Bicycle route choice model developed using revealed preference GPS data

Joseph Broach
PhD Student
Nohad A. Toulan School of Urban Studies and Planning
Portland State University
PO Box 751
Portland, OR 97207-0751
jbroach@pdx.edu
phone: 971-285-7045
fax: 503-725-8480
corresponding author

John Gliebe
Assistant Professor
Nohad A. Toulan School of Urban Studies and Planning
Portland State University
PO Box 751
Portland, OR 97207-0751
gliebej@pdx.edu
phone: 503-725-4016
fax: 503-725-8770

Jennifer Dill
Associate Professor
Nohad A. Toulan School of Urban Studies and Planning
Portland State University
PO Box 751
Portland, OR 97207-0751
jdill@pdx.edu
phone: 503-725-5173
fax: 503-725-8770

Paper submitted for presentation and publication at the 90th Annual Meeting of the Transportation Research Board January 23-27, 2011, Washington D.C.

Original submission date: August 1, 2010

Word count: 5464 (text) + 3 x 250 (tables) = 6,214

ABSTRACT

To better understand bicyclists' preferences, we used bicycle-mounted GPS units to observe the behavior of 162 bicyclists for several days each. Trip purpose and several other trip-level variables were recorded by the cyclists, and the resulting trips were coded to a highly detailed bicycle network. We then used the 1,449 valid non-exercise trips to estimate a bicycle route choice model. As part of this research, we developed a choice set generation algorithm based on multiple permutations of labeled path attributes, which seemed to out-perform comparable implementations of other route choice set generation algorithms. The choice model was formulated as a Path-Size Logit model to account for overlapping route alternatives. Estimation results are intuitive and suggest that cyclists are sensitive to the effects of distance, turn frequency, slope, intersection control, and traffic volumes. In addition, cyclists appear to place relatively high value on off-street bike paths, enhanced neighborhood bikeways (bicycle boulevards), and bridge facilities. Finally, estimation results support segmentation by commute versus non-commute trip types. The route choice model presented in this paper is currently being implemented as part of the regional travel forecasting system for Portland, Oregon, U.S.A.

INTRODUCTION

Non-motorized travel options have been largely ignored in regional transportation planning studies in the U.S., where decisions on more resource-intensive investments in highway and transit facilities have been of primary concern. Recently, however, policy-maker interest in sustainable transportation systems and healthier lifestyles has shifted some of the decision-making focus to bicycling and walking and the extent to which the urban travel environment supports these alternative modes.

To our knowledge, the practice in operational travel forecasting models used in North America has been to assume that cyclists choose the minimum-distance path between origins and destinations using a fixed travel speed, without consideration of network attributes. Travel environment attributes, such as slope, traffic volumes, and the presence of on and off-street bikeways are not considered. Moreover, extant models do not differentiate between bicycle trip purposes.

From March through November, 2007, we collected detailed survey data revealing the actual paths taken by 164 bicyclists over the course of several days, using global positioning system (GPS) tracking devices. The data were mapped to transportation network facilities, creating an enhanced bicycle network Geographic Information System (GIS) map file showing facility types, bike lanes and off-road trails. Each participant also provided trip purpose and weather conditions.

In this paper, we present a bicyclist route choice model developed from the data gathered in 2007. The model is currently being implemented in the regional travel forecasting framework for Oregon Metro, the metropolitan planning organization for the Portland region and a regionally-elected governing body. Metro Council is keenly interested in the capability of the modeling tool to project use of bicycle infrastructure investment alternatives and to derive economic welfare measures from such analysis. To our knowledge this is the first bicycle route choice model to be developed from revealed preference data that were generated through GPS methods.

In the remainder of this paper we review the existing literature on bicycle route choice modeling; describe the person, GPS, and network data used in model development; briefly describe important modeling assumptions regarding choice set generation and overlapping alternatives; and present the model specification and estimation results. Finally, we conclude with an assessment of what we believe to be the important modeling and policy implications of this research and suggest possible avenues for further development of the model.

REVIEW OF EXISTING BICYCLE ROUTE CHOICE MODELING EFFORTS

Most existing work on bicyclist route choice consists of small, targeted studies focusing on only a few variables. Sener et al. (1) provided a recent comprehensive review of published work. The primary data collection strategies have been recalled paths and binomial or trinomial choice stated preference surveys.

Stated Preference Studies

Stated preference studies have dominated the literature due to several appealing characteristics. Detailed travel network data are unnecessary. There is no need to solve the formidable problem of generating alternative routes. Model specification and estimation are also simpler due to the “clean” data and limited number of alternatives. From a policy perspective, the usual advantage of stated preferences for testing rare or nonexistent features also applies.

