Exploring non-attendance of situation attributes in choosing the bike as a transport mode: the case of Belgium

Abolghassem Jandari, Michel Meulders, Martina Vandebroek
Faculty of Economics and Business, KU Leuven

Abstract

Cycling is an important pillar of the global endeavour to have a more sustainable transportation system. A lot of information is available in the literature regarding the impact of different factors on cycling. However, no research has been conducted on heuristics individuals may use when choosing the bike as a transport mode. We try to answer this question in the current paper. We modelled the probability of cycling using a binary item response model where the choice is modelled as a trade-off between the individuals' tendency to cycle and the threshold of each cycling situation. Accounting for heuristics has been mainly studied in discrete choice modelling, but in this study, we will explore attribute non-attendance in the analysis of binary item responses. We classified the respondents into two groups of frequent and occasional cyclists. The occasional cyclists were more affected by adverse levels of situation attributes. Contrary to previous studies, a separate bike path was a stronger motivator for the group of frequent cyclists. The model fit was substantially improved after accounting for attribute non-attendance. The weather condition and the bike path had the highest and lowest attendance probability respectively.

1 Introduction

Finding ways of promoting people to cycle more has been an interesting objective for both researchers and policymakers in the past two decades. Cycling is a physical activity that some people may find troublesome, or that they have safety concerns about (Rodríguez and Joo, 2004). Therefore, in addition to the usual variables used in classic transport mode choice models, more variables are needed to explain the probability of choosing the bike.

Many studies attempted addressing people's propensity towards cycling using different types of variables. Heinen et al. (2010) introduced five groups of variables that affect cycling: the built environment, the natural environment (e.g., weather, landscape), socio-economic variables, psychological factors and aspects related to the cost, time, effort and safety. There is a significant amount of information available in the literature regarding the effect of the variables in each group and comprehensive literature reviews are provided in Buehler and Dill (2016), Handy et al. (2014), Fernández-Heredia et al. (2014), Heinen et al. (2010) and Pucher et al. (2010).

In this paper, we used a binary item response model to estimate the probability of cycling. In our model, the choice is explained as a trade-off between the individuals' tendency to cycle and the threshold of each cycling situation. Based on the results of a binary choice experiment, we analyze the effects of rain, wind, bike path, slope, light, and distance on the probability of taking the bike using the data of 562 respondents. Accounting for ANA has mainly been studied in discrete choice models. But here, we will use it to model binary item responses, i.e., whether or not a person will take the bike in a certain situation. Accounting for ANA when modelling choice behavior is important for two reasons. First, if individuals are ignoring attributes, it means that they have a non-compensatory choice behavior. So, any improvement in the ignored attribute will not compensate individuals for the negative effect of another attribute level. Lagarde (2013)

explains that this will violate the assumption about the continuity of individuals' preferences, and therefore, the preferences should not be represented by the conventional utility functions. Second, ignoring ANA heuristics could lead to biased coefficients and then biased policy recommendations (Lagarde, 2013). Many studies tried to understand cycling behavior, but none has investigated the role of heuristic rules in choosing the bike as a transport mode. The importance and originality of this study are that it explores ANA in cycling behavior when modelling binary item-responses. We studied cycling in Belgium, a country where most people are familiar with cycling, but there is still much room to increase the share of cycling. The research to date has mainly focused on encouraging non-cyclists to use the bike. Unlike previous studies, we are not only focusing on the ways of motivating non-cyclists to use the bike, but we also study those people who already cycle but whose level and purpose of using the bike are different.
Belgium is ranked top five in the EU regarding bike usage (Vandenbulcke et al., 2011). However, The share of cycling is still small compared to countries like the Netherlands, Denmark, and Germany (European Union, 2017; Pucher and Buehler, 2008). Based on the "FOD Mobiliteit en Vervoer" (Federal Public Service, Mobility and Transport) report, in 2014 9.5% of the commuting work trips in Belgium were done by bicycle. This number is 3% in Brussels, 1.5% in Wallonia (French-speaking region) and 14.9% in Flanders (Dutch-speaking region) (FODMobiliteit, 2017). More than half of the Belgians who live within 5 km of their workplace commute by private car while only 19% of them use the bike (Vandenbulcke et al., 2011). So there are many reasons to encourage Belgians to switch from taking the car to cycling.
The remainder of the paper proceeds as follows. Section 2 describes the data collection method and the statistics of the data. In section 3 and 4 we describe the methodology and report and discuss the estimated models. Finally, the fifth section presents a conclusion with the potential policy implications of the study.

2 Data

Desplenter and Meulders (2016) conducted the survey design and the data collection for this study. At first, they grouped all the critical factors that were mentioned in the literature into five categories including socio-demographic characteristics, route and trip characteristics, attitudes toward cycling, general transport costs for cyclists and habits. Then based on a preliminary study, they selected the most important attributes to be included in the survey design. In the preliminary study, 22 people were presented with a list of cycling situation attributes. They were asked to list the five most encouraging and the five most discouraging factors from that list. Based on the results of the preliminary study, it was decided to include the following attributes in the survey design: weather, bike path, road surface, distance, light, slope, and wind. Table 1 shows the attributes used in the survey and their levels. With these seven attributes, a fractional orthogonal design was constructed using the SAS macros (Kuhfeld, 2010). The survey consists of 24 situations in which people were asked if they would take the bike or not. We presented a table to the respondents to describe each situation in the experiment (Figure 1).

