Transportation (2013) 40:887–902
DOI 10.1007/s11116-013-9453-x

Bicyclist commuters' choice of on-street versus off-street route segments

Lei Kang · Jon D. Fricker

Published online: 17 February 2013
© Springer Science+Business Media New York 2013

Abstract When using limited funds on bicycle facilities, it would be helpful to know the extent to which a new facility will be used. If a bicycle lane is added to a street, how many bicyclists will no longer use the adjacent sidewalk? If a separate bicycle path is constructed, how many bicyclists will move from the street or sidewalk? This study seeks to identify factors that explain a bicyclist's choice between available facility choices—off-street (sidewalk and bicycle path) or on-street (bicycle lane and roadway). This paper investigates these issues through a survey of bicyclists headed to Purdue University in West Lafayette, IN, USA. The first data collected to address these questions were "site-based". Bicyclists were interviewed on campus at the end of their trips and asked which part of the cross-sections along their routes they had used—on-street or off-street. The characteristics of a particular cross-section of street right-of-way were then compared against the characteristics of each bicyclist and his/her observed choice of street, sidewalk, lane, or path. Later, "route-based" serial data were also added. The study developed a mixed logit model to analyze the bicyclists' facility preferences and capture the unobserved heterogeneity across the population. Effective sidewalk width, traffic signals, segment length, road functional class, street pavement condition, and one-way street configuration were found to be statistically significant. A bicycle path is found to be more attractive than a bicycle lane. Predictions from the model can indicate where investments in particular bicycle facilities would have the most desirable response from bicyclists.

Keywords Bicycle facilities · Bicycle lanes · Bicycle paths · Mixed logit · Marginal effects · Prediction analysis

L. Kang · J. D. Fricker (✉)
School of Civil Engineering, Purdue University, 550 Stadium Mall Drive, West Lafayette, IN 47907-2051, USA
e-mail: fricker@purdue.edu

L. Kang
e-mail: lkang119@gmail.com

Introduction

Bicycle commuting has been shown to be an effective way to reduce congestion and improve public health (de Geus et al. 2007; Lawlor et al. 2003). Dill and Carr (2003) and Nelson and Allen (1997) demonstrated that higher bicycle commuting rates were associated with higher levels of bicycle infrastructure (e.g., bicycle lanes and bicycle paths). Two national polls reported that 50 % of respondents supported requirements that streets include bicycle lanes or paths, even if it meant less space for cars and trucks (League of Illinois Bicyclists 2003). Although there has been a significant increase in funding for building new facilities for bicyclists, policymakers and city planners still need to know where to allocate resources towards bicycle facilities, because road and sidewalk space is often limited.

Given limited resources, it is probably impossible to add bicycle lanes or bicycle paths along every street. Thus, investments should be made where the new facility will be used. In most situations, bicyclists can either choose to ride on the street close to motor vehicles or ride on the sidewalk, where pedestrians and other factors reduce their speed. Sometimes, however, even though a street is equipped with a facility like a bicycle lane, some bicyclists continue to ride on the sidewalk, as in Davis, California and Ottawa and Toronto, Canada (HCM 2010; Aultman-Hall and Adams 1998). This study seeks to identify factors that explain a bicyclist's choice between available facility choices—off-street (sidewalk and bicycle path) or on-street (bicycle lane and roadway). This could help the facility planners make the best bicycle facility investment decision for a given situation.

Literature review

To build a bicycle-friendly community with limited funds, city engineers need to determine where best to invest those funds in bicycle facilities. It would be helpful to be able to predict how bicyclists would react to the introduction of a particular kind of bicycle facility, based on different environments and user populations.

For the purposes of our study, we consider two types of bicycle facilities—on-street and off-street. On-street facilities include bicycle lanes and roadways. A bicycle lane is defined "as a portion of roadway that has been designated by striping, signing, and/or pavement markings for the preferential and exclusive use of bicyclists." (Highway capacity manual 1994) Off-street bicycle facilities include bicycle paths and sidewalks. A bicycle path is defined as "a bikeway physically separated from motorized traffic by an open space or barrier" (Highway capacity manual 1994), either shared with pedestrians or limited to exclusive bicycle use. A sidewalk is a "path designed for pedestrians, usually paved and alongside a street".

The most direct impact of improved bicycle facilities is an increase in the number of persons using bicycles. According to a stated preference study conducted by the Federal Highway Administration (1992), people indicated that having a bicycle path or bicycle lane would encourage them to bicycle more. Based on analyses in 43 large cities, Dill and Carr (2003) built several regression models using percentage of workers commuting by bicycle as the dependent variable, and the number of bicycle lanes per square mile and various socio-economic measures as independent variables. They found positive association between bicycle facilities and bicycle commuting. Nelson and Allen (1997) indicated that each additional mile of bikeway per 100,000 residents would increase bicycle commuting 0.069 % while holding other factors constant. Krizek et al. (2009) found that areas

near new bicycle facilities showed considerably more of an increase in bicycle mode share than areas farther away. After a comprehensive review, Heinen et al. (2010) concluded that the presence of bicycle infrastructure might not only result in more cycling, but higher cycling frequency could also stimulate the construction of bicycle infrastructure. However, Parkin et al. (2008) found that the provision of infrastructure alone appears insufficient to engender higher levels of cycling. Buehler and Pucher (2012) collected length data for bicycle lanes and paths in the 90 largest USA cities. They concluded that cities with a greater supply of bicycle paths and lanes have significantly higher bicycle commute rates—even when controlling for land use, climate, socioeconomic factors, gasoline prices, public transport supply, and cycling safety. Their conclusions, based on city-level data, are in accordance with previous studies. What's more, through estimated elasticities, they indicated that both off-street paths and on-street lanes have a similar positive association with bicycle commute rates in USA cities.