There are drawbacks to stated preference data for cyclist route choice. The usual technique in existing studies has been to show respondents a sequence of side-by-side comparisons from which a binary choice is made (see, for example 2, 3, 4). Sener and Bhat (1) used this technique with three alternatives. It is difficult to know how well a participant can map these textual, or occasionally pictorial, representations to her preferences for real facilities. Many salient features of a route are sure to be missing on a piece of paper or computer screen. Also, although the choice set is in a sense controlled, it seems likely that respondents have in mind their own usual routes as points of comparison. Strategic bias is a possibility if participants think responses might influence policy outcomes. None of this is to say stated preference studies are not useful and the results valid, only that their advantages in execution involve tradeoffs.

Landis et al. (5) conducted an interesting variation on the typical stated preference method. Participants actually rode predefined alternative routes before evaluating each. There still may be a problem assuming cyclists can evaluate an unknown route in the same way they do a familiar one, but the technique does promise greater realism.

The stated preference projects most comparable with our research are two similar studies of route choice using web-based surveys (6, 1). Cyclists were provided with a base route and one or two alternatives with carefully designed characteristics. Binary Logit (6) and Mixed Multinomial Logit (1) models were estimated using the stated preference data along with personal characteristics of participants. Taking into account specific data and modeling differences, we found the results to generally agree with our own, with some interesting exceptions. More specific comparisons are provided in the model estimation section of this paper.

Revealed Preference Studies

A handful of revealed preference studies have been undertaken on this topic, but in general they are limited studies that do not estimate a full route choice model. Most commonly, cyclists have been asked to recall routes. The routes are then compared with pre-selected routes based on shortest paths or other definitions of optimal paths (7, 8). These studies have the advantage of using actual routes and network data. The ability of cyclists to accurately recall routes is a question, but it may be quite accurate for habitual routes like commute trips. The larger shortcoming of these studies is the limited choice sets and lack of compensatory choice models.

DATA DESCRIPTION

This research relies heavily on GPS data collected from March through November 2007, by 162 bicyclists recruited from throughout the Portland metropolitan area. After several screening steps, 1,449 non-exercise trips were available for the analysis.

This research also relies heavily on accurate geographic information system (GIS) mapping of an urban street network and off-street bike and multi-use paths, as well as related network attributes. The base network was provided by Oregon Metro, the regional planning organization, and we made numerous updates to make the network routable. In particular, ensuring that overpasses, underpasses, and one-way streets were coded correctly proved challenging. A full report describing the GPS data collection methods and the processes used to prepare the data for our research may be found in the report by Dill and Gliebe (9). A summary of the important features of the data for the purposes of route choice modeling is provided below.

Participants and Bicycle Trips

GPS participants were outfitted with small hand-held devices which they clipped onto their bicycles. They were instructed to enter both weather and trip purpose information and to record the beginning and end of a trip, defined by reaching a particular destination. They also completed demographic and attitudinal surveys.

The participants in this study were primarily regular bicyclists, who agreed to participate in the GPS portion of the study following an initial set of telephone interviews. Although regular cyclists are more likely to be male (80 percent according to the phone survey), we were able to recruit a GPS sample composed of 44 percent females. Among all respondents, 89 percent were between the ages of 25 and 64. Compared with the phone survey of bicyclists used to screen and recruit them, the GPS participants were slightly older, were more likely to have a college degree, had higher incomes, and were more likely to have full-time jobs. They were also more likely live in a two-person household, and only 7 percent lived in a household without a car. The phone survey participants had a demographic comparable to the general population.

GPS Survey Records

GPS tracks were matched to network links using ArcGIS and custom scripts written in the Python programming language (10). Especially challenging was eliminating spurious u-turns caused by GPS signal “bounce.” In some cases, links had to be added to the network where “cut-throughs” and other informal or unmapped links were used. Participants viewed the processed paths and noted GPS errors for manual correction. Further details of the GPS data are available in a separate report (9).

While participating in the study, GPS respondents recorded on average 12 total non-exercise bicycle trips at an average rate of 2.5 per day. About 30 percent of trips were commute trips (home to work or work to home). The average trip distance was 2.2 miles (3.5km) for non-commute trips and 3.7 miles (6km) for commute trips. About 80 percent of total miles recorded were bicycled within the Portland city limits with the balance located in the greater Portland region.

