Transportation Research Part A 67 (2014) 268–278

Contents lists available at ScienceDirect

Transportation Research Part A

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

Prioritizing bicycle paths in Belo Horizonte City, Brazil: Analysis based on user preferences and willingness considering individual heterogeneity

Dapeng Zhang [a], David José Ahouagi Vaz Magalhães [a,b,c], Xiaokun (Cara) Wang [a,*]

[a] Department of Civil and Environmental Engineering, Rensselaer Polytechnic Institute, USA
[b] Departamento de Engenharia de Transportes e Geotecnia, Universidade Federal de Minas Gerais, Brazil
[c] Bolsista da CAPES – Proc. no 3795/11-0, Brazil

A R T I C L E I N F O

Article history:
Received 17 October 2012
Received in revised form 10 June 2014
Accepted 28 July 2014
Available online 24 August 2014

Keywords:
Bicycle use
Random parameters
Ordered probit model

A B S T R A C T

Using bicycles as a commuting mode has proven to be beneficial to both urban traffic conditions and travelers' health. In order to efficiently design facilities and policies that will stimulate bicycle use, it is necessary to first understand people's attitudes towards bicycle use, and the factors that may influence their preferences. Such an understanding will enable reliable predictions of bicycle use willingness level, based on which cycling facility construction can be reasonably prioritized.

As people often have different perceptions on exercising, green transportation, and traffic conditions, effects of potentially influencing factors on people's willingness of using bicycles tend to be highly heterogeneous. This paper uses a random parameter ordered probit model to analyze how travelers' willingness of using bicycles is influenced by various socio-economic factors in Belo Horizonte City, Brazil, with the consideration of individual heterogeneity. The data was collected through the 2010 bicycle use survey in Belo Horizonte City. Results show that, first, the willingness of using bicycle is favored by middle income class household, and negatively related with commuting time. Second, people who rent apartments tend to be more willing to use bicycles. Third, if a person is currently walking a long time to work, he/she would be most willing to commute with a bicycle in the future. Those currently commuting a relatively short distance by motorcycle and bus follow this group in terms of willingness to commute by bicycle in the future. Car users seem to be difficult to convert to bicycle users. Moreover, the estimation shows clear evidence that significant individual heterogeneity indeed exists, especially for education level, necessitating the consideration of such an effect. With the calibrated model, residents' willingness of using bicycle commuting is then estimated for the entire Belo Horizonte City using the 2010 Census and the 2012 O/D survey data. The results are cross validated using the bicycle path preference information, also obtained from the 2010 bicycle use survey.

1. Introduction

The bicycle has always been considered an effective commuting mode because, in addition to its low cost, it improves the health of its user, reduces transportation expenses, and can be lifestyle enhancing (Marshall and Banister, 2000; Bauman,

* Corresponding author. Address: Department of Civil and Environmental Engineering, 4032 JEC Building, Rensselaer Polytechnic Institute, 110 8th Street, Troy, NY 12180-3590, USA. Tel.: +1 518 276 2098.

E-mail address: wangx18@rpi.edu (X. (C.) Wang).

http://dx.doi.org/10.1016/j.tra.2014.07.010
0965-8564/

2008; Shephard, 2008; Gordon-Larsen et al., 2009; Heinen, 2011). However, the use of bicycles has been quite limited in some regions. Commuting by bicycle can be seen as inconvenient, with long travel time and great physical effort required, especially during bad weather, or with heavy loads (Heinen et al., 2010). In recent years, as people become more aware of the externalities caused by motorized vehicles, more travelers are considering using bicycles again. Collectively, a greater share of bicycle use could reduce congestion, air pollution, noise, accidents, and infrastructure construction costs (Gardner, 1998; Pucher, 2010).

To promote the use of bicycles, and effectively design cycling facilities and management policies, it is critical to understand travelers' opinions towards bicycle use, especially those factors that may influence their willingness to use bicycles as their primary commuting mode. Heinen et al. (2010) group these potential factors into five categories: built environment factors; natural environment factors; socio-economic variables; psychological factors; and aspects related to cost, time, effort and safety. Many studies have investigated the effects of these factors on travel modes, and proposed approaches to improve the share of cycling trips. For example, studies found that a longer distance results in a lower possibility of bicycle use because it increases the time and effort for travelling (Moritz, 1998; Zacharias, 2005; Timperio et al., 2006; Parkin et al., 2008). Hilliness is found to have a negative effect on cycling, especially for inexperienced cyclists (Parkin et al., 2008). Rain, low temperature, darkness and wind make people give up or reduce the frequency of their use of bicycles (Heinen, 2011). Socio-economic factors have been confirmed to be tightly connected with cycling, although the factors are usually country-specific and unlikely causal (Moritz, 1998; Zacharias, 2005; Timperio et al., 2006; Parkin et al., 2008). Factors such as social norm and habits could be added into models in order to increase their explanatory power (Heinen, 2011). Studies have also confirmed cost, travel time, effort, and safety as crucial considerations for travelers (Heinen et al., 2011; Gordon-Larsen et al., 2009; Frank et al., 2008).

In recent years, Brazil has experienced rapid urbanization and motorization. Belo Horizonte City, the capital of the State of Minas Gerais, with a population of nearly 2.5 million, is one example of such development. The Belo Horizonte metropolitan region hosts one of the main industrial parks in Brazil. The volume of commuting trips to and from the central area is considerable. Falling behind of the rapidly developing economy, attractive transportation alternatives are still lacking: According to the 2002 household travel survey in Belo Horizonte, 71% of the daily trips are carried by motorized vehicles and only 0.7% by bicycles (BHTRANS, 2011a,b). As a consequence, significant traffic congestion problems have emerged: During peak hours, the average traveling speed in the urban center is approximately 9 km/h (BHTRANS, 2011a,b). In April 2012, the government of Brazil passed the Federal Law No. 12.587/2012, which targets at improving urban mobility and livability, and institutes the guidelines for the Urban Mobility National Plan (Política Nacional de Mobilidade Urbana). The key guidelines include integrating different transportation modes and prioritizing public transportation as well as non-motorized modes, featuring the creation of pedestrian streets and bicycle paths. In correspondence to these guidelines, public agencies in Belo Horizonte City have developed a program, Pedala BH (BHTRANS, 2011a,b), aiming at promoting bicycle use by constructing new facilities for cyclists. According to the program, there will be about 100 km of bicycle paths by the 2014 Soccer World Cup. The entire program will be fully implemented by 2020, by when a 365 km bicycle path network will cover most arterials in the city, as shown in Fig. 1. However, the program does not detail the construction priority of these bicycle paths. With limited budget and time, it is critical to prioritize the bicycle path construction properly so that the program can achieve maximum impacts early on, which will help secure supports for the continuation of such a long term policy.

