in a university setting: assessing commuting habits and potential for modal change. Transport Policy 13 (3), 240–253.
Stinson, M.A., Bhat, C.R., 2003. Commuter bicyclist route choice: analysis using a stated preference survey. Transport. Res. Rec. 1828, 107–115.
Tardiff, T.J., 1977. Causal inferences involving transportation attitudes and behavior. Transport. Res. 11 (6), 397–404.
Taylor, D., Mahmassani, H., 1996. Analysis of stated preferences for intermodal bicycle-transit interfaces. Transport. Res. Rec. 1556, 86–95.
Walker, J., Ben-Akiva, M., 2002. Generalized random utility model. Math. Soc. Sci. 43 (3), 303–343.
Whalen, K.E., Páez, A., Carrasco, J.A., 2013. Mode choice of university students commuting to school and the role of active travel. J. Transport Geogr. 31, 132–142.
Wilkinson, W.C., 1994. Selecting Roadway Design Treatments to Accommodate Bicycles. McLean, Va: U.S. Federal Highway Administration, Office of Safety and Traffic Operations Research and Development, Report FHWA-RD-92-073.