Various comparative studies had been done to discuss the safety and operational impact of bicycle facilities. One minute spent in a shared curb lane is 4.1 times as onerous as a minute spent on a designated bicycle lane (Hunt and Abraham 2007). Adding a bicycle lane would increase the perceived safety of a bicyclist, thus improving the level-of-service of a segment (Harkey et al. 1998). Based on the videotapes of almost 4,600 bicyclists from 48 sites in Santa Barbara, Gainesville, and Austin, Hunter et al. (1999) found that wrong-way riding was much more prevalent on sidewalks at wide curb lane sites than at bicycle lane sites. Moreover, bicyclists on a sidewalk or bicycle path would still incur, on average, 1.8 times as much risk as those on the roadway because of blind conflicts at intersections (Wachtel and Lewiston 1994). This finding is consistent with the work by Meuleners et al. (2007).

Several studies have indicated that bicyclists' preferences for bicycle facilities differ according to their purpose (e.g., recreation vs. commuting), riding skill (e.g., experienced vs. inexperienced). (Antonakos 1994; Taylor and Mahmassani 1996; Harvey et al. 2008; Stinson and Bhat 2005; Hunt and Abraham 2007) and sex (Krizek 2005). In general, bicyclists would prefer a bicycle path to a bicycle lane or roadway (Stinson and Bhat 2005), but in most cases, there is insufficient space in which to build a bicycle path. Based on 167 respondents' perceptions, Tilahun et al. (2007) found that users would pay the highest price for a designated bicycle lane, followed by the absence of parking on the street and by taking a bicycle lane off-road. More determinants for commuting by bicycle and a bicycle facility preference analysis can be found in comprehensive reviews by Heinen et al. (2010) and Pucher et al. (2010).

In a study by Ford et al. (2011), serial data were collected and two binary logit models were estimated to identify the critical segments where bicyclists prefer off-street to on-street components, and vice versa. However, it was found that bicyclists' preferences regarding bicycle facilities are greatly affected by path-level decisions. We need a better understanding of segment-level choice analysis, which could help city engineers decide the location of new facilities. This paper contributes to current knowledge by adding a disaggregate specification, to connect commuting bicyclist preferences with variations in bicycle facility supply and segment characteristics.

Data

Intercept surveys were conducted at bicycle racks throughout the campus of Purdue University, West Lafayette, Indiana USA during the fall semesters of years 2006–2008 to

obtain revealed preference cross-section choice data. The university-oriented database was created because (a) it was easy to conduct the interviews, (b) the university is the major destination in the area for bicycle commuters of various ages, and (c) there is a clear need for a systematic way to plan improvements for bicycling in the street network adjacent to campus. The respondents were bicyclists who began their trips off-campus and had just arrived at bicycle racks. The interview questions were:

1. When you bicycled to campus just now, where did that trip start? Give an address, living group name, or intersection of streets nearest the start.
2. What streets did you use to reach your destination on campus? List the streets or show on the map provided, if that is easier.
3. If, at any point during that trip, you rode on the sidewalk, indicate where that happened. Again, create a list or use a map, as preferred. Idea: Mark links where sidewalks were used with a "W".

The survey is described in more detail by Ford et al. (2011). If a bicycle lane or the street was chosen, it is regarded as on-street choice. If a bicycle path or a sidewalk was chosen, it is treated as off-street choice. This survey provides us with information on the series of on-street/off-street choices made by each bicyclist. In most cases, something about the bicycling environment changed during the trip. Residential streets fed into arterials, a bicycle lane began or ended, etc. Any such change offered an opportunity—or a reason—for the bicyclist to change his/her cross-section choice. Even if a bicyclist using a sidewalk continued to use the sidewalk as some observable factor changed, there may be information to be learned from that case.

A total of 931 'observations' were collected from 178 bicyclists. Each observation is associated with a segment of the bicyclist's trip. At the beginning of a segment, the bicyclist would make a decision whether to use an on-street facility or an off-street facility, based roadway characteristics, personal perceptions of safety, and other factors listed in Table 1. Each outcome comprised an observation. In our database, a "segment" is defined as that part of the route between two intersections at which a bicyclist would face a new choice of cross-section.

It would be also useful to investigate the choice between sidewalk and bicycle path when both these off-street options are available. However, in our database, only 14 cases provide both off-street options. Because such a small sample could not reveal any significant distinction between these two choices, we decide to aggregate them together as an off-street choice. For on-street choices, we assume that a bicyclist who chooses to use an on-street facility will use the bicycle lane where one exists. Observations with only one option (i.e., a street without bicycle lane, bicycle path, or sidewalk) are not included in the model.

Data on segment characteristics, including 85th percentile traffic speeds, curb lane widths, and average daily traffic, were obtained from the City of West Lafayette and the Indiana Department of Transportation's Annual Average Daily Traffic maps. Other segment characteristics data came from field visits. Weather data were added from online weather records for the days and times of the interviews. The bicycle compatibility index (BCI), originally developed by Harkey et al. (1998), was incorporated into our study. This index is associated with bicycle level of service on a roadway, which could be used to quantify a bicyclist's perception of safety (or risk) along a segment of roadway. Klobucar and Fricker (2007) have applied BCI in assessing the level of service offered to bicyclists in West Lafayette's street network. The typical BCI value would range from 1 to 5. The higher this value is, the more dangerous a bicyclist would perceive the on-street facility.