Observed paths were on average somewhat longer than the shortest network paths: by 12 percent for non-commute trips and 11 percent for commute trips. A little more than half (53%) of recorded miles were ridden on facilities with bicycle infrastructure, including bike lanes (29%), off-street paths (13%), and bike boulevards (11%). Bike boulevards are improved neighborhood bikeways with special features to reduce auto speeds and volumes while giving bicycles increased priority at intersections. Further descriptive analysis is available in a separate report (9).

Network

The network model developed for this research included about 88,000 undirected links and 66,000 nodes. This network was constructed as much as possible to include all links available for bicycle travel. This included a large number of links not usually found in an auto travel modeling network such as minor residential streets, off-street bike and multiuse paths, alleyways, and some private roads explicitly open to bicycles. The bike network did not include facilities where bicycle use was legally restricted, mainly urban freeways.

The City of Portland provided interpolated average daily traffic volumes for nearly all streets in the study area based on hose count data. Where missing, volumes were estimated based on functional class. Turns were calculated using a combination of street name and geometry. A 10 meter digital elevation model (DEM) was used to measure elevation gain and loss at roughly 10 meter increments along each link. Bicycle facilities, grade separation, intersection control, and one-way restrictions were provided by Oregon Metro. When generating route alternatives, one-way streets were treated as open to bicycling but with additional impedance based on observed speed reduction (70 percent).

MODELING ASSUMPTIONS

Choice Set Generation

Generating the set of alternative routes considered for each trip was the most difficult and time-consuming part of our analysis. The size and density of the Portland bicycle travel network greatly increased the task's complexity. In addition, the lack of existing revealed preference bicycle route choice studies demanded a careful rethinking of existing generation techniques. Common algorithms based on travel time and street hierarchy were not directly applicable, since bicycle travel times are not affected by speed limits, congestion, and functional class in the same ways as auto travel times.

We experimented with three common choice set generation methods: K-shortest paths, simulated shortest paths, and route labeling, none of which proved entirely satisfactory. Based on these experiments and our own hypotheses about bicyclists' choice set generation process, we developed a modified method of route labeling (11). Route alternatives were chosen by maximizing individual criteria, subject to a flexible, calibrated distance constraint. The new method appeared to outperform existing techniques developed for auto route choice on several key criteria. Readers are referred to the cited paper for further description of the Calibrated Labeling Method.

Applied to our bicycle travel network, the calibrated labeling algorithm produced a median of 20 alternative paths for each trip. The number of alternatives varied across choice situations, increasing with both trip distance and network density. The chosen alternative was not always reproduced exactly by the algorithm, and in such cases it was added to the choice set. For a small number of trips (15 out of 1,464), no alternative to the chosen route was found. These "captive" trips were not included in the model estimation.

Overlapping Alternatives

Where alternate paths share common links, they also presumably have correlated error components. This violates the multinomial logit (MNL) model assumption of independently distributed errors across alternatives. From a statistical point of view, an MNL route choice model will tend to assign counter-intuitively high probabilities to routes that share common network links. From a behavioral point of view, we might say that the MNL considers overlapping routes distinct alternatives; whereas, cyclists may consider such routes jointly as minor variants of a single alternative.

There are two options to overcome the overlapping routes problem (12). A correction factor can be applied to partially adjust the utilities for overlap, leading to the Path-Size Logit (PSL) model. Alternatively, more complex model forms may be specified that allow for correlated errors, including the multinomial probit model, mixed logit models, and closed-form members of the generalized extreme value (GEV) class of models.

Due to the very large number of potential alternatives, we chose the PSL approach, retaining the underlying MNL structure. We recognized the need to be able to apply the model for prediction across a very complex, detailed network. This requirement made the specifications of overlapping route calculations and nest memberships needed for the various probit, mixed logit, and GEV models seem somewhat intractable over such a large computational space. In addition, it has been shown that more complex model forms may not yield consistent estimates when only a small proportion of potential alternatives can be sampled (13).

A path size factor was calculated directly from route alternatives and network geometry, avoiding direct calculation of correlations across alternatives. The general form for the j alternatives in choice set C_n is specified as:

$$PS_{in} = \sum_{a \in \Gamma_i} \frac{l_a}{L_i} \frac{1}{\sum_{j \in C_n} \left(\frac{L_i}{L_j} \right)^{\gamma} \delta_{aj}} \qquad (1)$$

where Γ_i are the links in alternative i, l_a is the length of link a, L_i is the length of alternative i, and δ_{aj} equals 1 if j includes link a (12). The parameter γ is a positive scaling term meant to penalize very long routes in a choice set. Fixing or estimating $\gamma>0$ has been shown empirically to improve route choice model fit (14, 15, 16, 17); however, it has recently been shown that $\gamma>0$ can result in questionable utility corrections and illogical path probabilities (12). In addition, our choice set generation method makes it unlikely that improbably long alternative paths will be included in our analysis. For these reasons, the path-size correction factor in equation 1 is used with $\gamma=0$, essentially dropping the long-path correction factor and yielding the basic Path Size Logit (PSL) model (18).

