Piras, F., Sottile, E., & Meloni, I. (2021). Do psycho-attitudinal factors vary with individuals' cycling frequency? A hybrid ordered modeling approach. Travel behaviour and society, 22, 186-198.

The publisher's version is available at:

http://dx.doi.org/10.1016/j.tbs.2020.09.007

Title Page (with Author Details)

Do psycho-attitudinal factors vary with individuals' cycling frequency? A hybrid ordered modelling approach

Francesco Piras
Corresponding author

University of Cagliari

Via San Giorgio 12, Cagliari, 09124, Italy

Tel: (+39) 070-6756405; Email: francesco.piras@unica.it

Eleonora Sottile

University of Cagliari

Via San Giorgio 12, Cagliari, 09124, Italy

Tel: (+39) 070-6756405; Email: esottile@unica.it

Italo Meloni

University of Cagliari

Via San Giorgio 12, Cagliari, 09124, Italy

Tel: (+39) 070-6756403; E-mail: imeloni@unica.it

AKNOWLEDGEMENTS

This research was funded by the Sardinian Regional Government and Azienda Regionale Sarda Trasporti SpA. The second author gratefully acknowledges financial support provided by the Italian Ministry of University and Research (MIUR), within the Smart Cities framework (Project Netergit, ID: PON04a200490).

Do psycho-attitudinal factors vary with individuals' cycling frequency? A hybrid ordered modeling approach

ABSTRACT

The purpose of the present study was to investigate specifically whether psycho-attitudinal factors could differ for people with different cycling frequency levels and to quantify the determinants influencing the propensity to cycle. To perform our analysis, we developed a hybrid choice modeling approach with a generalized ordered probit choice kernel, using the information collected in 2016 for 2,128 individuals in two mid-size urban areas in Sardinia (Italy). Our results indicate that the latent variables *perception of cycling benefits*, *perception of cycling comfort* and *perceived importance of bike infrastructure* positively influence the propensity to cycle, supporting the idea of a relationship between attitudes and cycling frequency. In addition, the model shows a link between different socio-demographic variables (gender, age, Body Mass Index, education level, number of cars per household, number of household members), built environment characteristics and bike usage. Computation of the pseudo-elasticity effects indicates that strategies focusing only on the physical part of the problem, such as the expansion and improvement of proper infrastructure, might not be sufficient to encourage bike use. At the same time our findings stress the importance of considering people's psychological characteristics when implementing policies aimed at promoting cycling. This can be helpful for identifying, depending on the population segment that is targeted, the most appropriate advertising/information strategy for convincing people to cycle, as well as the most effective marketing messages.

Keywords: cycling behavior, cycling frequency, psycho-attitudinal factors, hybrid choice models, generalized ordered probit

1. INTRODUCTION

In recent years, with a view to reducing car dependence and encouraging the use of more sustainable modes of transport, cycling mobility has been receiving considerable attention in many countries. A growing body of literature has investigated which factors influence the propensity to use the bike (Heinen *et al.,* 2010; Handy *et al*., 2014; Muñoz *et al.*, 2016; Ton *et al.*, 2019). Indeed, the transport sector is a primary cause of the observed deterioration in urban air quality (European Environmental Agency, 2019), road transport having increased significantly. For example, in 2018 in Italy, where the number of cars saw an increase of 4.1% over 2014, air quality worsened, with levels of PM_{10} and $PM_{2.5}$ far higher than the standards set by both the European Union and the World Health Organization (ISFORT, 2018). Not less importantly, transportation also generates several other issues that impact on the environment and urban life, including noise pollution, public health and safety.

One particularly interesting aspect of recent research on cycling mobility is the recognition that, in addition to objective variables (such as socio-demographics, built environment and trip characteristics), psycho-attitudinal factors can contribute to influencing the choice to travel by bike (Ewing and Cervero, 2010; Willis *et al.*, 2015; Muñoz *et al.*, 2016; Arroyo *et al.*, 2020; Gutierrez *et al.*, 2020). This recognition may be attributed to the general finding that people may have different perceptions of cycling infrastructure attributes (Ma *et al.*, 2014). Also, Heinen *et al.* (2011) observed that individuals with similar socio-economic and cycling environment characteristics are likely to show quite different cycling behavior.