To help determine which paths should be constructed first, an online survey was conducted in Belo Horizonte City to assess citizens' perception of cycling as the major commuting mode. The interviewees were asked to indicate their willingness levels of using bicycles for commuting as well as their preferred bicycle paths. The survey also collected respondents' basic demographic and socio-economic information.

The wealth of information provided by this survey allows for the rigorous analysis of bicycle use issues within an econometric framework. To accommodate the nature of the responses, ordered discrete outcome models can be used. Ordered models are normally used when outcomes are ordinal, such as quantitative ratings, ordered attitudinal opinions, and categorical frequencies (Washington, 2010). An ordered model defines a set of cut-points to divide the continuous outcome into several segments in order. The model can also quantify the effects of influential factors on the outcome by estimating coefficients of independent variables. However, studies have found that individual heterogeneity exists when different people have different perceptions of one specific factor (Conner and White, 1999; Arias, 2002). For factors influencing attitudinal opinions, such as the bicycle use willingness in this study, individual heterogeneity is very likely to exist. Such individual heterogeneity will lead to different effects generated by this factor, and eventually different outcomes.

Standard econometric models consider the parameter vector β to be fixed values, which implies that the effects of variables are constant across all observations. When individual heterogeneity needs to be addressed, a random parameter model is more reasonable. A random parameter model allows β to vary across observations, following given random distributions. In practice, researchers often assume a subset of parameters to be random based on their prior knowledge, and set remaining parameters to be fixed. A number of studies have utilized random parameter models, including Anastasopoulos and Mannering (2009), who used a count model with random parameters to explain vehicle accident frequencies, and Malyshkina and Mannering (2010), who also included random parameters into their model investigating the severity of vehicle accidents.

This paper also incorporates the consideration of individual heterogeneity in the ordered probit model. The resulted random parameter ordered probit model is statistically rigorous and behavior-consistent. Such a model will help obtain insights into people's perception towards bicycle use. The calibrated model also enables the estimation of bicycle path use intensity

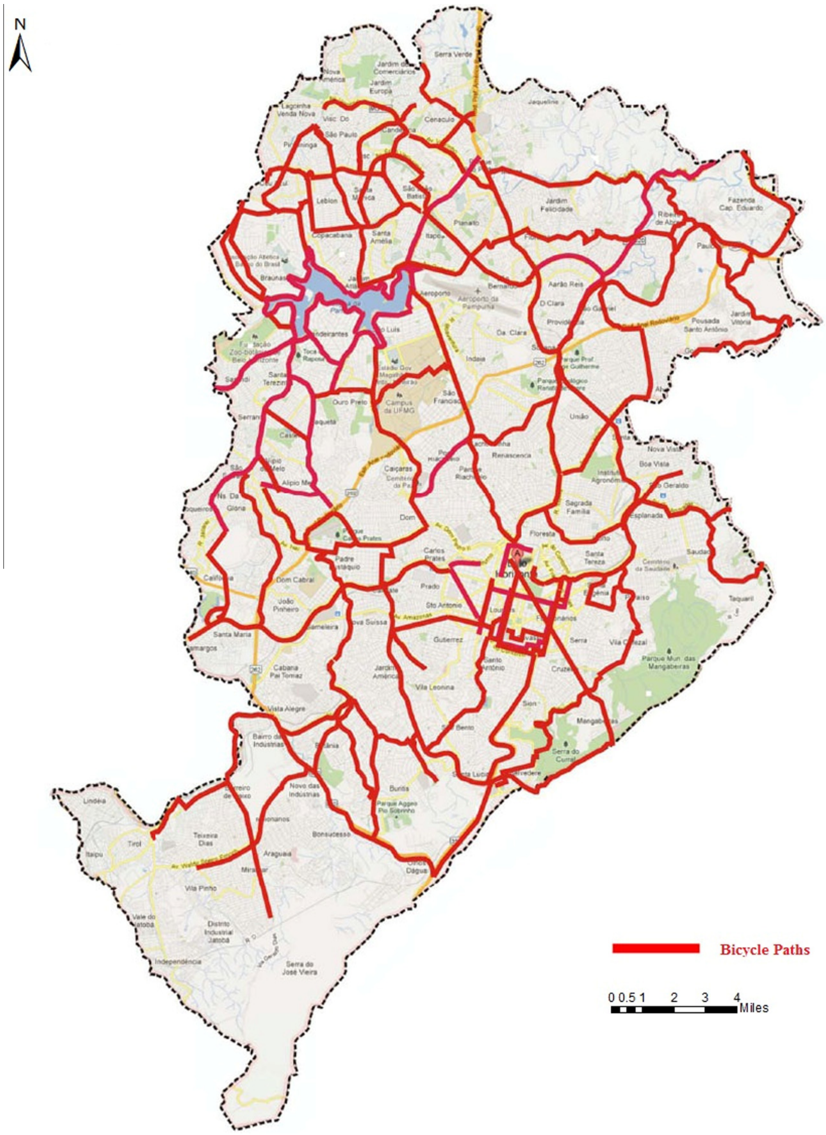

Fig. 1. Pedala BH bicycle path implementation plan (BHTRANS, 2011a,b).

in Belo Horizonte City: With the 2010 Census data and the 2012 O/D data (State of Minas Gerais, 2012), a composite bicycle use willingness index can be calculated for each potential cycling path. The resulted willingness index is then ranked and validated using information from a cycling path preference survey, facilitating the determination of cycling path construction priority.

The rest of this paper is structured as follows: the next section presents the specification of the random parameter ordered probit model, followed by a description of the survey and discussion of data. Analysis results are then discussed and applied to assess bicycle path construction priority, and important findings are summarized in conclusion.