While relatively simple, the PSL model has been shown to perform well relative to more complex model forms such as the cross-nested logit (CNL), although existing comparisons were performed with the generalized PS factor including $\gamma>0$ (15,16,17). While nested logit models should outperform the PSL specification, they are limited in real network applications due to the huge number of parameters that would have to be estimated to exploit their full flexibility (15, 12). Some promising work has been presented recently on using sub-network components as an improvement to the PSL which maintains much of its estimation simplicity (12). This latter method has not yet been applied to a real network problem but merits further research attention.

The remainder of this paper presents results obtained from the following specification of the Path Size Logit probability:

$$\Pr(i \mid C_n) = \frac{\exp^{V_{in} + \ln(PS_{in})}}{\sum_{j \in C_n} \exp^{V_{jn} + \ln(PS_{jn})}} \quad (2)$$

where *PS* is the path size factor from equation (1) with γ=0. Since *PS* will always fall between 0 and 1, ln(*PS*) will be negative, consistent with a utility reduction proportional to the degree of overlap.

Panel Effects

Our estimation dataset includes observations for 154 participants over multiple trips. It is likely that an individual's series of route choices are correlated to some extent. The inclusion of multiple trip purposes and the generally short period of observation probably limit such correlation. Furthermore, an investigation of commute trip sequences, which we might expect to be the most regular, showed noticeable route choice variation across trips. It did not seem as though these cyclists were "locked in" to a fixed route. For simplicity, trips were assumed to be independent and pooled for analysis. An obvious, but non-trivial future extension would be to consider different specifications including individual-specific effects.

MODEL RESULTS

Table 1 describes the variables used in the route choice model. Table 2 presents the full estimation results from our final model specification. Path-size correction factors were calculated using a custom Python script. Choice model estimation was performed using the freely available BIOGEME package (19).

Distance, slope, and turns

As expected, cyclists prefer shorter routes, all else equal. Log distance outperformed other distance specifications, suggesting that relative rather than absolute route deviations are what matter to cyclists. This result has some behavioral appeal. Implied is that a cyclist would be equally likely to go 1 mile out of her way on a 5 mile trip as 0.2 miles out of her way on a 1 mile trip. A fixed distance, say one mile, is perceived as more costly the shorter the trip. On a longer trip, even a one mile increment might not always be discernible. All else equal, a 1 percent increase in distance reduces the probability of

TABLE 1 Variable Descriptions

Variable	Description	Mean	Present in proportion alts (n=29,090)
Bridge w/ bike lane	bridge with on-street bike lane	dummy variable	0.05
Bridge w/ sep. facility	bridge with improved, separated bike facility	dummy variable	0.22
Prop. upslope 2-4%	Proportion of route along links with average upslope (gain/length) of 2-4%	0.10	0.90
Prop. upslope 4-6%	Proportion of route along links with average upslope (gain/length) of 4-6%	0.03	0.70
Prop. upslope 6%+	Proportion of route along links with average upslope (gain/length) of 6%+	0.02	0.68
Distance (mi)	distance of route in miles	4.48	1.00
Path size (0-1, 1=unique)	path size (see section 4 for formula)	0.31	1.00
Left turn, unsig., AADT 10-20k (/mi)	left turn without traffic signal and parallel traffic volume 10,000-20,000 per day	0.11	0.36
Left turn, unsig., AADT 20k+ (/km)	left turn without traffic signal and parallel traffic volume 20,000+ per day	0.08	0.18
Prop. bike boulevard	proportion of route on designated bicycle boulevard (improved neighborhood bikeway with traffic calming, diversion, and enhanced right of way)	0.10	0.53
Prop. bike path	proportion of route on off-street, regional bike path (i.e. not minor park paths, sidewalks, etc.)	0.04	0.41
Prop. AADT 10-20k w/o bike lane	proportion of route on streets with traffic volume 10,000-20,000 per day without a bike lane	0.08	0.73
Prop. AADT 20-30k w/o bike lane	proportion of route on streets with traffic volume 20,000-30,000 per day without a bike lane	0.04	0.46
Prop. AADT 30k+ w/o bike lane	proportion of route on streets with traffic volume 30,000+ per day without a bike lane	0.02	0.26
Traffic signal exc. right turns (/mi)	left turns and straight movements through traffic signals per mile	1.84	0.90
Stop signs (/mi)	turns or straight movements through stop signs per mile	3.12	0.95
Turns (/mi)	left and right turns per mile	3.64	1.00
Unsig. cross AADT 10k+ right turns (/mi)	right turns at unsignalized intersections with cross traffic volume 10,000+ per day	0.16	0.44
Unsig. cross AADT 5-10k exc. right turns (/mi)	left turns and through movements at unsignalized intersections with cross traffic volume 10,000-20,000 per day	0.56	0.72
Unsig. cross AADT 10-20k exc. right turns (/mi)	left turns and through movements at unsignalized intersections with cross traffic volume 10,000-20,000 per day	0.42	0.72
Unsig. cross AADT 20k+ exc. right turns (/mi)	left turns and through movements at unsignalized intersections with cross traffic volume 20,000+ per day	0.16	0.52