The data collection was done through an online survey with people who live in Belgium. A total of 794 people participated in the survey. Persons who did not complete the survey were omitted from the analysis. It was also decided to remove both those who do not own a bike and those for whom the bicycle is the only mean of transport. This was done to consider only those people who can choose between the bike or any other transport mode. It follows that the data analysis is based on the answers of 562 respondents. Table 2 and Table 3 show the distribution of the socio-demographic variables and the variables about the cycling background of the respondents.

Table 1: Trajectory attributes and their levels in the survey design

Attribute	Levels
Weather	Rain Dry and sunny Dry and cloudy
Bike Path	None Separate bike path Marking on the Road
Road Surface	100% asphalt 50% asphalt, 25% clinkers, 25% cobblestones 25% asphalt, 25% clinkers, 25% cobblestones
Distance	15 km 10 km 5 km
Light	Daylight Dark
Slope	Completely flat path Two slopes of 500 meters uphill at 10%
Wind	Powerful (10.8 to 13.8 m/s) Weak winds (1.6 to 3.3 m/s)

Figure 1: An example situation in the online survey

Would you like to take bike for the following route in the following circumstances?

Weather	Rain
Bike Path	Separate bike path
Road Surface	100% asphalt
Distance	5 km
Light	Daylight
Slope	2 slopes of 500 meters uphill at 10%
Wind	Powerful (10.8 to 13.8 m/s)

- Yes
- No

Table 2: Socio-demographic characteristics of respondents

Variable	Levels	Percentage
Gender	Male	43%
	Female	57%
Region of residence	Brussels	4%
	Flanders	65%
	Wallonia	31%
Birthplace	Brussels	4%
	Flanders	66%
	Wallonia	26%
	Abroad	4%
Education	None/Primary	2%
	Secondary	56%
	Bachelor	24%
	Master	19%
Occupation	Student	1%
	Unemployed	1%
	Laborer	1%
	Private company employees	17%
	Government employees	72%
	Self-employed	3%
	Retired	4%
Marital status	Single/Widow	23%
	Married	77%
Have kids	Yes	74%
	No	26%

Table 3: Personal cycling history among the respondents

Variable	Levels	Percentage
Frequency of cycling	Approximately daily	33%
	Only on weekends	8%
	Several times a week	17%
	Several times a month	10%
	Several times a year	23%
	Never	7%
Purpose of cycling	Work	34%
	School	1%
	Recreation	27%
	Shopping	8%
	Relaxation	20%
	Other	10%
Self-assessment of cycling skill	Not cyclist	20%
	Beginner	17%
	Advanced	63%
Using/used bike commuting to school	Yes	64%
	No	36%
Combine bike with other transportation modes	Yes	32%
	No	68%

3 Method

3.1 The latent class model

The response variable, Y_{pi}, is a binary variable which equals 1 if person p indicated that he/she would take the bike in situation i and 0 otherwise. In a first step, a random intercept logistic regression including all the trajectory attributes as fixed effects was estimated. Equation 1 shows the individual linear predictor for the corresponding logistic regression model including an error term (ϵ_{pi}) with a logistic distribution. Equation 2 shows the probability of taking the bike for person p in situation i.

$$\eta_{pi} = \theta_p - \delta_i + \epsilon_{pi} \tag{1}$$

$$\mathrm{P}(Y_{pi} = 1) = \frac{\exp(\theta_p - \delta_i)}{1 + \exp(\theta_p - \delta_i)} \tag{2}$$

The underlying theory of this model comes from the theory of descriptive and explanatory item response models (De Boeck and Wilson, 2004). In Equation 1, θ_p is the tendency to choose the bike in a situation for individual p, and δ_i is the threshold for situation i and is explained using the situation attributes. There are M discrete attributes for each situation and attribute m ($m = 1, ..., M$) has N_m levels indicated as k. Effect coding was used for the discrete attributes, so we define X_{imk} for the level k of attribute m in situation i as: $\sum_{k=1}^{N_m} X_{imk} = 0$. Equation 3 shows the threshold of situation i:

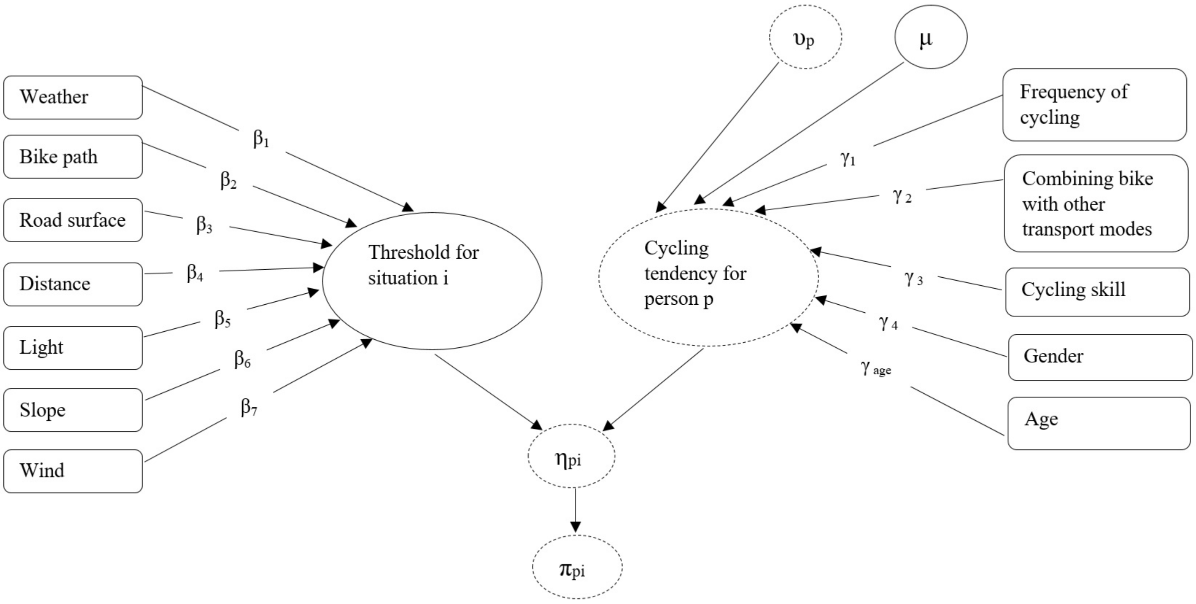
$$\delta_i = \sum_{m=1}^{M} \sum_{k=1}^{N_m - 1} \beta_{mk} X_{imk} \tag{3}$$

First we assume a Normal distribution for θ_p ($\theta_p \sim N(0, \sigma_\theta^2)$) and estimate a mixed logistic regression. Then we use the individual level estimates for θ_p as the response variable in a normal regression model where the independent variables are the individual attributes. We used a stepwise variable selection approach in both direction (Ripley et al., 2013) to select a group of individual attributes with the best AIC value for the normal regression model. Selected individual attributes will be included in the final model. The final model contains L discrete individual attributes and one continuous individual attribute (age) for each person. Attribute l ($l = 1, ..., L$) has N_l levels indicated as j. We will use effect coding for discrete individual attributes and define Z_{plj} for the level j of attribute l for person p as: $\sum_{j=1}^{N_l} Z_{plj} = 0$. Equation 4 shows the cycling tendency for person p:

$$\theta_p = \upsilon_p + \mu + \gamma_{age} Z_p^{age} + \sum_{l=1}^{L} \sum_{j=1}^{N_l - 1} \gamma_{lj} Z_{plj} \qquad \upsilon_p \sim N(0, \sigma_\upsilon^2) \tag{4}$$

A graphical representation of this model is shown in Figure 2. Dotted circles (or ellipses) represent random variables and elements that are a function of these.

Figure 2: Graphical representation of the latent class model



Previous studies have shown that people differ with respect to considering using the bike (Motoaki and Daziano, 2015; Heinen et al., 2011; Li et al., 2013). Classifying people into groups and studying each group separately, has gained much attention in recent years (Dill and McNeil, 2013). We added latent classes to explain the possible existing heterogeneity. In a latent class model, each class will have different parameter estimates. In our model, U_{pq} is a binary latent variable which equals 1 if person p belongs to the latent class q ($q = 1, ..., Q$). We denote the probability that person p belongs to the latent class q by ξ_q:

$$\mathrm{P}(U_{pq} = 1) = \xi_q \qquad \text{with} \quad \sum_q \xi_q = 1 \tag{5}$$

Equation 6 and Equation 7 respectively show the linear predictor and the probability of taking the bike in situation i for person p, given that the person p belongs to class q.

$$\eta_{pi|U_{pq}=1} = \upsilon_p + \mu_q + \gamma_q^{age} Z_p^{age} + \sum_{l=1}^{L} \sum_{j=1}^{N_l-1} \gamma_{qlj} Z_{plj} - \sum_{m=1}^{M} \sum_{k=1}^{N_m-1} \beta_{mkq} X_{imk} \qquad \upsilon_p \sim N(0, \sigma_\upsilon^2) \tag{6}$$

$$\mathrm{P}(Y_{pi} = 1 | U_{pq} = 1, \boldsymbol{\beta}_q, \boldsymbol{\gamma}_q) = \frac{\exp(\eta_{pi|U_{pq}=1})}{1 + \exp(\eta_{pi|U_{pq}=1})} \tag{7}$$

Notice that in Equation 6, the fixed part of the intercept (μ_q) is class dependent, while the random part ($\upsilon_p \sim N(0, \sigma_\upsilon^2)$) is independent of the latent class. The likelihood for the latent class model then becomes:

$$\mathrm{L}(\boldsymbol{Y} | \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\xi}, \sigma_\upsilon^2) = \prod_p \int_{\upsilon_p} \sum_q \xi_q \prod_i \mathrm{P}(Y_{pi} = 1 | U_{pq} = 1, \boldsymbol{\beta}_q, \boldsymbol{\gamma}_q) \phi(0, \sigma_\upsilon^2) d(\upsilon_p) \tag{8}$$