2. Model specification

As discussed, the model must consider individual heterogeneity and the ordered discrete nature of the responses. The corresponding model specification starts from a standard ordered probability model, which captures the influence of independent factors on the ordinal ranking outcomes. The ordinal feature is derived from a latent variable y^*, which is typically specified as a linear function for each observation, such that

$$y^* = X\beta + \varepsilon \tag{1}$$

where X is a set of independent variables such as interviewee's socio-economic factors, β is the corresponding estimable parameters, and ε is the set of random disturbance terms representing the unobserved effects. The actual observed responses, y, defined as following:

$$
\begin{aligned}
&y = 1 \quad \text{if } y^* \leq \alpha_1 \\
&y = 2 \quad \text{if } \alpha_1 < y^* \leq \alpha_2 \\
&\ldots \\
&y = l \quad \text{if } \alpha_{l-1} < y^* \leq \alpha_l \\
&\ldots \\
&y = L \quad \text{if } y^* > \alpha_{L-1}
\end{aligned}
$$

where α is a set of cut-points that define the integer ordinal outcomes, with $y = 1$ presenting the lowest level response, and $y = L$ the highest level response. When disturbance term ε is assumed to be normally distributed, the model becomes an ordered probit model.

As mentioned in the Introduction, one person's understanding of, and sensitivity to any one potentially influential factor may differ from another person's. For example, some people are sensitive to commuting time, and willing to use a bicycle if it can shorten their commuting time. Others may be less sensitive to commuting time, and prefer a more "effortless" transportation mode, such as bus or car. Such a character – individual heterogeneity – will also be captured in the model.

Let indicate the index of observations ($q = 1, 2, 3, \ldots, Q$). Here Q denotes the total number observations). The random parameter β_q can be denoted by two components, a deterministic mean vector b and stochastic term $\tilde{\beta}_q$ following a multivariate normal distribution with mean $\boldsymbol{0}$ and variance Ω. The stochastic term $\tilde{\beta}_q$ represents the unique preference of individual q. Thus, the random parameter ordered probit model is specified as:

$$y_q^* = X_q\beta_q + \varepsilon \tag{2}$$

$$y_q = l, \quad \text{if } \alpha_{l-1} < y^* \leq \alpha_l$$

where the term $X_q\beta_q$ represents the effects of exogenous variables on the ordinal outcome. X_q is a $(1 \times K)$ vector of variables that corresponds to observation q. β_q is a $(K \times 1)$ vector of random coefficients that follows a K-dimensional multivariate normal distribution. Theoretically, all β_q can be treated as random coefficients, but due to consideration of empirical identifiability, most studies select part of β_q as random and hold others deterministic (Dugundji, 2005). In this study, coefficients for education level, commuting time, and the indicator variable for using car as current mode are set to be random because previous studies on their effects seem to be inconclusive, implying the possibility of high individual heterogeneity.

Let $\phi(\cdot)$ and $\Phi(\cdot)$ denote probability distribution and the cumulative distribution of a normal distribution. Let d_{ql} be an indicator variable taking 1 if observation q selects willingness level of l, and 0, otherwise. Then, the likelihood function of the qth observation can be written as

$$L_q = \int_\beta \left(\Phi(\alpha_l - \beta' x_q) - (\Phi(\alpha_{l-1} - \beta' x_q)\right)^{d_{qk}} \phi(\beta) d\beta \tag{3}$$

The parameters can be estimated by maximizing the product of L_q over the entire sample. This paper uses Maximum Simulated Likelihood Estimation method to approximate the value of integrals with Halton draws.

3. Data description

The data used in this study come from a 2010 urban life quality survey in Belo Horizonte (BH) City, Brazil. Magalhães (2008) led this survey, which consists of several groups of questions, including one concerning the respondent's willingness to use bicycles as their main commuting mode. The interviewees were asked "Would you use a bike to go to work if there were infrastructure needed for biking?" There were five potential willingness levels available to select, coded from 1 (not willing) to 5 (very willing). If the respondents gave positive answers, they were then asked to list five preferred streets/avenues for bicycle path construction. These preferred paths will be later used to validate the results obtained from model application. Before the survey was conducted, it was broadly advertised across the 67 planning zones of the city. 3421 interviewees took part in this survey, and 2134 of them provided complete information including demographic and socio-economic information, as well as opinions towards bicycle use. A preliminary check shows that the geographic distribution of observations is random, and covers the study area reasonably well. However, as an online survey, the pre-selection bias could exist. During the same time period, Brazil also conducted its national census, which is therefore used here to cross-validate the survey data; respondent age and household income distribution in the survey are compared to those in the Census. A chi-square test is also employed to check consistency and the results show that the overall distribution (indicated by mean and variance) of age and household income are consistent in these two datasets.

Furthermore, as the study focuses on bicycle use as the commuting mode, only observations with working ages (between 18 and 60) are considered, leaving 2,008 observations in the final analysis.

Table 1
Summary statistics of used variables.

Variable name	Dependent variable	Min	Max	Mean	Std. deviation
Will	Willingness in use of bikes (1, not willing; 2, somewhat unwilling; 3, neutral; 4, somewhat willing; 5, willing.)	1	5	2.88	1.53
	Independent variables				
Edu	Level of education (Binary: 1, advanced education and above; 0, high school and below.)	0	1	0.67	0.47
Time	Time for commuting (units: 10 min)	1	6	3.62	1.80
HH_inc1	Household income group 1 (Binary: 1, if income is between R$1531 and R$5100; 0, if not.) R$: Real (Brazilian currency) Purchasing power parity (PPP) factor: 0.8	0	1	0.42	0.50
HH_inc2	Household income group 2 (Binary: 1, if income is between R$5101 to R$10,200; 0, if not.)	0	1	0.26	0.44
HH_inc3	Household income group 3 (Binary: 1, if income is above R$10,201; 0, if not.)	0	1	0.18	0.38
House	Living in a house (Binary: 1, if the respondent lives in a house; 0, apartments and others.)	0	1	0.38	0.49
Own	Property owned (Binary: 1, if the respondent own the property; 0, rent and others.)	0	1	0.73	0.44
C_car	Current mode: car or taxi (Binary: 1, if current mode is car or taxi; 0, if not.)	0	1	0.43	0.50
C_trnst	Current mode: transit (Binary: 1, if current mode is transit; 0, if not.)	0	1	0.45	0.50
C_mtr	Current mode: motor (Binary: 1, if current mode is motorcycle; 0, if not.)	0	1	0.03	0.18
C_wlk	Current mode: walk (Binary: 1, if current mode is walk; 0, if not)	0	1	0.06	0.25
T_car	Interaction term Time × C_car	0	6	1.29	1.84
T_trnst	Interaction term Time × C_trnst	0	6	2.05	2.48
T_mrt	Interaction term Time × C_mrt	0	6	0.09	0.56
T_wlk	Interaction term Time × C_wlk	0	6	0.11	0.52
Kid	Kids in household (Binary: 1, there are kids in household; 0, no kids in household.)	0	1	0.06	0.24
NetDen	Network density (units: km/km^2)	8.5	92.8	47.00	9.76