Table 1 Variables available for modeling

Dependent variable	Cross-section choice: 1 if on-street facility being selected, 0 off-street facility being selected
Bicyclist characteristics	Sex indicator variable: 1 if female, 0 if male
	Bicyclist age
	Helmet indicator variable: 1 if helmet used, 0 otherwise
Time-dependent characteristics	Time of day interview conducted: 1 if 5:00–9:59 a.m., 2 if 10:00 a.m.–2:59 p.m., 3 if 3:00–7:59 p.m., 4 if 8:00 p.m.–4:59 a.m.
	Wind speed (mph)
	Wind direction (degrees)
	Wind gust (mph)
	Air temperature (degrees in Fahrenheit)
	Saturation vapor pressure (mb)
Segment characteristics	Functional class of segment: 1 if local/collector, 2 if minor arterial, 3 if major arterial
	Number of lanes on segment (2-way)
	Segment length (mile)
	Curb lane width (ft)
	Effective sidewalk width (ft)
	Bicycle compatibility index (BCI)
	Direction traveled on segment: 1 if north, 2 if east, 3 if south, 4 if west
	One-way against traffic indicator:1 if use one-way segment against traffic direction, 0 otherwise
	Slope indicator variable: 1 if steep slope, 2 steep downhill, 0 otherwise
	Curb indicator variable: 1 if curb present, 0 otherwise
	Parked vehicles indicator: 1 if on-street parking allowed, 0 otherwise
	Pedestrian traffic indicator: 1 if typically congested sidewalk, 0 otherwise
	Average daily traffic (vph)
	Pavement condition (PCR—in/mi)
	Encounter traffic signal at start of one segment indicator: 1 if encounter traffic signal at start point, 0 otherwise
	Encounter traffic signal at end of one segment indicator: 1 if encounter traffic signal at end point, 0 otherwise
	Bicycle lane indicator: 1 if bicycle lane presence, 0 otherwise
	85th percentile traffic speed (mph)
	Residential development indicator: 1 if residential presence, 0 otherwise
	Segment level of service (A-1, B-2, C-3, D-4, E-5, F-6)
	Bicycle path indicator: 1 if bicycle path presence, 0 otherwise

Road pavement surface conditions of West Lafayette were evaluated by applying the Pavement Surface Evaluation and Rating (PASER) system (Walker 2002), which rates pavement surface condition from 1 (failed) to 10 (excellent) in terms of surface distress.

Bicyclist and network characteristics are summarized in Table 2. Among the bicyclists interviewed, 45 were female and 133 were male, with ages ranging from 18 to 30. All of them are commuters who are familiar with conditions on their routes to campus.

Table 2 Sample statistics summary

Variable	Mean	Median	SD	Minimum	Maximum
Bicyclist age	21.3	21.0	1.90	18	30
Segment length (mile)	0.11	0.07	0.08	0.01	0.68
Effective sidewalk width (feet)	6	5	2.23	0	12
Average daily traffic	8,585	5,800	6,799	327	26,962
Curb lane width (feet)	11.6	11.0	1.30	7	20
85th percentile speed (mph)	32.7	35.0	6.29	20	40
Wind speed (mph)	13.4	14.7	5.33	3.1	19.6
Wind gust (mph)	22.9	24.8	6.85	11.4	31.0
Pavement condition rating	6.00	6.00	0.93	4.00	8.00
Air temperature (Fahrenheit)	63.8	59.0	11.078	49.0	85.0
Saturation vapor pressure (mb)	21.6	16.8	9.203	11.9	41.5
Bicycle compatibility index	3.72	3.34	1.06	0.92	5.45

Methodology

Because cross-section choice behavior is an individual choice, a logit model is suitable for our analysis. In the standard binary logit model, only the utility difference of two choices matters. Therefore, the utility difference could be captured by a constant term, covariates that determine the outcome for observations, and a Gumbel distributed disturbance term:

$$y_{in}^{*} = \alpha + \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}_{in} + \varepsilon_{in} \tag{1}$$

$$Y_{in} = \begin{cases} 1, & y_{in}^{*} > 0 \\ 0, & y_{in}^{*} \leq 0 \end{cases} \tag{2}$$

The subscript i denotes an individual bicyclist, n denotes the number of choices she/he had made along the chosen route, and $y*$ is a latent variable that represents the utility difference between two alternatives. If $y*$ is greater than zero, an on-street facility will be selected. Otherwise, an off-street facility will be selected. $\boldsymbol{X}$ is the vector of variables determining the discrete choice for observation n made by bicyclist i, $\boldsymbol{\beta}$ is the vector of estimated parameters, and ε is the random disturbance. The probability $P_n(Y_{in} = 1|x_{in}, \boldsymbol{\beta})$ that an on-street facility is chosen in the nth observation is represented in Eq. 3 as (Ben-Akiva and Lerman 1985):

$$P_n(Y_{in} = 1|x_{in}, \boldsymbol{\beta}) = P\left(y_{in}^{*} > 0\right) \tag{3}$$

Because, in our data base, each respondent generates multiple observations, these observations will likely share unobserved effects (unobserved characteristics relating to each respondent). These shared unobserved effects violate the assumption of independent disturbances made in Eq. 1. If this is not accounted for, disturbance correlation will violate the disturbance independence assumption and result in potentially erroneous parameter estimates (Washington et al. 2011). A random effects approach could be introduced to solve this problem by adding individual random effect terms φ_i (assumed to be normally distributed with mean = 0 and variance = σ^2).

$$y_{in}^{*} = \alpha + \boldsymbol{\beta}_i^{\mathrm{T}} x_{in} + \varepsilon_{in} + \phi_i \tag{4}$$