Most studies focusing on cycling attitudes and perceptions have examined the difference between cyclists and non-cyclists (Gatersleben and Appleton, 2007). However, it may be useful to distinguish among different types of cyclists (Dill and McNeil, 2013). For instance, one would expect that individuals who are more aware of the benefits associated with cycling, do so more frequently. Some empirical studies have investigated, from a modeling perspective, the association between the effect of psycho-attitudinal variables and cycling frequency. However, several of these have considered the frequency variable as a continuous variable, despite it being measured in ordinal discrete categories, which is inappropriate from an econometric point of view (Bhat *et al.*, 2017). In other works, psycho-attitudinal factors were obtained through some scoring scheme or directly included as exogenous variables in the empirical models, an approach that can potentially lead to measurement errors and result in inconsistent estimates (Walker, 2001).

A much more systematic approach would be advantageous for both policy makers and planners. In fact, the implementation of policies aimed at promoting bike use, such as marketing/information campaigns, can benefit from a deeper understanding of this phenomenon. This in turn would help to avoid misspending financial resources on inadequate measures and to diminish the risk of failures that result in reduced public support (Handy *et al.*, 2014).

In light of these considerations, the aim of our work is to study whether psycho-attitudinal factors vary among people with different frequencies of cycling for any purpose and to explore the effects of socio-demographic and built environment characteristics on the propensity to cycle. In an effort to better understand to what extent psycho-attitudinal variables affect cycling frequency, we estimated a Hybrid

Choice Model (Vij and Walker, 2016) with a generalized ordered probit choice (Greene and Hensher, 2010) kernel. Generalization of the ordered response model renders the thresholds themselves dependent upon both the objective and psycho-attitudinal variables, allowing us to account for systematic heterogeneity across individuals. Further, in our approach, psycho-attitudinal variables are specified as latent variables dependent upon some sociodemographic variables. This approach to modeling the latent variables allowed us to understand the direct and indirect effects of sociodemographic variables on the propensity to cycle. Running different test scenarios, revealed that the probability to cycle with a certain frequency changes following a change in the sociodemographic variables defining the latent variables.

The econometric model is estimated using a dataset collected in the urban areas of Cagliari and Sassari, main cities in Sardinia (Italy), where, despite the implementation of policies encouraging bike use, cycling levels are still low. The model clearly provides evidence that individuals with a greater perception of cycling benefits and comfort and who attach greater importance to the presence of proper infrastructure, cycle more. In addition, the model reveals a link between different socio-demographic variables, built environment characteristics and cycling frequency.

The remainder of the paper is structured as follows. In section 2 we provide a literature review to set the context of the current investigation. Section 3 presents the exploratory data analysis. Next, in sections 4 and 5, we illustrate the methodological approach used to perform our analysis and discuss the model estimation results. Finally, section 6 provides conclusions and identifies the study's limitations.

2. CONTEXT OF CURRENT INVESTIGATION

The relationship between psycho-attitudinal factors and cycling frequency has been reported in different studies. Some works investigated how individuals' propensity to cycle is influenced by their perceptions of traffic risks and cycling facilities. Gatersleben and Appleton (2007) investigated attitudes and perceptions of people in different behavioral stages of change of the Transtheoretical Model in relation to cycling to work. They found that as the level of cycling experience increases so their perceptions of various personal and external barriers change, and they are more likely to recognize the benefits of cycling. Sener *et al.* (2009) showed that the perception of the quality of cycling facilities significantly affects commute cycling frequency, but this variable was not relevant for non-commuting. Handy *et al.* (2010) indicated that people with higher levels of cycling comfort are more likely to cycle regularly for transportation than non-cyclists and non-regular cyclists. Ma *et al.* (2014) tried to disentangle the effect of objective built environment and perception thereof. They showed that the perception of the cycling environment has a positive and direct impact on frequency, while the direct effect of the objective environment became insignificant when including the perception effect. Regarding the perception of safety, Sallis *et al.* (2013) found that this construct correlated with cycling frequency, and individuals would have cycled more had it been safer than the car. Recently, Kelarestaghi *et al.* (2019) indicated that the latent risk factor, related to indicators such as theft and road safety, has a negative effect on cycle-to-campus frequency for university students in the Maryland Metropolitan Area.