3.2 Accounting for attribute non-attendance

There are two different approaches to study attribute non-attendance: the stated and inferred non-attendance approach (Mariel et al., 2013). In the stated approach the respondents state the ANA rules they employed to make their choice. However, Campbell and Lorimer (2009) showed that the respondents' declarations about their ANA rules might be unreliable. Contrary

to the stated approach, the inferred approach uses analytical methods to infer the rules used by respondents. In this approach, the ANA rules are explored using a latent class model where each class represents a certain non-attendance decision rule (Hess and Rose, 2007; Scarpa et al., 2009; Hensher and Greene, 2010).
We define H_{pm} as a binary latent variable which equals 1 if person p considers attribute m and 0 otherwise. H_{pm} is assumed to have a Bernoulli distribution with probability ζ_m. Then the probability that person p has the ANA pattern $\boldsymbol{H}_p = \boldsymbol{h}_p$ equals (Hole, 2011):

$$\mathrm{P}(\boldsymbol{H}_p = \boldsymbol{h}_p|\boldsymbol{\zeta}) = \prod_{m}^{M}(\zeta_m)^{h_{pm}}(1-\zeta_m)^{1-h_{pm}} \tag{9}$$

Then we have the linear predictor and the probability of taking the bike for person p in situation i as shown in Equations 10 and 11:

$$\eta_{pi|U_{pq}=1,\boldsymbol{h}_p} = \upsilon_p + \mu_q + \gamma_q^{age} Z_p^{age} + \sum_{l=1}^{L}\sum_{j=1}^{N_l-1}\gamma_{qlj}Z_{plj} - \sum_{m=1}^{M} h_{pm}\sum_{k=1}^{N_m-1}\beta_{mkq}X_{imk} \qquad \upsilon_p \sim N(0,\sigma_\upsilon^2) \tag{10}$$

$$\mathrm{P}(Y_{pi}=1|U_{pq}=1, \boldsymbol{H}_p = \boldsymbol{h}_p, \boldsymbol{\beta}_q, \boldsymbol{\gamma}_q) = \frac{\exp(\eta_{pi|U_{pq}=1,\boldsymbol{h}_p})}{1+\exp(\eta_{pi|U_{pq}=1,\boldsymbol{h}_p})} \tag{11}$$

Equation 12 shows the likelihood function for the ANA model:

$$\mathrm{L}(\boldsymbol{Y}|\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\zeta},\sigma_\upsilon^2) = \prod_p \int_{\upsilon_p} \sum_q \xi_q \sum_h \mathrm{P}(\boldsymbol{H}_p = \boldsymbol{h}_p|\boldsymbol{\zeta}) \prod_i \mathrm{P}(Y_{pi}=1|U_{pq}=1, \boldsymbol{H}_p = \boldsymbol{h}_p, \boldsymbol{\beta}_q, \boldsymbol{\gamma}_q)\phi(0,\sigma_\upsilon^2)d(\upsilon_p) \tag{12}$$

4 Results

4.1 Latent class model ignoring attribute non-attendance

Several latent class models containing various numbers of latent classes were estimated using LatentGOLD software (Vermunt and Magidson, 2016). It was decided to select the model with two latent classes because it has a straightforward interpretation and because we noticed that the conclusions about people's attitude toward cycling would not change significantly by having more classes in the model. The results are shown in Table 4. Notice that 66.7% of the sample was classified in class 1 and 33.3% in class 2. Remember that effect coding was used for categorical variables and that coefficients of situation attributes have a negative sign in our linear model predictor, so a larger value for these parameters means a greater barrier for cycling. The last column of the Table 4 shows the p-values for testing if the parameters in the two classes are significantly different. We can reject the $H_0 : \mu_{q=1} \leq \mu_{q=2}$ against the $H_a : \mu_{q=1} > \mu_{q=2}$. This is a one-sided test with p-value: $0.02/2 = 0.01$. Therefore we can say that class 1 are the group of people who on average are more interested in cycling and will be called the "frequent cyclists" in contrast to the "occasional cyclists" in class 2.
In general, the effect of the variables is the same for both classes, but for some variables, there are significantly different. The estimated thresholds for different levels of weather, bike path, light, slope, and wind are significantly different between the two classes. But, for road surface and distance this difference is not significant. The adverse levels of all the situation attributes, except the bike path, have a more negative effect on the probability of cycling in class 2. Almost all of the previous studies have shown that a separate bike path is much more critical for inexperienced cyclists and the absence of separate bike path will not affect experienced cyclists significantly. However, contrary to the expectations, our results showed that the frequent cyclists give more weight to a separate bike path, and the absence of any bike path is a greater barrier for them.