Before we admit the random effects of unbalanced panel data generated by the bicyclists, we need to realize that there are also many other sources of unobserved heterogeneity. It is likely that different people behave differently given the same conditions, such as two people making different choices on the same segment. The same road may have several segments that may share unobserved effects. People may behave differently during a longer trip than during a shorter trip, etc. Unobserved heterogeneity makes invalid the assumption that the effect of any individual explanatory variable is the same. This problem may be addressed by a more generalized random parameters approach, which assumes that the estimated parameters vary across the population. It allows individuals to have heterogeneous responses with respect to changes in a dependent variable. To some extent, a random effects model is a special case of a random parameters model, by allowing the "constant" term to be random for each bicyclist. Based on the groups established by the random effects model, we estimate parameters for each bicyclist and the probabilities of on-street facility being chosen with a mixed model $P_n^r(y_i = 1|x_{in})$ are defined as

$$P_n^r(y_i = 1|x_{in}) = \int_x P_n(y_i = 1|x_{in}) f(\boldsymbol{\beta}\phi) d\boldsymbol{\beta} \tag{5}$$

where $f(\boldsymbol{\beta}|\phi)$ is the density function of $\boldsymbol{\beta}$ with φ referring to a vector of parameters of that density function (mean and variance). The $\boldsymbol{\beta}$ new contains a common mean plus a randomly distributed term (e.g., a normally distributed term with mean zero and variance σ^2), representing a standard deviation of the mean for each bicyclist. $P_n^r(y_i = 1|x_{in})$ were approximated by drawing values of $\boldsymbol{\beta}$ from $f(\boldsymbol{\beta}|\phi)$, given values of φ, and using these drawn values to estimate the random effects logit probability $P_n(Y_{in} = 1|x_{in}, \boldsymbol{\beta})$ (Washington et al. 2011). To estimate this model, a log-likelihood transformation is conducted to produce a *LL* function

$$LL = \sum_{n=1}^{N} \left[\delta_{in} LNP_n^r(y_i = 1|x_{in})\right] \tag{6}$$

where δ_{in} is 1 if observation n is the choice of an on-street facility, and 0 otherwise. Estimation of Eq. 6 would be undertaken by a simulation-based maximum likelihood method with a random draws or Halton draws (Halton 1960) procedure. More discussion about this estimation technique can be seen in Bhat (2001, 2003) and Train (1999).

In addition to parameter estimation, we calculated marginal effects for indicator variables and elasticity for continuous variables. There are two ways to calculate marginal effects: marginal effects at the means (MEMS) and average marginal effects (AME). (Bartus 2005) To calculate MEMS for an indicator variable, one needs to hold all the other variables at their means, or simply take the derivative with respect to the binary variable as if it were continuous, which provides an approximation that is often surprisingly accurate (Greene 2002). This method works well for linear models; however, in the logit model, AME is superior due to the nonlinearity of the outcome probabilities with respect to each variable and because MEMS suffers aggregation bias (Washington et al. 2011). Instead of computing only one marginal effect while setting other variables at their means, AME computes the marginal effect for each observation while other variables stay fixed and then takes the average. For an indicator variable x_k, its AME would be computed as:

$$\frac{\Sigma_N[P_n(Y_n|x_k = 1) - P_n(Y_n|x_k = 0)]}{N} \tag{7}$$

where N is the number of observations. Marginal effects for binary indicator variables are interpreted as a change in the outcome probability given this variable changing from zero

to one. For a continuous variable, marginal effects are the outcome probability change given one unit change in that variable. However, sometimes it is difficult to quantify "one unit" of a continuous variable. Rather than applying an impractical concept, instead, elasticity is adopted. Elasticity values are interpreted as the percent effect that a 1 % change in the targeted variable has on the outcome probability, which is shown as

$$E_{x_k}^{P(n)} = \frac{\partial P(n)}{\partial x_k} \times \frac{x_k}{P(n)} \tag{8}$$

If the computed elasticity value is greater than one, then this variable is said to be elastic, meaning that a 1 % change in this variable will bring more than a 1 % change in the probability of an on-street facility being chosen. If the computed elasticity value is less than one, then this variable is said to be inelastic. A 1 % change in this variable would bring less than 1 % probability change in the outcome.

Model estimation

Best model specification

Halton draws and random draws were both tested as part of a simulation-based maximum likelihood method, and 8,500 Halton draws were found to produce stable results. Random draws would produce results similar to Halton draws, but with a much larger number of simulations. This result is consistent with Bhat's (2003) suggestion that using Halton draws is more efficient than random draws. Therefore, only the Halton draws' results are presented. Likelihood ratio tests were conducted to determine whether the traditional random effects or more general random parameters model provided the best statistical fit. It was found that the mixed logit model provided a statistically superior fit in that the null hypothesis that the traditional random effects and the mixed models were statistically the same could be rejected with over 99.99 % confidence.

To save space, only the results of the mixed logit model with parameters varying across respondents using Halton draws are presented in Table 3. In the mixed logit model, the *t*-stat for a parameter's mean is not of interest. Instead, we care more about the standard deviation of the estimated parameters for each variable. If the standard deviation is significant, then the random parameter approach is warranted to help capture the unobserved heterogeneity.

Turning to the estimation results in Table 3, the parameters that were found to be random (varying across respondents) were: the constant term, bicycle lane indicator; major arterial indicator; minor arterial indicator; BCI; effective sidewalk width; high average daily traffic indicator, and good pavement condition indicator. All of these parameters were found to be normally distributed with their standard deviations significantly different from zero (log-normal, uniform, Weibull, exponential, and triangle distribution were also tried but provided inferior statistical fits relative to the normal distribution).