Other works have focused on the effects of cycling attitudes, which refer to how people generally view bike use. Heinen *et al.* (2011) studied differences in attitudes and norms between full-time and part-time commuter cyclists, for different distance categories. They found that habits, subjective norms and the recognition of cycling benefits positively influence the likelihood of cycling full-time to work. Further, they showed that individuals who commute over longer distances have, on average, a more positive attitude towards cycling. A study by Swiers *et al.* (2017) of a sample of University students in the UK found that regular cyclists (daily/weekly) were significantly more likely to perceive health benefits as a motivator than monthly/annual cyclists. Interestingly, Kroesen *et al.* (2017), using a panel dataset gathered in the Netherlands, indicated that cycling behavior and attitudes mutually influence each other over time, and the effect of cycling on attitudes is stronger than *vice versa*. Kaplan *et al.* (2019) found that cycling has the potential for making people feel better about themselves from a physical and psychological perspective.

Recently some papers have introduced into their analysis the notion of experience, that represents a concept closely related with frequency of use. A person who rides frequently can be considered an experienced cyclist who is likely to feel more confident (Namgung and Jun, 2019). The same authors examined attitudes towards cycling of cyclists with different experience levels among a sample of students and staff members at Ohio University. They showed that the more experienced group of cyclists exhibited more positive attitudes, while users with less experience are more likely to perceive barriers to cycling. Thigpen (2019) explored whether attending a bike-friendly university, like UC Davies, led to high levels of cycling and a change in cycling attitudes, and to what extent changes are influenced by personal cycling experience. He found that riding a bike at any point during college increases both pro-bike attitudes and cycling skills, while exposure to high levels of cycling appears not to influence attitudes or skills.

From a methodological viewpoint, different approaches have been used to explore the link between exogenous variables and cycling frequency (see Table 1). Some studies have employed linear regression models for their analysis, the dependent variable being the number of trips made by bike in a certain period of time (Sallis *et al.*, 2013, Stinson *et al.*, 2014) or the weekly miles of transportation and recreation cycling (Xing *et al.*, 2010). Other studies have used econometric models including multinomial logit, ordered logit and ordered probit. Stinson and Bhat (2004) were the first to estimate an ordered logit model to investigate which factors influence frequency of bike use for commute to and from work. Similarly, Noland *et al.* (2010), estimating an ordered probit model, investigated which factors are associated with cycling behavior in New Jersey. Manaugh *et al.* (2017) categorized cycling frequency into four different classes of *never*, *rarely*, *usually*, and *always* and used a multinomial logit to run their analysis. Interestingly, Bhat *et al.* (2017) proposed a new spatial generalized ordered response model with skew-normal kernel error terms and applied it to the analysis of cycling frequency. Recently, Oliva *et al.* (2019) estimated a latent class ordered logit to segment neighborhoods, according to the cycling behavior observed in them, as a function of built environment attributes.

The first to include psycho-attitudinal characteristics in their modeling framework were Sener *et al.* (2009). They estimated two panel ordered probit models for analyzing the frequency of bike commuting and non-commuting. In a similar vein, Fu and Farber (2017) used an ordered probit model to study the influence of cycling safety, benefits and comfort on cycling frequency for commuting. Handy *et al.* (2010) estimated a nested logit to examine the objective and subjective factors influencing bike ownership, use and frequency. Heinen *et al.* (2011) and Namgung and Jun (2019) employed multiple binary logit models to explore the difference in attitudes between cyclists and non-cyclists, and between experienced and inexperienced ones. One limitation of the studies described above is the use of a two-stage sequential approach for their model estimation. Typically, they first obtain the value of psycho-attitudinal factors, identified by an exploratory factor analysis, through some scoring scheme, for example Principal Component Regression (PCR), and then include them in the estimation phase of the discrete choice model. This methodology may result in biased estimators for the parameters involved (Walker, 2001) or in estimators with a statistical significance greater than their real contribution to the model (Raveau *et al.*, 2010).