TABLE 2 Route Choice Model Estimation Results

Variable	Est. coeff.	t-stat
Ln(distance)	-5.22	*-10.9*
Ln(distance) * commute	-3.76	*-5.14*
Turns (/mi)	-0.371	*-15.4*
Prop. upslope 2-4 %	-2.85	*-4.57*
Prop. upslope 4-6 %	-7.11	*-6.11*
Prop. upslope >= 6 %	-13.0	*-8.57*
Traffic signal exc. right turns (/mi)	-0.186	*-5.73*
Stop sign (/mi)	-0.0483	*-2.10*
Left turn, unsig., AADT 10-20k (/mi)	-0.782	*-4.19*
Left turn, unsig., AADT 20k+ (/mi)	-1.87	*-4.70*
Unsig. cross AADT >= 10k right turn (/mi)	-0.338	*-2.32*
Unsig. cross AADT 5-10k exc. right turn (/mi)	-0.363	*-5.39*
Unsig. cross AADT 10-20k exc. right turn (/mi)	-0.516	*-5.39*
Unsig. cross AADT 20k+ exc. right turn (/mi)	-2.51	*-11.5*
Prop. bike boulevard	1.03	*5.17*
Prop. bike path	1.57	*4.64*
Prop. AADT 10-20k w/o bike lane	-1.05	*-3.02*
Prop. AADT 10-20k w/o bike lane * commute	-1.77	*-2.28*
Prop. AADT 20-30k w/o bike lane	-4.51	*-6.04*
Prop. AADT 20-30k w/o bike lane * commute	-3.37	*-2.24*
Prop. AADT 30k+ w/o bike lane	-10.3	*-4.67*
Prop. AADT 30k+ w/o bike lane * commute	-8.59	*-1.96*
Bridge w/ bike lane	1.81	*-4.71*
Bridge w/ sep. bike facility	3.11	*-4.96*
Ln(path size)	1.81	*20.78*
Number of observations	1,449	
Null log-likelihood	-4058.7	
Final log-likelihood	-3020.0	
Rho-square	0.256	

choosing a route by about 5 percent and 9 percent for non-commute and commute trips, respectively. That cyclists are highly sensitive to distance is consistent with the observed data. Half of all observed trips were less than 10 percent longer than the shortest path, and 95 percent of trips were less than 50 percent longer.

Travel times in our sample were highly correlated with distance (r = 0.93) such that the two were more or less interchangeable. That said, there are probably some minor travel time effects embedded in some of the non-distance variables as well. Non-distance variables should be interpreted as the combination of travel time and non-travel time (e.g. perceived safety, effort, pleasantness, etc.) effects.

Turns likely delay cyclists, and they also add the mental cost of having to remember the correct sequence of turns. As expected, turn frequency is a significant negative factor in route choice. Once difficult left turns (across moderate to heavy traffic without a traffic signal) were accounted for, left and right turns were not significantly different, which seems reasonable.

Since distance enters the model in log form, each of the other variables in the model have marginal rates of substitution that are constant with respect to *percent* distance. Table 3 presents the estimation results in terms of this distance trade-off, which may be thought of as the distance value of each variable. In the case of turns, for instance, the model predicts that an additional turn per mile (0.6 turns/km) is equal to a 7.4 percent increase in non-commute distance and a 4.2 percent increase in commute distance.