Table 1 summarizes variables statistics for these 2008 observations. It is worth noting that the survey recorded commuting time using 20-min categories (0–20 min, 20–40 min, 40–60 min, and above 60 min). In the analysis, commuting time is approximated using the mean of these categories. In other words, the commuting time is treated as a continuous variable with 10 min as the unit. Household income is categorized into four groups, with income lower than R$1531 as the base case. The base of current mode is cycling.

4. Results analysis

The random parameter ordered probit model is then applied to the bicycle use data. Parameters of education level, commuting time and the indicator variable for "using car as current mode" are selected to be random parameters following normal distributions. These three parameters were selected because previous studies on their effects are rather inconclusive and their estimates changed significantly when different subsamples were used with the standard ordered probit model. The estimation result of the model and its comparison with a standard ordered probit model are shown in Table 2. The standard ordered probit model is run with both MATLAB code and STATA, as a way of validation. The results are consistent.

As Table 2 shows, in terms of the overall model goodness of fit, the random parameter model does not show significant superiority. However, the model finds that the effect of individual heterogeneity is indeed significant, as indicated by the statistical significance of parameters' variances. The consideration of such effect provides better understanding of factors' impacts. Table 3 summarizes the variables' marginal effects derived from the random parameter ordered probit model.

Estimation results of the random parameter model provide some interesting insights into the bicycle use problem. The parameter of education level is normally distributed with mean −0.066 and variance 0.886. Both the mean and variance are statistically significant, indicating that the effect of parameter varies across observations. With the estimated parameter, 52.79% of the distribution is less than 0, and 47.21% is greater than 0. This implies that the level of education has a bimodal effect on willingness to use bicycles. Such heterogeneous effect is also indicated by its marginal effect. If one's education increases from 0 (high school and below) to 1 (advanced education and above), the probability that people selecting 1 (not willing) and 5 (very willing) will both increase (by 10.59% and 5.33%, respectively), while the probabilities of the intermediate willingness levels (2, 3 and 4) will all drop. Apparently, because of the accommodation of individual heterogeneity, a variable's average marginal effect in this random parameter model does not necessarily present the one-directional shift as in a standard ordered probit model (which would be indicated by marginal effects' opposite signs in the two extreme levels). In most current literature, education was found to have one-directional effect in travel demand models. For example, Goetzke and Weinberger (2012) found that non-car-ownership was a status symbol for people with post baccalaureate education. Boarnet and Crane (2001) found that people with college and higher degree were associated with high trip frequency. Choo and Mokhtarian (2004) also found that less-educated people tend to drive pickups and large cars. Some other studies have discussed the heterogeneity associated with educational level. For example, Brand (2010) found that education has heterogeneous effects on civic participation. Brand and Davis (2011) found that higher-educated women are more heterogeneous in terms of family formation patterns. Brand (2012) found that the heterogeneity effect exist in the decision of attending four-year colleges or community colleges. In short, people with higher education tend to be more heterogeneous in terms of their socio-economic behaviors.

Table 2
Summary of model estimation results.

Dependent variable: willingness in use of bikes	Model (1) Standard ordered probit		Model (2) Random parameter ordered probit	
Variable name	Parameters estimated	*t*-Stat	Parameters estimated	*t*-Stat
Edu	−0.082	−1.51	−0.066	−1.08
Time	−0.166	−1.85	−0.057	−0.57
HH_inc1	0.035	0.46	0.022	0.26
HH_inc2	−0.017	−0.20	−0.051	−0.52
HH_inc3	−0.004	−0.05	−0.023	−0.22
House	−0.035	−0.67	−0.043	−0.73
Own	−0.046	−0.82	−0.057	−0.89
C_car	−0.703	−2.08	−0.271	−0.74
C_trnst	−0.378	−1.08	0.037	0.10
C_mtr	−0.329	−0.75	0.095	0.20
C_wlk	−0.263	−0.71	0.236	0.59
T_car	0.046	0.50	−0.086	−0.82
T_trnst	0.037	0.40	−0.080	−0.77
T_mtr	0.100	0.78	−0.006	−0.04
T_wlk	0.151	1.26	0.023	0.17
Kid	−0.004	−0.08	−0.008	−0.14
NetDen	0.0005	0.20	−0.019	−0.10
Variance of Edu			0.886	2.34
Variance of Time			0.034	2.56
Variance of C_car			0.390	1.60
Threshold 1	−1.81	−5.56	−1.60	−2.05
Threshold 2	−1.25	−5.66	−0.96	−1.96
Threshold 3	−0.98	−6.99	−0.65	−2.40
Threshold 4	−0.35	−6.76	0.07	−2.03
Log simulated likelihood (constant only)	−3236		−3236	
Log simulated likelihood	−3081		−3079	

Table 3
Marginal effects of variables in random parameter ordered probit model.