Interpretation of estimation results

The mixed logit model estimates provide information about how user characteristics and segment characteristics are associated with previously-defined cross-section choices and help capture unobserved heterogeneity. A positive parameter sign in Table 3 indicates that the corresponding factor makes a bicyclist more likely to use an on-street facility. Besides

Table 3 Estimation results (all random parameters are normally distributed)

Variables	Parameter	*t*-Stat
Random parameters		
Constant (*standard deviation of parameter distribution*)	5.992 (1.754)	5.95 (6.54)
Bicycle lane indicator (1 if bicycle lane present, 0 otherwise) (*standard deviation of parameter distribution*)	1.392 (5.092)	2.90 (6.14)
Major arterial indicator (1 if major arterial, 0 otherwise) (*Standard deviation of parameter distribution*)	−2.717 (6.270)	−4.09 (7.84)
Minor Arterial Indicator (1 if minor arterial, 0 otherwise) (*Standard deviation of parameter distribution*)	0.609 (8.118)	1.63 (8.53)
Bicycle compatibility index (*standard deviation of parameter distribution*)	−0.325 (1.031)	−2.14 (8.70)
Effective sidewalk width (feet) (*standard deviation of parameter distribution*)	−0.300 (0.923)	−3.14 (8.87)
High average daily traffic indicator (1 if ADT > 8,600, 0-otherwise) (*standard deviation of parameter distribution*)	−1.832 (6.881)	−2.94 (7.80)
Good pavement condition indicator (1 if rate > 6, 0-otherwise) (*standard deviation of parameter distribution*)	0.893 (5.224)	2.12 (6.88)
Fixed parameters		
Bicycle path indicator (1 if bicycle lane present, 0 otherwise)	−10.101	−3.95
Signalized intersection at beginning of segment indicator(1-if there is a signalized intersection at beginning of one segment, 0 otherwise)	−1.100	−2.03
Segment length (mile)	−5.337	−1.89
Signalized intersection at end of segment indicator(1-if there is a signalized intersection at the end of one segment, 0 otherwise)	−1.274	−3.17
One-way against traffic indicator(1 if use one-way segment against traffic direction, 0 otherwise)	−2.075	−3.27
Log likelihood at convergence	−295.55	
Log likelihood at zero	−637.69	
Number of observations	931	
McFadden pseudo ρ^2	0.54	

Note There is no multicollinearity problem in these models diagnosed by variance inflation factor test and all VIFs are smaller than 5. 5,500–9,500 Halton draws were tested and results became stable since 8,500 Halton draws. Among the several choices of a bicyclist, the previous cross-section choice would influence the next choice. However, this kind of state dependence variable may also capture residual heterogeneity, which would lead to one to observe spurious state. Therefore, this variable was not included in the model. More discussion about this could be found in Washington et al. (2011) and Heckman (1981)

the randomly distributed variable stated above, the parameters for the "signalized intersection at the beginning of segment" indicator variable are fixed across the population, because the estimated standard deviation of this parameter distribution was not significantly different from zero at the 95 % confidence level. Marginal effects for indicator variables and the elasticity for continuous variables are displayed in Table 4. The percentages of the random parameter distributions that are above and below zero, given the estimated mean and standard deviation of the random parameters, are displayed in Table 5.

A constant term is estimated for each bicyclist, to pick up the mean unobserved effects in the error term. Without it, the error term may not have a mean of zero, which violates the regression model assumption. The constant term in Table 3 is assumed to be normally

distributed with a mean of 5.992 and a standard deviation of 1.754. Its sign is positive for all the bicyclists. This result suggests, in our case study environment, that nearly all bicyclists have an initial preference to ride on the street. This is because, under ideal conditions, such as not being threatened by high-volume or fast-moving vehicle traffic and having good curb lane pavement condition, bicyclists riding on the street can comfortably maintain higher speeds. However, as the bicycling environment departs from the ideal, a bicyclist may adjust her/his facility type choice correspondingly.

The bicycle lane indicator produces random parameters with a mean of 1.392 and a standard deviation of 5.092. Its positive mean value indicates that the introduction of a bicycle lane would encourage bicyclists to use it. The bicycle path variable has a fixed negative parameter, which suggests that adding a bicycle path would decrease, on average, by 0.29 the probability that an on-street facility would be selected. These two results are intuitive and in line with other researchers' results (Krizek et al. 2009; Sener et al. 2009), because a bicycle lane or a bicycle path would provide bicyclists with a specified right-of-way. Based on a link-level analysis of bicyclist behavior, our results suggest that adding a bicycle lane would be viewed favorably by 61 % of bicyclists, increasing by 0.30 the probability that an on-street facility is chosen. The bicycle lane and bicycle path variables are demonstrated to have similar strength in terms of their AME—+0.30 and −0.29, respectively. However, when we look at individual choices (to be discussed in Sect. 5.3), the bicycle path is found to be more attractive. Although building bicycle lanes has been found to increase bicycle ridership (Krizek et al. 2009), being able to estimate the extent to which bicyclists would use a new bicycle lane or bicycle path can guide bicycle infrastructure investment at specific locations. Adding a bicycle lane without considering other factors like roadway characteristics and safety concerns may not be a good strategy. The models developed in this study permit the incorporation of multiple factors.

The estimated parameter for major arterial is assumed to be normally distributed with a mean of −2.717 and a standard deviation of 6.270. This suggests that, for 67 % of

Table 4 Average marginal effects and elasticity

Variable	Elasticity/marginal effects
Effective sidewalk width (feet)	-1.93^{e}
Segment length (mile)	-0.26^{e}
Bicycle compatibility index	-1.25^{e}
Major arterial indicator (1 if major arterial, 0 otherwise)	-0.11^{m}
Minor arterial indicator (1 if minor arterial, 0 otherwise)	-0.05^{m}
High average daily traffic indicator	-0.07^{m}
Bicycle lane indicator (1 if bicycle lane present, 0 otherwise)	0.30^{m}
Bicycle path indicator (1 if bicycle path present, 0 otherwise)	-0.29^{m}
One-way against traffic indicator (1 if use one-way segment against traffic direction, 0 otherwise)	-0.05^{m}
Signalized intersection at end of segment indicator(1-if there is a signalized intersection at the end of one segment, 0 otherwise)	-0.03^{m}
Good Pavement Condition Indicator (1 if rate > 6, 0-otherwise)	0.03^{m}
Signalized intersection at beginning of segment indicator (1-if there is a signalized intersection at beginning of one segment, 0 otherwise)	-0.03^{m}