Table 4: Estimation results for the latent class model with 2 classes and a random intercept

Variable	Levels	Class 1	Class 2	p-value: test the difference between classes
Intercept	μ	1.67*	-0.65	0.02
	σ^2	2.17*	2.17*	
Weather	Rain	0.79*	4.26*	0.00
	Dry and sunny	-0.44*	-2.50*	
	Dry and cloudy	-0.35*	-1.76*	
Bike Path	None	0.52*	0.16	0.01
	Separate bike path	-0.54*	-0.31*	
	Mark on the Road	0.02	0.16	
Road Surface	100% asphalt	-0.18*	-0.34*	0.29
	50% asphalt	0.06	0.06	
	25% asphalt	0.11*	0.28*	
Distance	15 km	0.46*	0.64*	0.15
	10 km	0.03	0.09	
	5 km	-0.5*	-0.74*	
Light	Daylight	-0.6*	-0.85*	0.01
	Dark	0.6*	0.85*	
Slope	Completely flat	-0.39*	-0.61*	0.01
	2 slopes of 10%	0.39*	0.61*	
Wind	Powerful wind	0.48*	0.93*	0.00
	Weak wind	-0.48*	-0.93*	
Cycling frequency	Daily	2.08*	1.86*	0.03
	Weekends	-0.03	0.23	
	Weekly	1.15*	-0.17	
	Monthly	0.38	-0.23	
	Yearly	-0.63*	-1.12*	
	Never	-2.96*	-0.56	
Cycling skill	Not cyclist	-0.8*	-0.32	0.21
	Beginner	0.24	-0.5	
	Advanced	0.56*	0.82*	
Combine bike with other transport modes	Yes	-0.47*	0.10	0.06
	No	0.47*	-0.10	
Gender	Male	0.54*	0.24	0.26
	Female	-0.54	0.00	
Age		-0.04*	0.01	0.05
AIC = 10903, BIC = 11102				

*: statistically significant at an alpha level of 0.05

This inconsistency can be explained by the fact that the two classes here are not so different as the experienced cyclists and the inexperienced ones in previous studies. But, they are two groups of cyclists where in general one group cycles more than the other one.
Figures 3 to 7 depict detailed comparisons of how situation attribute levels affect the probability of taking the bike for persons in each class. The results for distance and road surface are not shown in these figures because they do not have significantly different parameters between the two classes. In these figures, the occasional cyclists are shown with dashed-lines and frequent cyclists with full-lines. The minimum tendency to choose the bike in class 1 is -0.796 and the maximum in class 2 is 1.447. The lines are only drawn for those tendency values that occur in the dataset. there are persons in that class who have such tendency to choose the bike. Notice that in Figure 3 the estimates for “dry and sunny”, and “dry and cloudy” in class 1 is almost equal and their lines are overlapped. Also in Figure 4 the estimates for “mark on the road” and “no bike path” in class 2 is equal and so they have overlapping lines.