	Willingness level (%)				
	1	2	3	4	5
Edu	10.59	−3.09	−3.01	−9.82	5.33
Time	3.14	−0.09	−0.50	−2.54	−0.01
HH_inc1	−0.44	−0.23	−0.09	−0.05	0.80
HH_inc2	1.03	0.54	0.18	0.08	−1.82
HH_inc3	0.46	0.25	0.08	0.05	−0.83
House	0.86	0.45	0.16	0.07	−1.54
Own	1.15	0.60	0.21	0.08	−2.03
C_car	10.15	0.07	−1.13	−6.24	−2.85
C_trnst	−0.72	−0.40	−0.13	−0.11	1.36
C_mtr	−1.79	−1.06	−0.33	−0.35	3.53
C_wlk	−4.20	−2.57	−0.94	−1.15	8.88
T_car	1.75	0.91	0.28	0.08	−3.03
T_trnst	1.63	0.84	0.27	0.08	−2.82
T_mtr	0.11	0.07	0.02	0.02	−0.22
T_wlk	−0.46	−0.24	−0.09	−0.05	0.84
Kid	0.15	0.09	0.03	0.02	−0.29
NetDen	0.38	0.20	0.07	0.04	−0.69

The parameter of commuting time is normally distributed with mean −0.057 and variance 0.034. With the estimated parameter, 62.14% of the distribution falls lower than 0, and 37.86% is greater than 0. This implies that commuting time has a mixed effect on bicycle use, and negative for most respondents. If commuting time increases by 10 min, the probability that people select 1 (very unwilling) will increase by 3.14%. The probability that people select 2 (somewhat unwilling, 3 (neutral), 4 (somewhat willing), and 5 (very willing) will decrease by 0.09%, 0.50%, 2.54%, and 0.01%, respectively. The reason is that long commuting time means intensive physical effort from the cyclists and a high propensity to use cars or public transit. Such a result shows a one-direction effect of commuting time toward willingness level, although the corresponding parameter is set to be random.

The coefficient of "car as current commuting mode" is normally distributed with mean −0.271 and variance 0.390. With this estimate, 66.78% of the distribution is less than 0, indicating that most current car users are much less likely to adopt bicycle use than are people who currently use bicycles. Compared to a current bicycle user, the probability of a car user selecting 1 (very unwilling) is higher by 10.15%. It seems that car users tend to stay with their current mode. Furthermore, the interaction term of commuting time and current car mode shows that people are even less likely use bicycles when they commute by cars for a long time. For car users, 10 min increase in travel time will reduce the utility by an average of 0.143, with a variance of 0.034. Assuming 50 km/h average travel speed, this result can be interpreted as every 1 km increase of travel distance will lower the utility of car users by an average of 0.017, with a variance of 4.81×10^{-4}. Accordingly, if car users' commuting distance increases by 1 km, the probability that they select 1 (very unwilling) will increase by 0.58%, and the probabilities of selecting 2, 3, 4, and 5 will decrease by 0.002%, 0.20%, 1.05%, and 0.34%, respectively. Apparently, in the cases of long travel time and distance, cars provide a more comfortable environment and require less physical efforts. For current transit users and motorcyclists, the coefficients are not significantly different from 0, indicating that current transit users and motorcyclists may have similar behavior as cyclists. And generally speaking, people are unwilling to use bicycles when their commuting time is long by these two modes, as indicated by the respective interaction terms. Compared to current bicycle users, the probability of people currently walking selecting 5 (very willing) is higher by 8.88%. They are very likely to use bicycles if proper cycling infrastructure is provided. And as indicated by the interaction term of commuting time and current walking mode, the longer the current walking time is, the higher the possibility that they are willing to shift to bicycles. When appropriate travel speeds for walking, transit and motorcycles are assumed, the different effects of commuting distance can also be derived in a way that is similar to the car user analysis. Collectively, current commuting modes and the corresponding commute time and distance have effects on people's willingness towards bicycle use. People may be attracted by the convenience and healthy life-style provided by cycling, but deterred by the physical efforts associated with long commuting. These findings suggest that policies aiming at increasing the share of bicycle use should first target people who commute a long distance by walking, and then motorcyclists and bus riders with short commuting distance. Car users are much less likely to be converted than users of all other modes.

The coefficients of household income groups also provide some interesting insights. Group 1 (household income between R$1531 and R$5100, around the same purchasing power of $1225 to $4080 in the U.S., accounting for 42% of the entire sample) has the largest positive coefficient. Compared to the low income group (household income below R$1531), the probability of a middle-level income household selecting 5 (very willing) are higher by 0.76%. The high income group (groups 2 and 3) seems to be less enthusiastic about using bicycles than the middle-level income group, compared to the low income group, the probability of choosing 5 (very willing) are lower by 1.82% and 0.83%. This result is consistent with findings in previous studies that cycling is often perceived as a middle class transportation mode (Steinbach et al., 2011).

Even with current transportation modes and household income controlled, people living in houses are still less willing to use bicycles than those living in apartments, with probabilities of choosing 5 (very willing) lower by 1.54%. It seems that houses, implying communities with low residential density, are not optimal targets for developing bicycle facilities, as compared to high density residential communities which contain mostly apartments.

People who own their property are less willing to use bicycles than are renters, with probabilities of choosing 5 (very willing) levels lower by 2.03%. As this indicator variable is likely to be a proxy of age and lifestyle, the estimation result may suggest that people who have a settled lifestyle are less likely to become bicycle users.

Having kids has a negative effect on bicycle use. Compared to households with no kids, the possibility of a household with kids choosing 5 (very willing) decreases by 0.29%. This result implies that households with kids tend to use other modes of transportation. Parents may combine their commuting trip with trips involving kids. Thus, physical efforts and safety issues turn to be the major concerns. Therefore, family structure is an important part of considerations about their commuting mode.

Network density has a negative effect on bicycle use. For 1 km additional length of street in an area, the probability of people living there selecting 5 (very willing) is 0.69% lower than other areas. This finding indicates that other modes of transportation are more competitive in locations with dense road network.

In summary, bicycle use is favored by households with middle-level income. It is largely negatively influenced by commuting time and residence ownership. If a person is currently walking a long time to work, he/she would be the most willing to use a bicycle as their major commuting mode in the future, followed by motorcyclists and bus riders in short time. Car users seem to be the most difficult group to convert to bicycle use. Moreover, the different estimations show clear evidence that individual heterogeneity indeed exists, as indicated by the significant estimate variance of education level, travel time and "car as current commuting model." Thus it is necessary to consider such an effect in the analysis.

5. Bicycle path prioritization based on model outputs

The outputs of the above model can be used to estimate people's tendency of using bicycles on each potential bicycle path. Paths with higher usage intensity could be given higher priority in the construction plan. This section explains how each path's potential usage intensity can be estimated and ranked.