Note superscript *e* denotes elasticity for that variable. Superscript *m* denotes calculated marginal effects

Table 5 Random parameters percent of distribution

Variable	Random parameters Above zero (%)	Percent of distribution Below zero (%)
Bicycle lane indicator (1 if bicycle lane present, 0 otherwise)	61	39
Major arterial indicator (1 if major arterial, 0 otherwise)	33	67
Minor arterial indicator (1 if minor arterial, 0 otherwise)	53	47
Bicycle compatibility index	38	62
Effective sidewalk width (feet)	37	63
High average daily traffic indicator (1 if ADT > 8,600, 0-otherwise)	39	61
Good pavement condition indicator (1 if rate > 6, 0-otherwise)	57	43

bicyclists are more likely to choose the off-street facility on the major arterial. Similarly, the estimated parameter for minor arterial is tested to produce a normal distributed parameter with a mean of 0.609 and a standard deviation of 8.118, which suggests 53 % of bicyclists are more likely to use on-street facility on the minor arterial. Riding on a major arterial would result in the probability that an on-street facility is chosen decreases by 0.11. Meanwhile, riding on a minor arterial would decrease the probability that an on-street facility is chosen by only 0.05. Higher class roads have higher vehicle speeds and higher traffic volumes, making a bicyclist feel less safe. This outcome suggests that an on-street bicycle facility investment along a higher functional class segment may not be suitable; rather, a bike path should be considered, if space permits.

The "BCI" variable yields a random parameter with a mean of −0.325 and a standard deviation of 1.031. This outcome suggests that 62 % of bicyclists are more likely to use an off-street facility when they feel more dangerous riding on the street. It is intuitive that most bicyclists avoid vehicle threats by switching to an off-street facility. However, it should be noted that 38 % of bicyclists are less sensitive to a dangerous riding environment. A 1 % increase in this variable would cause a 1.25 % decrease in the probability of on-street facility usage. This value suggests BCI is elastic and that somehow improving bicyclists' perceived on-street safety level would keep many of them off the sidewalk.

The effective sidewalk width is tested to be significant at 95 % confidence level and produces random parameters with a mean of −0.300 and a standard deviation of 0.923. This indicates that 63 % of bicyclists would be attracted by a wider sidewalk. The elasticity value shows that this variable is elastic; a 1 % increase in this variable value would increase by 1.93 % the probability that an off-street facility is chosen. However, wider sidewalks tend to exist where there are more pedestrians to serve. If the objective of the city planner is to keep bicyclists off the sidewalk, an attractive alternative for bicyclists should be provided. A bicycle path could be provided to avoid pedestrian and bicyclist conflicts if space permits.

A threshold value of 8,600 vpd for the "average daily traffic" indicator variable was chosen after several trials. A street with ADT greater than 8,600 produces a normally distributed parameter with a mean of −1.832 and a standard deviation of 6.881. This suggests that, for 61 % of bicyclists, high ADT would have negative impact on them, which decreases by 0.07 the probability that an on-street facility is used. The threshold may vary between communities.

A bad "curb lane pavement condition" negatively influences a bicyclist's decision to ride on the street. According to the PASER system (Walker 2002), if the pavement condition rating is greater than 6, only minor patching or routine maintenance is needed, which is good pavement surface performance. In our case, a good "pavement condition" indicator yields random parameters with a mean of 0.893 and a standard deviation of 5.224. This suggests that 57 % of bicyclists are more likely to use an on-street facility if the pavement condition is good or better. Improving pavement condition on the street can increase the likelihood of bicyclists leaving the sidewalk by 0.03.

The "one-way against traffic" indicator variable has a negative parameter. This suggests that bicyclists would be more likely to use an off-street facility when they are using a one way street and riding against vehicle traffic. The probability that an off-street facility will be chosen by a bicyclist is 0.05 higher when traveling against the direction of vehicles. This result suggests that "contra-flow" bicycle movements are best accommodated with an off-street facility such as a bicycle path, if space is available. Providing an official contra-flow bicycle lane against a one-way street's vehicle direction requires specific signs and markings (NACTO 2012), which are not present in our case, for example, along one-way northbound University Street.

The "segment length" variable yields a negative parameter, indicating that bicyclists are more likely to choose an off-street facility on a longer segment. This may be because a longer segment length is associated with higher exposure to vehicle traffic. The segment length variable is inelastic; a 1 % increase in segment length leads to only a 0.26 % decrease in the probability of an on-street facility being selected.

The "signalized intersection at end of segment" indicator variable produces a fixed negative parameter that indicates that bicyclists would prefer to use an off-street facility (with 0.03 lower probability) when they will confront a signalized intersection at the end of the segment. A traffic signal tends to be located at a busy intersection. The most common manifestation of the model results is the use of crosswalks and pedestrian signals by many bicyclists.

The "signalized intersection at beginning of segment" indicator variable yields a fixed negative parameter, indicating that facing a signalized intersection would decrease by 0.03 the likelihood that an on-street facility is chosen. Because traffic signals tend to be located at busy intersections, bicyclists tend to use crosswalks and pedestrian signals there, and the sidewalks they lead to.

Application of model results

From a city engineer's perspective, we might want to measure the impact of adding a new bicycle path or a bicycle lane along a segment, such that one can predict the probability of each bicycle facility within a cross-section will be chosen. Therefore, prediction analysis based on the model specification would provide more insights about the relative strength of adding bicycle path/lane and give guidance on making bicycle facility improvements. To illustrate the potential usefulness of the model, the impact of different facility improvement scenarios and strategies are considered below.