Figure 3: Probability of taking the bike in different weather conditions

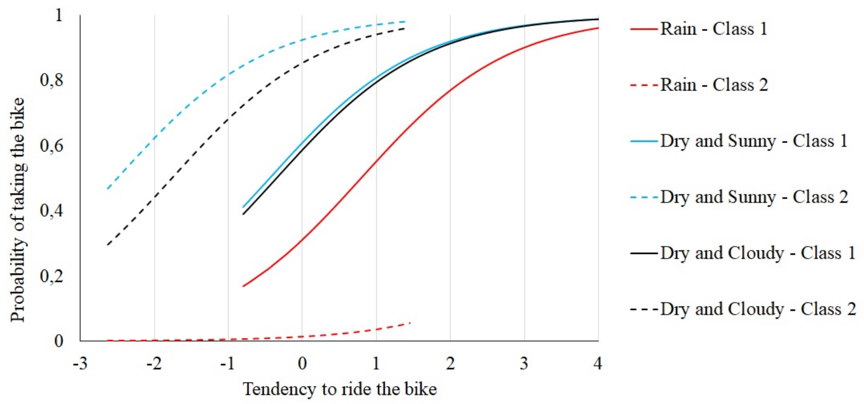


Figure 4: Probability of taking the bike for different types of bike paths

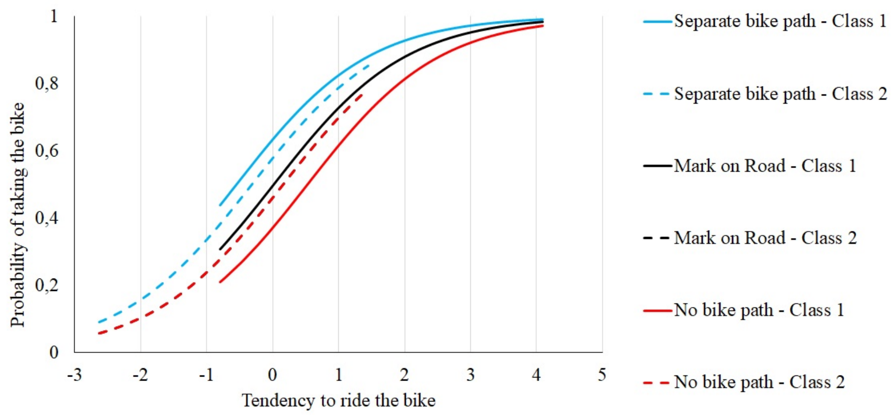


Figure 5: Probability of taking the bike for different levels of light

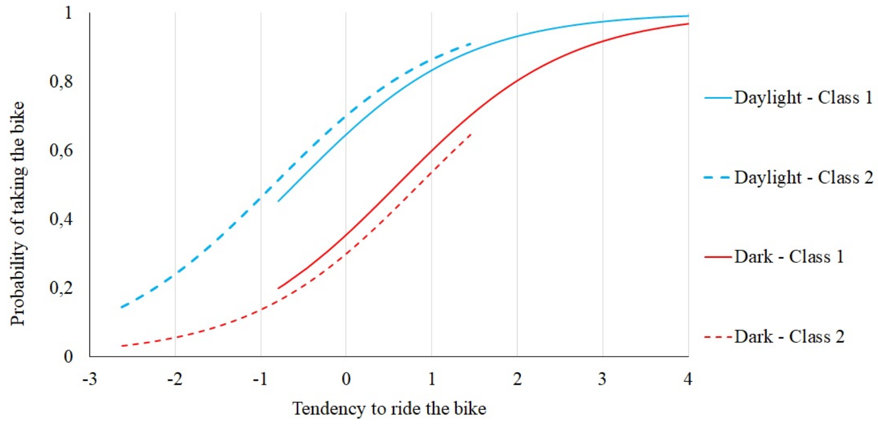


Figure 6: Probability of taking the bike for different levels of slope

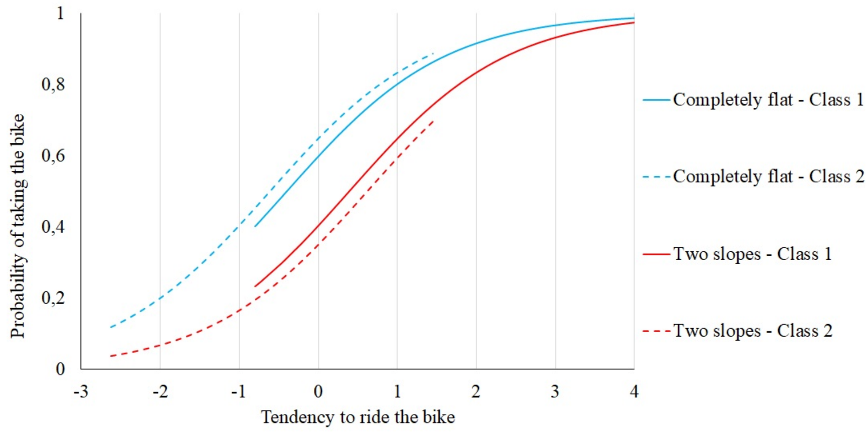
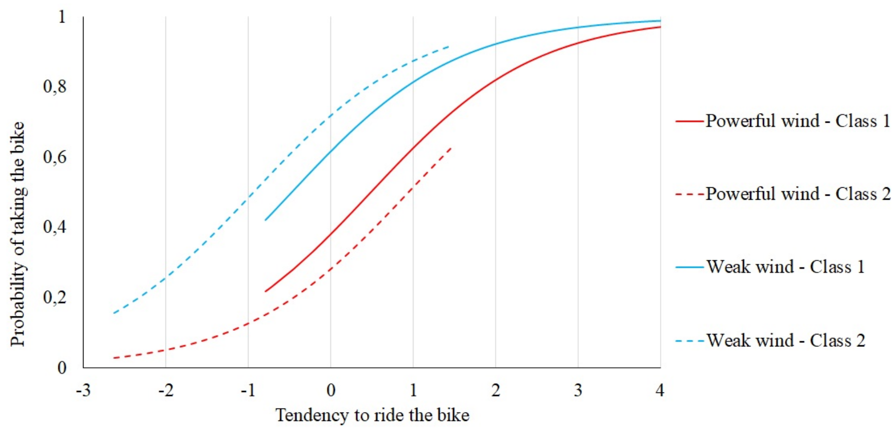


Figure 7: Probability of taking the bike for different levels of wind



4.2 Attribute non-attendance model

We extended the model discussed in our previous analysis to account for ANA. Because of the computation issues in the LatentGOLD, it was not possible to include the individual attributes in this model. So we will continue only using the situation attributes. The AIC value for the ANA model was 9727 and the BIC value was 9870. The AIC and BIC values for our previous model (which did not account for ANA) were 10903 and 11102. So there has been a significant improvement in the model fit after accounting for the ANA in the model showing that not everybody takes all the attributes always into account when deciding on choosing the bike. The estimated latent variable for the attendance of each situation attribute and the probabilities of attendance are shown in Table 5. The estimated latent variables were all significantly different from zero at an alpha level of 0.05.