The other main approach used in the field is structural equation modeling (SEM). In structural equation modeling psycho-attitudinal variables, specified as a linear combination of observed variables, are not directly observed from individuals but are considered as functions of original statement variables. Examples of research employing SEM include Ma *et al.* (2014), Kelarestaghi *et al*. (2019) and Zhang *et al.* (2019). The main issue with these studies is that they treat the frequency of cycling as a continuous variable, despite being measured as an ordered variable, which is improper from an econometric point of view as they assume that the numerical distance between each set of categories is the same.

Table 1. Methodology used for modeling cycling frequency

Authors	Year	Include psycho-attitudinal factors	Methodology	Location
Stinson and Bhat	2004	No	Standard ordered logit	USA
Sener *et al.*	2009	Yes	Standard ordered probit	USA
Handy *et al.*	2010	Yes	Nested logit	USA
Xing *et al.*	2010	Yes	Linear regression	USA
Heinen *et al.*	2011	Yes	Binary logit	Netherlands
Noland *et al.*	2011	No	Standard ordered logit	USA
Sallis *et al.*	2013	Yes	Linear regression	USA
Ma *et al.*	2014	Yes	Structural equation modeling	USA
Stinson *et al.*	2014	No	Linear regression	USA
Bhat *et al.*	2017	No	Spatial generalized ordered probit	USA
Fu and Farber	2017	Yes	Standard ordered probit	USA
Manaugh *et al.*	2017	No	Multinomial logit	USA
Kelarestaghi *et al.*	2019	Yes	Structural equation modeling	USA
Namgung and Jun	2019	Yes	Binary logit	USA
Oliva *et al.*	2019	No	Latent class ordered logit	Chile
Zhang *et al.*	2019	Yes	Structural equation modeling	China

The modeling approach adopted in the current paper is distinct from earlier works for three important aspects. First, the use of a Hybrid Choice Model allowed us to include psycho-attitudinal variables in a discrete choice analysis, in which the choice model and the latent variable model are integrated into a single structure that is estimated simultaneously. The simultaneous estimation yields consistent and efficient estimates of the parameters (Walker, 2001). Further, with this approach we are able to correlate psycho-attitudinal variables with observed socio-economic variables. This feature is particularly relevant because it can help to identify which marketing/information strategy is more effective in convincing people to use the bike along with which messages to deliver to different segments of the population. For example, if the study finds that women are less likely to cycle because they have a greater perception of the limitations of the bike as a means of transport than men, then a marketing campaign promoting bike use could be adapted to target women. For example, a video campaign showing a number of women who commute or go shopping by bike or using as testimonials celebrities who cycle. Second, the estimation of an ordered choice model overcomes the limitations of structural equations modeling (SEM) in handling ordered variables. In fact, it has been demonstrated that using an ordinal approach, instead of a continuous numerical approach, yields more accurate results (Allen *et al.*, 2018). Third, the use of a generalized ordered probit as kernel relaxes the thresholds making them the same for all individuals. The standard ordered logit/probit assumes that the threshold values are fixed across observations, which might be not appropriate (Eluru *et al.*, 2008; Balusu *et* *al.*, 2018). In fact, imposing this restriction might produce inconsistent latent propensity and threshold values, and consequently inconsistent effects of the independent variables on the likelihood of different categories of frequency (Eluru *et al.*, 2008).

3. DATA ANALYSIS

The data used in this study come from a survey conducted by the Regional Government of Sardinia and the Research Centre for Mobility Models (CRiMM) at the University of Cagliari (Italy) in the urban areas of Cagliari and Sassari, main cities in Sardinia (Italy). The survey, called "BIKE I LIKE YOU", was carried out between 2014 and 2016 and targeted local authority employees. We intercepted potential respondents both via mailing lists and through a promotional campaign. The mailing lists were provided by the universities of Cagliari and Sassari, the Regional Government of Sardinia and the municipalities of Cagliari and Sassari (around 9,600 invitation mails were sent). The promotional campaign was conducted via traditional communication channels and social media, inviting people to complete an on-line questionnaire using the WUFOO survey platform (for more details see Sottile *et al.*, 2019). In particular, the questionnaire was organized into 4 sections:

1. Bike use section aimed to identify for what purpose and how frequently people choose to cycle.
2. Cycling perceptions section (Likert scale from 1 to 5, 1=Totally disagree to 5 = Totally agree, intended to:
    - Measure positive and negative aspects of cycling in general.
    - Measure the perception of safety of bike lanes and paths.
    - Measure the perceived importance of context characteristics, intended as the importance assigned to policies for increasing bike use.
3. Description of home-work commute trip.
4. Socio-demographic information section.

A total of 4,691 individuals completed the survey. However, after careful screening - excluding records with incomplete socio-economic information and residence location - the final sample size included 2,128 individuals (corresponding to 45.4% of respondents).

Regarding individual characteristics, the analysis revealed that the sample is almost equally divided between males and females (48.4% *vs* 51.6%). 73.3% of those surveyed are aged between 41 and 60. The majority of individuals are highly-educated (57.7% have a bachelor's degree or higher). Average number of household members is 2.88. As for personal monthly income, 6.6 % stated they earned less than € 1,000 a month, 64.9% € 1,001-2,000, 9.6% € 2,001-3,000, 14.1% >€ 3,000. Note that, in terms of socio-economic characteristics, the sample is not representative of the entire working population in Sardinia vis-a-vis gender (for the share of male individuals we have 48.4% in our sample *vs* 58.0%), age (for the age category *41-60* we have 73.3% in our sample *vs* 58.3% in Sardinia) and the level of education (only 18.9% of workers in Sardinia have a bachelor's degree). However, the sample can be considered as representative of employees in the service sector, which in Sardinia accounts for 58.7% of the entire working population (Sardegna Statistiche, 2018).

Participants were asked to report their cycling frequency for the three different purposes -commuting, errands and leisure/sport - in five ordinal categories from “I never cycle” to “I cycle every day” (the same categorization has been used in other works, *e.g.* Noland *et al.*, 2010; Bhat *et al.*, 2017). Analysis of the questionnaire revealed the following share by frequency of people using the bike: (1) I never use the bike (50.1% of the sample); (2) I use the bike 1-10 times per year (14.6% of the sample); (3) I use the bike 1-5 times per month (14.2% of the sample); (4) I use the bike more than once a week (14.7% of the sample); (5) I use the bike every day (6.5% of the sample). Note that the trips considered here are not restricted to utilitarian cycling alone but include recreational cycling as well.

Table 2. Socio-demographic characteristics

Variables	N.	[%]	AVG.
Total sample	2,128		
Gender			
Male	1,029	48.4%	-
Female	1,099	51.6%	-
Age	-	-	48.02
Age 18-30	82	3.9%	-
Age 31-40	341	16.0%	-
Age 41-60	1,559	73.3%	-
Age > 60	146	6.9%	-
Level of education			
Low (High school and lower)	901	42.3%	-
Medium (Graduate)	738	34.7%	-
High (Higher than Master's degree)	489	23.0%	-
Marital status			
Married	1,550	72.8%	-
Not married	578	27.2%	-
Presence of children in the household			
Yes	1,159	54.5%	-
No	969	45.5%	-
# of household members	-	-	2.88
Driving license			
Yes	2,098	98.6%	-
No	30	1.4%	-
Personal car available			
Yes	1,930	90.7%	-
No	198	9.3%	-
# of cars per household	-	-	1.72
# of bikes per household	-	-	1.54
Personal income per month			
Income 0-1,000 €	140	6.6%	-
Income 1,001-2,000 €	1,382	64.9%	-
Income 2,001-3,000 €	205	9.6%	-
Income >3,000 €	301	14.1%	-
Cycling frequency			
Never	1,065	50.1%	-
1-10 times per year	310	14.6%	-
1-5 times per month	303	14.2%	-
More than once a week	311	14.6%	-
Every day	139	6.5%	-

3.1. Comparisons: socio-demographic characteristics for different levels of cycling

Table 3 gives the socio-demographic variables of respondents for different levels of cycling. Some differences were detected among different categories of cycling frequency. Males and low-educated individuals tend to use the bike more frequently. The most interesting difference concerns car ownership and bike ownership per household. Indeed, bike users tend to own more bikes and less cars in their households than individuals who do not choose to cycle.