We examine the predicted probabilities with respect to "bicycle lane" and "bicycle path", while other variables stay constant. The probability of an on-street facility being selected is predicted for each observation. To avoid losing any information for each observation, the summation of predicted probabilities for all observations is used as the predicted number of using on-street facility. In the sample database, the on-street facility was chosen 525 times and the off-street facility was chosen 406 times. According to the mixed logit model prediction, on-street facility is estimated to be selected 559 times. Compared to

Table 6 Model prediction with respect to different scenarios

	Model prediction	All bicycle lane	No bicycle lane
On-street facility	559	607	327
Off-street facility	372	324	604

	All bicycle path	No bicycle path	All bicycle lane and bicycle path
On-street facility	291	561	352
Off-street facility	640	370	579

the observed cases, the model overestimates somewhat the use of on-street facilities. However, the prediction error is within 5 % [(559 − 525)/931 = 3.6 %]. Under five different scenarios, the corresponding indicator variable is set to be one or zero to represent different scenarios while other variables stay unchanged. The results are shown in Table 6.

If a bicycle lane were added to every segment, while holding other factors unchanged, in 607 − 559 = 48 cases would bicyclists switch to the bicycle lane. If all bicycle lanes were removed, on-street facilities would lose 232 users, which is 24.9 % of the 931 total observations. Similarly, if all segments are equipped with a bicycle path, while holding other variables constant, in 268 cases bicyclists would switch to off-street facilities. In the scenario which all the segments have both a bicycle lane and a bicycle path, while holding other factors constant, in 207 cases bicyclists would switch to off-street facilities. If we evaluate only the AME of two variables, we draw the conclusion that these two facilities are equally attractive. However, when we look at individuals' predicted response to both bicycle facilities, we would learn that a bicycle path is more attractive than a bicycle lane. This result is in line with Stinson and Bhat (2005). The seemingly different result drawn from AME and prediction analysis can be explained by the large standard deviation of the parameter estimates. Different bicyclists perceive safety differently and have a large variance in facility preferences. This result also reveals the preferences of each bicyclist and demonstrates the strength of using a random parameters model, which can help capture unobserved heterogeneity. Furthermore, this aggregate prediction procedure puts into better perspective the potential effect of investments in certain bicycle facilities.

Conclusion

This paper looks at bicyclists' choice of bicycle facility types within the street right-of-way cross-section, with the intent of developing a way to guide public investment decisions. The methodology adopted in this paper has produced useful findings. The mixed logit model established in this paper can capture bicyclist preferences and unobserved heterogeneity. In general, bicycle facility planners would prefer to keep bicyclists off the sidewalk, perhaps by providing more bicycle lanes. Some of our study's findings are expected:

- Under ideal conditions, bicyclists prefer riding on the street.
- A bicycle path is more attractive than a bicycle lane.
- Bicyclists tend to use off-street facilities along higher functional class roads or streets with high ADT.
- If a bicyclist's feeling of safety (as measured by the BCI) is improved, the use of on-street facilities is increased.

- Poor road surface conditions drive bicyclists to off-street facilities.

Other findings were less obvious:

- In many situations, bicyclists use sidewalks, despite potential conflicts with pedestrians.
- Adding a bicycle lane may not be the best use of limited bicycle facility funds.

The mixed logit model estimates provide information about how user characteristics and segment characteristics are associated with previously defined cross-section choices along a bicyclist's route and help capture unobserved heterogeneity (indicated by the large standard deviation of the parameter estimates). In addition to parameter estimation, we calculated average marginal effects for indicator variables and elasticity for continuous variables. The results from such models, of course, may vary as the database varies, but the mixed logit model allows us to make full use of a database that includes multiple observations for each bicyclist during his/her trip. Considering the large standard deviation of the parameter estimates, a prediction analysis is conducted to test different scenarios with respect to bicycle facility improvement. A bicycle path was found to be more attractive than a bicycle lane. Using the model results to test potential scenarios can help guide investments in bicycle facilities.

Acknowledgments The authors thank Professor Fred Mannering at Purdue University for his guidance and suggestions on model estimation. The authors appreciate the helpful comments of three anonymous reviewers on an earlier version of the paper. The authors acknowledge the efforts of Jennifer Brown and Kevin Ford in preparing the original data set. Yingge Xiong also provided some thoughtful ideas for this paper.

References

Antonakos, C.L.: Research issues on bicycling, pedestrians, and older drivers. Environmental and travel preference of cyclists. Transp. Res. Rec. **1438**, 29–31 (1994)

Aultman-Hall, L., Adams, J.R.M.F.: Sidewalk bicycling safety issues. Transp. Res. Rec. **1636**, 71–76 (1998)

Bartus, T.: Estimation of marginal effects using margeff. Stata J **3**, 309–329 (2005)

Ben-Akiva, M., Lerman, S.R.: Discrete choice analysis: theory and application to predict travel demand, pp. 69–70. MIT Press, Cambridge (1985)

Bhat, C.: Quasi-random maximum simulated likelihood estimation of the mixed multinomial Mogit model. Transp. Res. B **37**, 677–693 (2001)

Bhat, C.: Simulation estimation of mixed discrete choice models using randomized and scrambled Halton sequences. Transp. Res. B **37**, 837–855 (2003)

Buehler, R., Pucher, J.: Cycling to work in 90 large American cities: new evidence on the role of bicycle paths and lanes. Transportation **39**, 409–432 (2012)

de Geus, B., De Smet, S., Nijs, J., Meeusen, R.: Determining the intensity and energy expenditure during commuter cycling. Br. J. Sports Med. **41**, 8–12 (2007)

Dill, J., Carr, T.: Bicycle commuting and facilities in major U.S. cities: if you build them, commuters will use them. Transp. Res. Rec. **1828**, 116–123 (2003)

Federal Highway Administration: National bicycling and walking study, case study no. 1: reasons why bicycling and walking are not being used more extensively as travel modes. U.S. Department of Transportation, Washington (1992)