Table 5: Estimated probabilities of attendance for each situation attribute

	Latent variable	Standard error	Probability of attendance
Weather	0.53*	0.08	0.63
Bike Path	-0.98*	0.08	0.27
Road Surface	-0.95*	0.14	0.28
Distance	-0.44*	0.07	0.39
Light	-0.24*	0.07	0.44
Slope	-0.33*	0.09	0.42
Wind	-0.18*	0.07	0.46

*: statistically significant at an alpha level of 0.05

The weather is taken most into account for choosing the bike. Except for the weather, all other attributes have an attendance probability lower than 0.5. It is striking to notice that only 27% in the sample considered the bike path. Many of the previous studies have introduced the bike path as the most important factor to increase the share of cycling but, our results showed that 73% of the sample ignored the bike path. One possible explanation is once more that the majority of the people in the sample are almost experienced cyclists and the type of the bike path is not their concern. Another possible reason is the impact of respondents' background memory about their own cycling experience. For instance, if someone has always used the bike in a calm and quiet neighborhood without any heavy traffic, his judgment about the bike path can be influenced by his personal experience.

Another possible explanation for the low probabilities of attendance might be the relatively high share respondents who chose the bike in almost every situation. As 20% of the sample selected the bike in more than 20 situations out of the total 24 situations. This means that most of the situation's thresholds were not high enough to overcome their tendency to ride the bike. So they selected the bike almost always without really making a trade-off between situation attributes.

Table 6 shows the estimation result for the ANA model. Class 1 contains 57% of the sample and class 2 has 43%. In Table 6, similar to the model shown in Table 4, class 1 has a positive intercept which is significantly different from the intercept for class 2. Class 2 has a negative intercept which can be treated as the average cycling tendency in class 2. So persons in class 2 are the occasional cyclists and persons in class 1 are the frequent cyclists. There are three major differences with the results that we obtained earlier for the latent class model without ANA. The first difference is in the parameters for the bike path. Unlike before, here the "no bike path" is a stronger barrier for occasional cyclists. But, the separate bike path is more important to the

Table 6: Estimation results for the ANA model

Variable	Levels	Class 1	Class 2	Probability of attendance	p-value: test the difference between classes
Intercept	μ	1.95*	-1.41*		0.00
	σ^2	4.58*	4.58*		
Weather	Rain	1.99*	8.99*	0.63	0.00
	Dry and sunny	-1.25*	-5.43*		
	Dry and cloudy	-0.74*	-3.56*		
Bike Path	None	2.99*	7.76*	0.27	0.00
	Separate bike path	-4.82*	-4.18*		
	Mark on the Road	1.83*	-3.58*		
Road Surface	100% asphalt	-0.73*	-4.68*	0.28	0.00
	50% asphalt	-0.28	2.98*		
	25% asphalt	1.01*	1.70*		
Distance	15 km	2.63*	2.81*	0.39	0.00
	10 km	0.12	1.74*		
	5 km	-2.75*	-4.55*		
Light	Daylight	-2.37*	-3.40*	0.44	0.00
	Dark	2.37*	3.40*		
Slope	Completely flat	-1.32*	-3.43*	0.42	0.00
	2 slopes of 10%	1.32*	3.43*		
Wind	Powerful wind	1.45*	4.03*	0.46	0.00
	Weak wind	-1.45*	-4.03*		
AIC = 9727, BIC=9870					

*: statistically significant at an alpha level of 0.05

frequent cyclists as we also had in our previous analysis. The second difference is that generally, parameters are larger than what we had in the latent class model without ANA. But, we should notice that these parameter estimates hold given that a person attends to all attributes, or $\boldsymbol{h}_p$ is a vector of ones. By averaging over all $\boldsymbol{h}_p$ patterns, we obtain average β parameters that hold for the entire sample. And the third difference is the fact that in the ANA model all the parameters are significantly different between the two classes but in the latent class model, some parameters were not significantly different.

5 Conclusion

This study investigates attribute non-attendance in a binary item response model to explain the propensity of cycling in Belgium, one of the top-ranked countries in Europe regarding bicycle usage. Exploring ANA in a cycling context and with binary item responses had not been conducted before.

The respondents were classified into two groups of frequent and occasional cyclists. There were significant differences between the sensitivity of the two groups to the situation thresholds. In terms of the effect of the weather, uphill slope and darkness our results are in line with previous findings from past research (Motoaki and Daziano, 2015; Heinen et al., 2011; Li et al., 2013). But for the bike path, our results show that frequent cyclists are more concerned about their type of bike path and a separate bike path is more important for them.

We tried to extend our model to account for ANA. This improved the model fit significantly. However, the results for the attendance probabilities were striking. Only weather had a probability of attendance higher than 0.5. Moreover, the probability of attendance for the bike path was only 0.27. One possible explanation is the characteristics of our sample where the majority of the people are experienced cyclists, and 20% of them chose the bike in almost every situation in the survey. Another reason might be the impacts of individuals' background memory about cycling environment on their judgment about each situation.

References

Buehler, R. and J. Dill (2016): "Bikeway Networks: A Review of Effects on Cycling," *Transport Reviews*, 36, 9–27.

Campbell, D. and V. S. Lorimer (2009): "Accommodating attribute processing strategies in stated choice analysis: do respondents do what they say they do," in *Comunicación presentada a la XVII Conferencia anual de la Asociación Europea de Economía Ambiental y de los Recursos Naturales, Amsterdam.*

De Boeck, P. and M. Wilson, eds. (2004): *Explanatory Item Response Models*, Springer-Verlag New York, URL `https://www.springer.com/gp/book/9780387402758`.