Many permutations of elevation change and slope were tested, and the best performing specification was proportion of route length within three categories of average positive slope (gain/distance): 2-4 percent, 4-6 percent, and 6 percent. For example, a link traversing 500 ft (150m) with 10 ft (3m) gross gain along the traversal would have an average upslope of 2 percent and would be coded as 500 ft (150m) in the 2-4 percent upslope category. The consistently negative and strong sensitivity to slope contrasts with stated preference work (1,6).

Intersections

Data on intersection control and traffic volumes allowed us to construct a number of detailed intersection variables. Stop signs and traffic signals are delay factors for cyclists. At the same time, depending on the amount of conflicting traffic, signals might also be attractive features for cyclists trying to travel through or make turns across busy intersections.

As anticipated, in general traffic signals and, to a lesser extent, stop signs decrease the utility of a route. However, where conflicting traffic volumes are high, the positive effects of signals outweigh the negative. Whether this is because signals actually reduce delay at busy intersections, because they increase perceived safety, or some combination of the two is unclear. Right turns were excluded from most of the variables because such movements avoid most of the traffic conflicts and delays. Model fit with different specifications supported this distinction. To our knowledge, this is the first such result demonstrating the importance to cyclists of signalized intersections at busy street crossings.

Facility Types

Four bike-specific facility types were included in the final model: bike boulevards, off-street bike baths, bike lanes, and separated bike facilities on bridges. In addition, bike lanes were further divided into categories based on traffic volumes, and a separate category for bridge bike lanes was included. Bike boulevards are always on low traffic, neighborhood streets. Bike paths by definition have no motorized traffic. In addition to the facility types in the final model, designated bike routes were also tested as a facility

TABLE 3 Relative attribute values

Attribute	Distance value (% dist)	
	Non-commute	Commute
Turns (/mi)	7.4	4.2
Prop. upslope 2-4 %	72.3	37.1
Prop. upslope 4-6 %	290.4	120.3
Prop. upslope >= 6 %	1106.6	323.9
Traffic signal exc. right turns (/mi)	3.6	2.1
Stop sign (/mi)	0.9	0.5
Left turn, unsig., AADT 10-20k (/mi)	16.2	9.1
Left turn, unsig., AADT 20k+ (/mi)	43.1	23.1
Unsig. cross AADT >= 10k right turn (/mi)	6.7	3.8
Unsig. cross AADT 5-10k exc. right turn (/mi)	7.2	4.1
Unsig. cross AADT 10-20k exc. right turn (/mi)	10.4	5.9
Unsig. cross AADT 20k+ exc. right turn (/mi)	61.7	32.2
Prop. bike boulevard	-17.9	-10.8
Prop. bike path	-26.0	-16.0
Prop. AADT 10-20k w/o bike lane	22.3	36.8
Prop. AADT 20-30k w/o bike lane	137.3	140.0
Prop. AADT 30k+ w/o bike lane	619.4	715.7
Bridge w/ bike lane	-29.3	-18.2
Bridge w/ sep. bike facility	-44.9	-29.2

type. As expected, these unimproved bike routes were insignificant factors once the other variables were included in the model.

Bike boulevards and bike paths hold significant residual value even after controlling for all of the other variables in the model. For non-commute trips, travel on bike boulevards is equivalent to decreasing distance by almost 18 percent; by 26 percent on bike paths.

We tried many iterations of bike lane variables. Because bike lanes in Portland are almost exclusively on busy arterial streets, it was difficult to tease out the effect of bike lanes from that of traffic volume. In the final specification, bike lanes more or less exactly offset the negative effects of adjacent traffic but had no residual value of their own. This is consistent with the idea that bike lanes provide cyclists their own space separate from traffic but beyond this are no more or less attractive than a basic low traffic volume street. All else equal, the estimation suggests cyclists are willing to go considerably out of their way to use a bike boulevard or bike path rather than an arterial bike lane. This is not to suggest bike lanes are not valuable; if the alternative is an arterial street without a bike lane, then a designated lane has considerable value to cyclists. These results may not transfer to places where bike lanes are placed on low traffic volume streets.

On streets without bike lanes, cyclists are highly sensitive to high traffic volumes. In fact, the estimation suggests that for non-commute trips, streets with traffic volumes in excess of 20,000 vehicles per day would be used only if lower traffic alternatives required very long detours (in excess of 100 percent) or other strong deterrents such as steep hills. Within the city, Portland's bike network is fairly dense and well-developed, and it is not clear that this result would hold in places with a sparser network of bike facilities. It seems unlikely that a cyclist would actually choose a route seven times longer to avoid traveling on a highway without a bike lane; more likely, he would not travel by bicycle at all. Nonetheless, the result underscores the sensitivity of cyclists to high volumes of mixed traffic.