The Belo Horizonte City's 2010 Census divides the city into 67 spatial study zones, which is the smallest spatial unit in all available datasets. For each zone, the census data provides a representative sample with individual socioeconomic and commuting time information. This dataset is supplemented by the 2012 O/D survey, which further provides the zonal level travel mode information. The calibrated model and the O/D survey data allow the estimation of average willingness level for each zone z, denoted as WL_z.

$$WL_z = \sum_{k=1}^{K} X_k \beta_k \tag{4}$$

where X_k is the k-th explanatory variable in the estimated model, and β_k is the associated estimated parameter. WL_z can be seen as a unitless latent value that measures the utility for residents in zone z to use bicycles for commuting. A higher WL_z implies a higher tendency for zone z's residents to use bicycles. Fig. 2 shows the distribution of this average willingness level. Apparently, residents in the northeast and southwest of the city are more willing to use bicycle commuting when provided necessary facilities.

Residents are assumed to have access to a bicycle path if they live close to the path. Therefore, buffer areas can be generated along the potential bicycle paths to estimate each path's overall popularity based on the willingness and the total number of the residents it serves. This study calculates and compares three buffer widths: 0.25 miles, 0.5 miles, and 1 mile, which represent the range of walkable to bikable distances. The usage intensity index of bicycle path p, UI_p, can be derived as

$$UI_p = \sum_{z=1}^{Z_p} S_{zp} WL_z \tag{5}$$

where S_{zp} denotes the size of area where path p's buffer zone and spatial study zone z intersects. Z_p is the total number of spatial study zones covered by path p's buffer zone. If it is assumed that residents in each spatial study zone are evenly

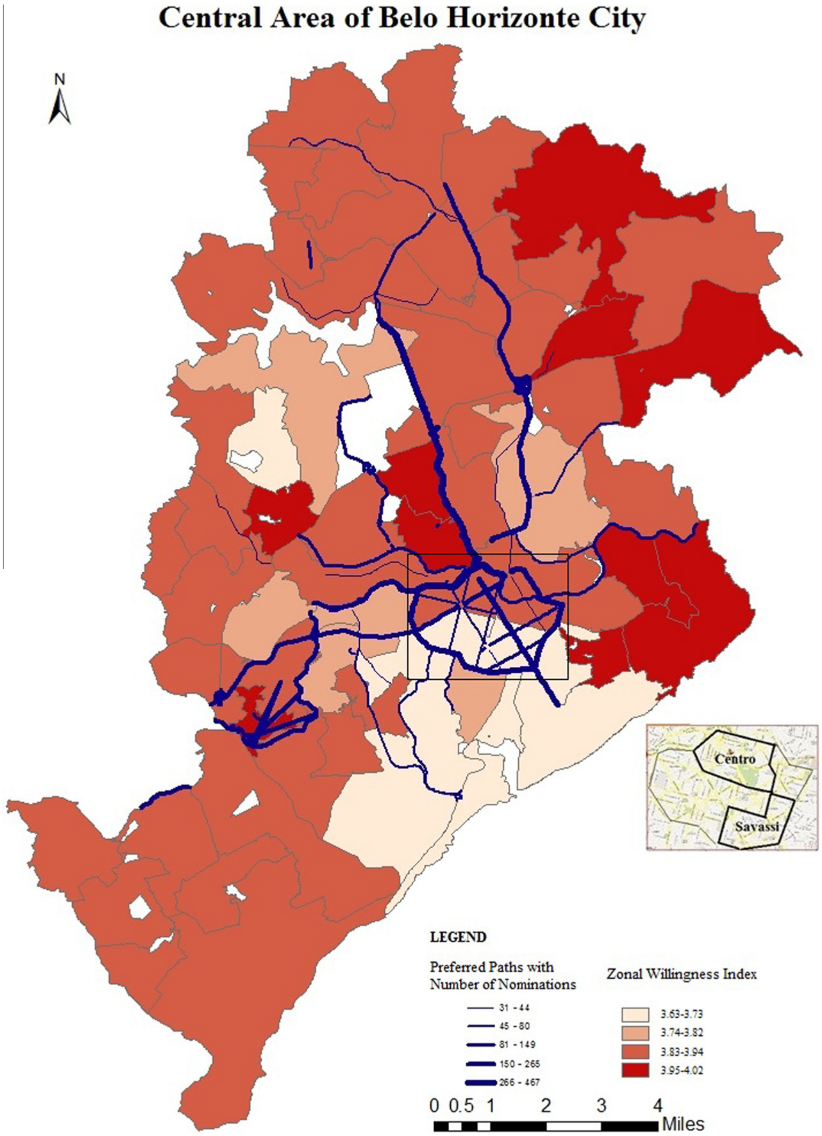

Fig. 2. Preferred bicycle paths and zonal willingness indexes.

distributed in space, S_{zp} is proportional to the population served by path p in zone z. If it is further assumed that residents' willingness to use bicycle commuting are homogeneous in each zone, i.e., WL_z is representative for zone z, UI_p can be interpreted as the all path p users' total utility, which roughly reflects the path's total potential usage. Given the fact that the spatial study zones are small with homogeneous land use, such assumptions should be reasonable. As UI_p is derived from the unitless latent variable WL_z, the absolute value of UI_p cannot be interpreted numerically. However, the UI_p value can be ranked and used to compare the proposed bicycle paths.

Meanwhile, as mentioned in Data Description, the 2010 urban life quality survey also asked respondents to nominate up to five "preferred roads" for cycling path construction. Each time a road is recommended by a respondent, it is counted as one "nomination". Overall, the top 30 preferred roads (17% of all candidate roads) account for about 80% of all nominations. Fig. 2 shows the locations of these top 30 roads. The UI_p for these roads are calculated and ranked, as shown in Table 4. Table 5 further summarizes the basic statistics of WL_z and UI_p.

The correlation between the UI_p ranking and the surveyed ranking is examined using the Spearman's rank correlation coefficient (Spearman, 1904), the Kendall's Tau rank correlation coefficient (Kruskal, 1958), and the Goodman and Kruskal's Gamma rank coefficient (Goodman and Kruskal, 1954). The results are shown in Table 6.