Desplenter, S. and M. Meulders (2016): "Determinanten van pendelgedrag per fiets," *M.Sc. Thesis, Faculty of Economics and Business, Campus Brussels, KU Leuven.*

Dill, J. and N. McNeil (2013): "Four Types of Cyclists?" *Transportation Research Record: Journal of the Transportation Research Board*, 2587, 90–99, URL `http://trrjournalonline.trb.org/doi/10.3141/2387-15`.

European Union (2017): *EU TRANSPORT in figures*, European Union, URL `https://publications.europa.eu/en/publication-detail/-/publication/2e046bd0-b542-11e7-837e-01aa75ed71a1/language-en/format-PDF/source-search`.

Fernández-Heredia, Á., A. Monzón, and S. Jara-Díaz (2014): "Understanding cyclists' perceptions, keys for a successful bicycle promotion," *Transportation Research Part A: Policy and Practice*, 63, 1–11.

FODMobiliteit (2017): "Kerncijfers van de Mobiliteit," Technical report, Federale Overheidsdienst Mobiliteit en Vervoer, URL `https://mobilit.belgium.be/nl`.

Handy, S., B. van Wee, and M. Kroesen (2014): "Promoting Cycling for Transport: Research Needs and Challenges," *Transport Reviews*, 34, 4–24, URL `http://dx.doi.org/10.1080/01441647.2013.860204`.

Heinen, E., K. Maat, and B. Van Wee (2011): "The role of attitudes toward characteristics of bicycle commuting on the choice to cycle to work over various distances," *Transportation Research Part D: Transport and Environment*, 16, 102–109, URL `http://dx.doi.org/10.1016/j.trd.2010.08.010`.

Heinen, E., B. van Wee, and K. Maat (2010): "Commuting by bicycle: An overview of the literature," *Transport Reviews*, 30, 59–96.

Hensher, D. A. and W. H. Greene (2010): "Non-attendance and dual processing of common-metric attributes in choice analysis: A latent class specification," *Empirical Economics*, 39, 413–426.

Hess, S. and J. M. Rose (2007): "A latent class approach to modelling heterogeneous information processing strategies in SP studies," in *Oslo workshop on valuation methods in transport planning, Oslo*.

Hole, A. R. (2011): "A discrete choice model with endogenous attribute attendance," *Economics Letters*, 110, 203–205, URL `http://dx.doi.org/10.1016/j.econlet.2010.11.033`.

Kuhfeld, W. (2010): "Marketing research methods in SAS," *SAS Institute Inc.*, 1–1309, URL `http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.163.8176{&}rep=rep1{&}type=pdf`.

Lagarde, M. (2013): "Investigating attribute non-attendance and its consequences in choice experiments with latent class models," *Health Economics*, 1193, 544–567.

Li, Z., W. Wang, C. Yang, and D. R. Ragland (2013): "Bicycle commuting market analysis using attitudinal market segmentation approach," *Transportation Research Part A: Policy and Practice*, 47, 56–68, URL `http://dx.doi.org/10.1016/j.tra.2012.10.017`.

Mariel, P., D. Hoyos, and J. Meyerhoff (2013): "Stated or inferred attribute non-attendance? A simulation approach," in *Economia Agraria y Recursos Naturales*, volume 13, volume 13, 51–67.

Motoaki, Y. and R. A. Daziano (2015): "A hybrid-choice latent-class model for the analysis of the effects of weather on cycling demand," *Transportation Research Part A: Policy and Practice*, 75, 217–230, URL `http://dx.doi.org/10.1016/j.tra.2015.03.017`.

Pucher, J. and R. Buehler (2008): "Making cycling irresistible: Lessons from the Netherlands, Denmark and Germany," *Transport Reviews*, 28, 495–528.

Pucher, J., J. Dill, and S. Handy (2010): "Infrastructure, programs, and policies to increase bicycling: An international review," *Preventive Medicine*, 50, S106–S125, URL `http://dx.doi.org/10.1016/j.ypmed.2009.07.028`.

Ripley, B., B. Venables, D. M. Bates, K. Hornik, A. Gebhardt, D. Firth, and M. B. Ripley (2013): "Package 'mass'," *Cran R*.

Rodríguez, D. A. and J. Joo (2004): "The relationship between non-motorized mode choice and the local physical environment," *Transportation Research Part D: Transport and Environment*, 9, 151–173.

Scarpa, R., T. J. Gilbride, D. Campbell, and D. A. Hensher (2009): "Modelling attribute non-attendance in choice experiments for rural landscape valuation," *European Review of Agricultural Economics*, 36, 151–174.

Vandenbulcke, G., C. Dujardin, I. Thomas, B. de Geus, B. Degraeuwe, R. Meeusen, and L. I. Panis (2011): "Cycle commuting in Belgium: Spatial determinants and 're-cycling' strategies," *Transportation Research Part A: Policy and Practice*, 45, 118–137.

Vermunt, J. K. and J. Magidson (2016): "Technical Guide for Latent GOLD 5.0: Basic, Advanced, and Syntax," *Statistical Innovations Inc.*, 1–161, URL http://www.statisticalinnovations.com/products/LGtechnical.pdf.