The Willamette River splits Portland as it runs south to north, separating Portland's central business district from largely residential areas east of the river. A little more than a quarter of observed trips crossed the Willamette on one of eight bridges available to bikes. Sampled cyclists were quite sensitive to bridge bike facilities. For non-commute trips, a bridge with a bike lane was equivalent to a 29 percent reduction in distance up to almost 45 percent for a separated bridge facility. Clearly, bridge facilities have a strong influence on cyclists' route choices for trips crossing the river.

Non-commute Versus Commute Trips

In general, the model suggests that commuting cyclists are relatively more sensitive to distance and less sensitive to most other variables. Commuting cyclists are likely under greater time pressure to reach their destination in the work direction. It is also possible that the more habitual nature of commute trips makes commuters more aware of distance and time differences among competing routes. It is also possible that commuters' knowledge of the route allows them to mitigate some of the delay and safety issues on commute trips. For example, they may learn the timing of traffic lights, how best to navigate intersections, and where to make difficult turns.

Exceptions to the above are found in the facility traffic volume attributes. Commuting cyclists are somewhat more sensitive to riding in high volumes of mixed traffic. The finding is consistent with the fact commutes are more likely to occur during periods of peak traffic.

Path-size Parameter

The path-size parameter estimate's positive coefficient is consistent with theory. It is significantly different from 1.0, which would be the expected value if the path-size parameter captured only the statistical error introduced by the independence from irrelevant alternatives (IIA) property of the MNL model. It has been suggested that the path-size parameter should not be arbitrarily fixed to 1.0, since it may have a meaningful behavioral interpretation (12).

In our case, estimating the parameter significantly improved model fit. Fixing the path-size parameter to 1.0 has the effect of reducing the magnitude of the distance coefficient while leaving the other parameters more or less unchanged. Since generated alternatives tend to cluster around the shortest-path, the greater than expected path-size correction may indicate unobserved disutility factors along shortest-path corridors. One plausible explanation is that many shortest paths in Portland involve a handful of busy, diagonal arterials that cut across the otherwise regular grid. These streets have generally poor riding environments which may not be fully captured by our observed attributes.

Another interpretation is that cyclists in our sample are less likely to distinguish between overlapping routes than statistically expected. That is, cyclists may consider two routes that overlap for just 25 percent of their lengths to be more similar than the physical overlap suggests. Perhaps they tend to share particularly unpleasant segments such as the diagonal arterials mentioned in the previous paragraph. Consistent with this hypothesis, a multi-modal route choice study found that trip “legs” rather than distance may sometimes be a better overlap measure (14).

CONCLUSIONS

This paper outlined the development of a unique bicycle route choice model based on revealed preference GPS data. The endeavor was made particularly challenging by the unusually large and dense travel network required to capture the options open to cyclists. A new choice set generation algorithm, dubbed the Calibrated Labeling Method, was developed to generate reasonable sets of alternatives after existing methods proved unsatisfactory.

The final model specification resulted in a rich range of insights into cyclist preferences which we are still exploring. Distance, turn frequency, slope, intersections, facility types, and traffic volumes all contribute significantly to a route’s attractiveness to bicyclists. Results highlight the importance for policymakers and planners of not only building bike facilities, but building them well. Details like busy street crossing treatments, route “jogs” necessitating extra turns, and siting to avoid slopes greater than 2 percent may prove as or more important than the facility itself. That said, bike boulevards and off-street bike paths appear to have inherent value to cyclists beyond the detailed facility variables we were able to measure. In other words, there is something more to a bike boulevard than low traffic volumes, improved street crossings, and “flipped” stop signs. The something more may be explained by attributes we were unable to measure, such as parking and traffic speeds, or perhaps something more subtle like perceived safety in numbers or simplified navigation. The results leave this intriguing question for future research.

The authors look forward to refining the model further. They also await with interest the results of similar studies using revealed preferences in different locations. For now, the question of how the Portland-based data will generalize to other places remains an open one. The model presented here is currently being implemented as a component of the regional travel demand forecasting model for the Portland region.

ACKNOWLEDGMENTS

The authors thank Oregon Metro for providing the base street network, the City of Portland for providing traffic volume estimates, and Jack Newlevant for his tireless collection of bicycle network data. The Active Living Research program of the Robert Wood Johnson Foundation provided funding for much of the original data collection. Oregon Metro and the Oregon Transportation Research and Education Consortium

(OTREC) provided funding for the model development. Of course, special thanks is due to the 164 Portland cyclists who took time to gather the GPS data for the study.