Coefficients are found significantly positive for all combinations of buffer bandwidths and correlation coefficient calculation approaches. The results indicate that the rankings of the usage intensity indexes UI_p, which are derived from the model, and the surveyed ranking are highly correlated. In other words, the analysis and prediction framework developed in this paper is validated as a good approach to assess the potential usage popularity of bicycle paths, which can help determine the bicycle path construction priority. The model estimates, derived with disaggregate data covering the entire region, can be applied to assess the usage popularity of any proposed bicycle paths in this region. In other words, alternative bicycle path construction plans can also be evaluated using the model outputs, without relying on further surveys of "preferred roads."

The results suggest that the roads in the north (Av. Antonio Carlos and Av. Cristiano Machado), west (Av. Amazonas, Av. Tereza Cristina and via Expressa), northwest (Av. Dom Pedro II, Av. Pres. Carlos Luz) and east (Av. dos Andradas) should get the highest priority for bicycle path construction, in addition to some of the most important avenues in the Central Area (Av. do Contorno, Av. Afonso Pena, Av. Amazonas and Av. Getulio Vargas). Belo Horizonte City's Central Area houses the city's primary central business district (CBD), Centro, and the secondary CBD, Savassi. All of these roads are important links between residential areas and workplaces.

Table 4
Usage intensity analysis of top 30 nominated paths.

Path name	Number of nominations	Nomination ranking	0.25 miles		0.5 miles		1 mile	
			UI_p	UI_p ranking	UI_p	UI_p ranking	UI_p	UI_p ranking
Av. do Contorno	467	1	13.40	3	39.80	1	73.20	3
Av. Afonso Pena	361	2	5.32	14	15.66	13	41.15	11
Av. Antônio Carlos	339	3	10.00	6	20.81	8	45.46	8
Av. Amazonas	265	4	11.42	4	32.00	4	71.37	4
Av. Cristiano Machado	260	5	14.11	2	39.30	2	81.87	1
Av. Dom Pedro II	149	6	15.77	1	17.90	12	43.21	9
Av. dos Andradas	134	7	7.07	9	23.71	7	52.80	7
Av. Presidente Carlos Luz	131	8	10.18	5	38.50	3	75.47	2
Av. Teresa Cristina	112	9	4.29	16	19.50	10	40.73	12
Av. Getúlio Vargas	112	9	2.58	28	14.51	14	33.91	15
Av. Nossa Senhora do Carmo	107	11	6.83	10	4.96	30	15.34	30
Via Expressa	100	12	6.20	12	9.13	27	30.68	18
Av. Brasil	92	13	4.65	15	28.80	5	70.75	5
Av. Dom Pedro I	80	14	3.07	25	8.65	28	21.50	29
Av. Prudente de Morais	76	15	7.11	8	10.17	25	29.97	23
Av. Raja Gabaglia	66	16	6.32	11	19.10	11	39.72	13
Av. Barão Homem de Melo	62	17	4.03	19	20.46	9	42.78	10
Av. Bias Fortes	62	17	3.07	25	11.61	20	30.59	19
Av. Silviano Brandão	58	19	3.31	24	13.28	16	31.20	17
Av. Augusto de Lima	56	20	4.19	18	8.52	29	29.36	22
Av. José Cândido da Silveira	50	21	3.65	21	13.02	18	29.85	21
Av. Olegário Maciel	49	22	3.34	23	11.17	21	28.68	26
R. da Bahia	49	22	2.11	29	13.17	17	34.95	14
R. Jacuí	44	24	4.23	17	24.80	6	58.33	6
R. Padre Eustáquio	41	25	1.38	30	10.21	23	28.97	24
Av. Cristóvão Colombo	40	26	3.56	22	10.20	24	25.18	28
Av. Silva Lobo	37	27	6.10	13	14.46	15	31.93	16
Av. Portugal	36	28	7.88	7	9.93	26	25.76	27
Av. Vilarinho	33	29	3.86	20	10.61	22	28.75	25
Av. Abílio Machado	31	30	2.89	26	12.15	19	29.89	20

Table 5
Summary statistics of the willingness index and the usage intensity.

WL_z		Mean	Std	Min	Max
		2.94	0.26	2.23	3.70
Buffer width					
UI_p	0.25	6.06	3.68	1.38	15.77
UI_p	0.5	17.54	9.49	4.96	39.80
UI_p	1	40.78	17.44	15.34	81.87

Table 6
Summary of the rank correlation coefficients.

Buffer width (mile)	Spearman	*p*-Value	Kendall's Tau	*p*-Value	Goodman and Kruskal's Gamma	*p*-Value
0.25	0.60***	0.0005	0.43***	0.0009	0.15***	0.0001
0.5	0.54***	0.0020	0.36***	0.0051	0.18***	0.0000
1	0.63***	0.0002	0.44***	0.0007	0.14***	0.0003

*** $p < 0.001$.

6. Conclusions

This study investigates people's willingness of bicycle use in Belo Horizonte City, Brazil, using a random parameter ordered probit model.

In the random parameter ordered probit model, the random parameters ensure the flexibility to capture the individual heterogeneity that can influence outcomes. The estimation results disclose some interesting findings: the willingness level of using bicycles as the major commuting mode is influenced substantially by a range of socio-economic factors. People who are in middle income class, rent, live in apartment-like dwellings, walk a long time to work, and commute a shorter time by other modes are more likely to use bicycles. Current car users are very unlikely to be converted to bicycle users, compared to users of others. Education has a fairly dual effect, confirming the expectation of individual heterogeneity. These findings have important policy implications. In order to promote bicycle use in Belo Horizonte City efficiently, policy makers should focus primarily on providing necessary infrastructure to high-density communities, and targeting middle income commuters.

Based on the calibrated model, the usage intensity index for each path is also derived, which is then validated by the bicycle path preference stated by interviewees. The ideal bicycle path construction priority can be inferred from these results.

Thanks to the data unavailability and computation complexity, this study has rooms for improvements. In the survey, the commuting time was collected instead of commuting distance, which is believed as a more objective measurement. Furthermore, more parameters may be set as random when higher computation power is achieved.

The analysis framework developed in this study provides valuable insights into people's opinions towards the use of bicycles as a primary mode of transportation. And the model application provides a practical way to transfer people's attitude to optimize bike paths construction. The findings are especially important for both researchers and practitioners in countries where extensive infrastructure development is planned to encourage non-motorized transportation. Such an analysis enhances the understanding of bicycle use and will eventually improve the efficiency of infrastructure planning targeting a more sustainable urban transportation system.