Ford, K.M., Fricker, J.D., Brown, J.M.: Discrete choice analysis of bicyclist cross-sectional behavior. 90th Annual Meeting of the Transportation Research Board, Paper #11-0956, 23–26 Jan 2011

Greene, W.H.: Econometric analysis, 5th edn. Prentice-Hall, Upper Saddle River (2002)

Halton, J.: On the efficiency of evaluating certain quasi-random sequences of points in evaluating multi-dimensional integrals. Numer. Math. **2**(1), 84–90 (1960)

Harkey, D.L., Reinfurt, D.W., Knuiman, M., Stewart, J.R., Sorton, A.: Development of the bicycle compatibility index: a level of service approach. Federal Highway Administration, Report FHWA-RD-98-072. Chapel Hill (1998)

Harvey, F., Krizek, K.J., Collins, R.: Using GPS data to assess bicycle commuter route choice. Compendium of papers DVD, 87th Annual Meeting of the Transportation Research Board, Paper #08-2951, 13–15 Jan 2008

Heckman, J.: Statistical models for discrete panel data. In: Manski, C., McFadden, D. (eds.) Structural analysis of discrete data with econometric applications. MIT Press, Cambridge (1981)

Heinen, E., Van Wee, B., Maat, K.: Commuting by bicycle: an overview of the literature. Transp. Rev. **30**(1), 59–96 (2010)

Highway capacity manual: Special Report 209. National Research Council, Washington (1994)

Highway capacity manual: Transportation Research Board. National Research Council, Washington (2010)

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

Hunter, W.W., Stewart, J.R., Stutts, J.C.: Study of bicycle lanes versus wide curb. Transp. Res. Rec. **1674**, 70–77 (1999)

Klobucar, M., Fricker, J.D.: Network evaluation tool to improve real and perceived bicycle safety. Transp. Res. Rec. **2031**, 25–33 (2007)

Krizek, K.J., Barnes, G., Thompson, K.: Analyzing the effects of bicycle facilities on commute mode share over time. J. Urban Plan. Dev. **135**(2), 66–73 (2009)

Krizek, K.J., Johnson, P., Tilahun, N.: Gender differences in bicycling behavior and facility preferences. In: Rosenbloom, S. (ed.) Research on women's issues in transportation, special report, vol. 2, pp. 31–40. National Academy Press/Transportation Research Board, Washington (2005)

Lawlor, D.A., Ness, A.R., Cope, A.M., Insall, P., Riddoch, C.: The challenges of evaluating environmental interventions to increase population levels of physical activity: the case of the UK National Cycle Network. J. Epidemiol. Community Health **57**(2), 96–101 (2003)

League of Illinois Bicyclists.: Polls Show Americans Want Better Biking. http://www.bicyclelib.org/tbills/americanswant.htm (2003). Accessed 10 July 2009

Meuleners, L.B., Lee, A.H., Haworth, C.: Road environment, crash type and hospitalization of bicyclists and motorcyclists presented to emergency departments in western Australia. Accid. Anal. Prev. **39**, 1222–1225 (2007)

NACTO: NACTO urban bikeway design guide, contra-flow bike lanes, National Association of City Transportation Officials, http://nacto.org/cities-for-cycling/design-guide/bike-lanes/contra-flow-bike-lane/ (2012). Accessed 15 May 2012

Nelson, A.C., Allen, D.: If you build them, commuters will use them: association between bicycle facilities and bicycle commuting. Transp. Res. Rec. **1578**, 79–83 (1997)

Parkin, J., Wardman, M., Page, M.: Estimation of the determinants of bicycle mode share for the journey to work using census data. Transportation **35**, 93–109 (2008)

Pucher, J., Dill, J., Handy, S.: Infrastructure, programs, and policies to increase bicycling: an international review. Prev. Med. **50**, 106–125 (2010)

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

Stinson, M.A., Bhat, C.R.: A comparison of the route preferences of experienced and inexperienced bicycle commuters. Presented at 84th annual meeting of the Transportation Research Board, Washington, 9–13 Jan 2005

Taylor, D., Mahmassani, H.: Analysis of stated preferences for intermodal bicycle-transit interfaces. Transp. Res. Rec. **1556**, 86–95 (1996)

Tilahun, N.Y., Levinson, D.M., Krizek, K.: Trails, lanes, or traffic: valuing bicycle facilities with an adaptive stated preference survey. Transp. Res. A **41**, 287–301 (2007)

Train. K.: Halton Sequences for Mixed Logit. Working Paper, Department of Economics, University of California, Berkeley (1999)

Wachtel, A., Lewiston, D.: Risk factors for bicycle-motor vehicle collisions at intersections. ITE Journal **64**, 30–35 (1994)

Walker, D.: PASER manuals—concrete and asphalt. Wisconsin Transportation Information Center, 432 North Lake Street, Madison, WI 53706, USA. tic.engr.wisc.edu (2002). Accessed 16 April 2012

Washington, S., Karlaftis, M., Mannering, F.L.: Statistical and econometric methods for transportation data analysis, 2nd edn. Chapman & Hall/CRC, Boca Raton (2011)

Author Biographies

Lei Kang is a graduate of the Masters program in Transportation and Infrastructure Systems Engineering, School of Civil Engineering, Purdue University. He also earned a Graduate Certificate in Applied Statistics from the Department of Statistics at Purdue. Mr. Kang's Bachelors degree is in Transportation Engineering

from Tongji University in Shanghai, China. His research interests include statistical and econometric modeling of bicyclist and pedestrian behavior and safety.

Jon D. Fricker is Professor of Civil Engineering at Purdue University, where he regularly teaches courses on Transportation Engineering, Transportation Planning, Urban Planning and Analysis, and Public Mass Transportation. His current research interests include asset management, traffic impact analysis, and bicyclist response to bicycle facility alternatives.