REFERENCES

1. Sener, I.N., N. Eluru, and C.R. Bhat. An Analysis of Bicycle Route Choice Preferences in Texas, US. Transportation, Vol. 36, 2009, pp. 511-539.
2. Krizek, K.J. Two Approaches to Valuing Some of Bicycle Facilities' Presumed Benefits. *Journal of the American Planning Association*, Vol. 72:3, 2006, pp. 309-319.
3. Hunt, J. D., and J.E. Abraham. Influences on Bicycle Use. *Transportation*, Vol. 34:4, 2007, pp. 453-470.
4. Tilahun, N.Y., D.M. Levinson, and K.J. Krizek. Trails, Lanes, or Traffic: Valuing Bicycle Facilities With an Adaptive Stated Preference Survey. *Transportation Research Part A*, Vol. 41:4, 2007, pp. 287-301.
5. Landis, B.W., V.R. Vattikutti, and M. Brannick. Real-Time Human Perceptions: Towards a Bicycle Level of Service. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1578*, TRB, National Research Council, Washington, D.C., 1997, pp. 119-126.
6. Stinson, M.A., and C.R. Bhat. Commuter Bicyclist Route Choice: Analysis Using a Stated Preference Survey. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1828*, TRB, National Research Council, Washington, D.C., 2003, pp. 107-115.
7. Aultman-Hall, L., F.L. Hall, and B.B. Baetz. Analysis of Bicycle Commuter Routes Using Geographic Information Systems: Implications for Bicycle Planning. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1578*, TRB, National Research Council, Washington, D.C., 1998, pp. 102-110.
8. Howard, C., and E.K. Burns. Cycling to Work in Phoenix: Route Choice, Travel Behavior, and Commuter Characteristics. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1773, Part 2: Bicycle and Pedestrian Research*, TRB, National Research Council, Washington, D.C., 2001, pp. 39-46.
9. Dill, J. L. and J. P. Gliebe. Understanding and Measuring Bicycling Behavior: A Focus on Travel Time and Route Choice: Final Report, OTREC-RR-08-03. Oregon Transportation Research and Education Consortium, 2008.
10. Python Programming Language. http://www.python.org/
11. Broach, J., J.G. Gliebe, and J.L. Dill. Calibrated Labeling Method for Generating Bicyclist Route Choice Sets Incorporating Unbiased Attribute Variation. In *Transportation Research Record: Journal of the Transportation Research Board,* TRB, National Research Council, Washington, D.C., in-press.
12. Frejinger, E. and M. Bierlaire. Capturing Correlation with Subnetworks in Route Choice Models. *Transportation Research Part B*, Vol. 41, 2007, pp. 363-378.
13. Neralla, S. and C.R. Bhat. Numerical Analysis of Effect of Sampling of Alternatives in Discrete Choice Models. In *Transportation Research Record:*

Journal of the Transportation Research Board, No. 1894, TRB, National Research Council, Washington, D.C., 2004, pp. 11-19.

14. Hoogendoorn-Lanser, S. R. van Ness, and P. Bovy. Path Size Modeling in Multimodal Route Choice Analysis. In *Transportation Research Record: Journal of the Transportation Research Board*, *No.* 1921, 2005, pp. 27-34.
15. Bekhor, S., M.E. Ben-Akiva, and M. Scott Ramming. Evaluation of Choice Set Generation Algorithms for Route Choice Models. *Annals of Operations Research*, Vol. 144:1, 2006, pp. 235-247.
16. Prato, C.G., and S. Bekhor. Applying Branch-and-Bound Technique to Route Choice Set Generation. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1985*, TRB, National Research Council, Washington, D.C., 2006, pp. 19-28.
17. Prato, C.G., and S. Bekhor. Modeling Route Choice Behavior: How Relevant Is the Composition of the Choice Set? In *Transportation Research Record: Journal of the Transportation Research Board, No. 2003*, TRB, National Research Council, Washington, D.C., 2007, pp. 64-73.
18. Ben-Akiva, M.E., and M. Bierlaire. Discrete Choice Methods and Their Applications to Short Term Travel Decisions. In *Handbook of Transportation Science*, Kluwer, Dordrecht, The Netherlands, 1999, pp. 5-33.
19. Bierlaire, M. BIOGEME: A free package for the estimation of discrete choice models , *Proceedings of the 3rd Swiss Transportation Research Conference*, Ascona, Switzerland, 2003.