Acknowledgements

The writers would like to acknowledge Foundation for the Coordination of Improvement of Higher Education Personnel (CAPES, Brazil) and National Science Foundation (NSF) for the support on this collaborative research.

References

Anastasopoulos, P.C., Mannering, F.L., 2009. A note on modeling vehicle accident frequencies with random-parameters count models. Accid. Anal. Prev. 41 (1), 153–159.

Arias, O., Hallock, K., et al., 2002. Individual heterogeneity in the returns to schooling: instrumental variables quantile regression using twins data. In: Fitzenberger, B., Koenker, R., Machado, J.F. (Eds.), Economic Applications of Quantile Regression, Physica-Verlag, HD, pp. 7–40.

Bauman, A., Rissel, C., et al., 2008. Cycling: getting Australia moving – barriers, facilitators and interventions to get more Australians physically active through cycling.

BHTRANS, 2011. Program Pedala BH. Retrieved 09.06.14, from http://www.bhtrans.pbh.gov.br/.

BHTRANS, P.d.B.H., 2011. Empresa de Transportes e Transito de Belo Horizonte S.A./Observatório da Mobilidade Urbana Sustentável, Balanço da Mobilidade Urbana de Belo Horizonte – 2010.

Boarnet, M., Crane, R., 2001. The influence of land use on travel behavior: specification and estimation strategies. Transp. Res. Part A 35 (9), 823–845.

Brand, J.E., 2010. Civic returns to higher education: a note on heterogeneous effects. Soc. Forces 89 (2), 417–433.

Brand, J.E., Pfeffer, F.T., et al., 2012. Interpreting Community College Effects in the Presence of Heterogeneity and Complex Counterfactuals, California Center for Population Research, University of California, Los Angeles. http://papers.ccpr.ucla.edu/papers/PWP-CCPR-2012-004/PWP-CCPR-2012-004.pdf (accessed June 2013).
Brand, J.E., Davis, D., 2011. The impact of college education on fertility: evidence for heterogeneous effects. Demography 48 (3), 863–887.
Choo, S., Mokhtarian, P.L., 2004. What type of vehicle do people drive? the role of attitude and lifestyle in influencing vehicle type choice. Transp. Res. Part A 38 (3), 201–222.
Conner, M.M., White, G.C., 1999. Effects of individual heterogeneity in estimating the persistence of small populations. Nat. Res. Model. 12 (1), 109–127.
Dugundji, E., Walker, J., et al, 2005. Discrete choice with social and spatial network interdependencies: an empirical example using mixed generalized extreme value models with field and panel effects. Transp. Res. Rec. 1921 (-1), 70–78.
Frank, L., Bradley, M., et al, 2008. Urban form, travel time, and cost relationships with tour complexity and mode choice. Transportation 35 (1), 37–54.
Gardner, G., 1998. When cities take bicycles seriously. World Watch Magaz. 11, 16–22.
Goetzke, F., Weinberger, R., 2012. Separating contextual from endogenous effects in automobile ownership models. Environ. Plan. Part A 44 (5), 1032.
Goodman, L.A., Kruskal, W.H., 1954. Measures of association for cross classifications. J. Am. Stat. Assoc. 49 (268), 732–764.
Gordon-Larsen, P., Boone-Heinonen, J., et al, 2009. Active commuting and cardiovascular disease risk: the cardia study. Arch. Intern. Med. 169 (13), 1216–1223.
Heinen, E., Maat, K., et al, 2011. Day-to-day choice to commute or not by bicycle. Transp. Res. Rec. 2230 (-1), 9–18.
Heinen, E., Wee, B.V., et al, 2010. Commuting by bicycle: an overview of the literature. Transp. Rev. 30 (1), 59–96.
Heinen, E., Maat, K., et al, 2011. The role of attitudes toward characteristics of bicycle commuting on the choice to cycle to work over various distances. Transp. Res. Part D 16 (2), 102–109.
Kruskal, W.H., 1958. Ordinal measures of association. J. Am. Stat. Assoc. 53 (284), 814–861.
Magalhães, D.J.A.V., Oliveira, A.G.R., 2008. Desenvolvimento de indicadores municipais de satisfação da população quanto à localização residencial, mobilidade e acessibilidade no espaço urbano, XVI Encontro Nacional de Estudos Populacionais.
Malyshkina, N.V., Mannering, F.L., 2010. Empirical assessment of the impact of highway design exceptions on the frequency and severity of vehicle accidents. Accid. Anal. Prev. 42 (1), 131–139.
Marshall, S., Banister, D., 2000. Travel reduction strategies: intentions and outcomes. Transp. Res. Part A 34 (5), 321–338.
Moritz, W., 1998. Adult bicyclists in the United States: characteristics and riding experience in 1996. Transp. Res. Rec. 1636 (-1), 1–7.
Parkin, J., Wardman, M., et al, 2008. Estimation of the determinants of bicycle mode share for the journey to work using census data. Transportation 35 (1), 93–109.
Pucher, J., Dill, J., et al., 2010. Infrastructure, programs, and policies to increase bicycling: an international review. Prevent. Med. 50, Supplement (0): S106–S125.
Shephard, R.J., 2008. Is active commuting the answer to population health? Sports Med. 38 (9), 751–758.
Spearman, C., 1904. The proof and measurement of association between two things. Am. J. Psychol. 15 (1), 72–101.
State of Minas Gerais, 2012. The 2012 Greater Belo Horizonte Origin-Destination Survey, Belo Horizonte.
Steinbach, R., Green, J., et al, 2011. Cycling and the city: a case study of how gendered, ethnic and class identities can shape healthy transport choices. Soc. Sci. Med. 72 (7), 1123–1130.
Timperio, A., Ball, K., et al, 2006. Personal, family, social, and environmental correlates of active commuting to school. Am. J. Prev. Med. 30 (1), 45–51.
Washington, S.P., Karlaftis, M.G., et al., 2010. Statistical and Econometric Methods for Transportation Data Analysis, CRC Press.
Zacharias, J., 2005. Non-motorized transportation in four shanghai districts. Int. Plan. Stud. 10 (3–4), 323–340.