Table 3. Socio-economic characteristics for different levels of cycling

	Frequency of bike use									
	Never		*1-10 times per year*		*1-5 times per month*		*More than once a week*		*Every day*	
	N.	*%*	*N.*	*%*	*N.*	*%*	*N.*	*%*	*N.*	*%*
Total	1,065	50.1%	310	14.6%	303	14.2%	311	14.6%	139	6.5%
Gender										
Male	423	39.7%	128	41.3%	159	52.5%	212	68.2%	107	77.0%
Female	642	60.3%	182	58.7%	144	47.5%	99	31.8%	32	23.0%
Age (average)	49.07	-	46.13	-	46.49	-	47.46	-	48.82	-
18-30	24	2.3%	20	6.5%	14	4.6%	19	6.1%	5	3.6%
31-40	164	15.4%	59	19.0%	53	17.5%	44	14.1%	21	15.1%
41-60	786	73.8%	215	69.4%	223	73.6%	232	74.6%	103	74.1%
>60	91	8.5%	16	5.2%	13	4.3%	16	5.1%	10	7.2%
Body Mass Index (average)	23.82	-	23.06	-	23.19	-	23.75	-	23.93	-
Level of education										
Low (High school and lower)	451	42.3%	107	34.5%	94	31.0%	172	55.3%	77	55.4%
Medium (Graduate)	381	35.8%	110	35.5%	125	41.3%	85	27.3%	37	26.6%
High (Higher than Master's degree)	233	21.9%	93	30.0%	84	27.7%	54	17.4%	25	18.0%
Presence of children in the household										
Yes	581	54.6%	170	54.8%	163	53.8%	171	55.0%	74	53.2%
No	484	45.4%	140	45.2%	140	46.2%	140	45.0%	65	46.8%
Marital status										
Married	757	71.1%	230	74.2%	224	73.9%	236	75.9%	103	74.1%
Not married	308	28.9%	80	25.8%	79	26.1%	75	24.1%	36	25.9%
# of household members (average)	2.83	-	2.99	-	2.88	-	2.96	-	2.81	-
Driving License										
Yes	1,049	98.5%	308	99.4%	300	99.0%	305	98.1%	136	97.8%
No	16	1.5%	2	0.6%	3	1.0%	6	1.9%	3	2.2%
Personal car available										
Yes	978	91.8%	276	89.0%	274	90.4%	275	88.4%	127	91.4%
No	87	8.2%	34	11.0%	29	9.6%	36	11.6%	12	8.6%
# of cars (average)	1.71	-	1.82	-	1.72	-	1.73	-	1.52	-
# of bikes (average)	0.97	-	1.98	-	2.07	-	2.19	-	2.32	-
Monthly personal income										
Income 0-1,000 €	52	4.9%	37	11.9%	16	5.3%	22	7.1%	13	9.4%
Income 1,001-2,000 €	706	66.3%	183	59.0%	205	67.7%	206	66.2%	82	59.0%
Income 2,001-3,000 €	148	13.9%	46	14.8%	41	13.5%	50	16.1%	20	14.4%
Income >3,000 €	159	14.9%	44	14.2%	41	13.5%	33	10.6%	24	17.3%

One important issue concerns the association between built environment and cycling behavior (Cervero and Duncan, 2003; Wang *et al.*, 2016; Yang *et al.*, 2019). In particular one key factor in the choice to use the bike is that in many cases built-environment may represent a barrier to cycling. To analyze this correlation, the micro-environments of the place of residence (presence of bike lanes and percentage of green areas) were assessed within the GIS environment using a buffer of 400 m radius. Using the digital land use maps downloaded from the Sardinian Government website (<http://www.sardegnageoportale.it/areetematiche/databasegeotopografico>), it was possible to calculate the characteristics of the residence location (urban or suburban/rural). Table 4 provides a summary of respondents' built environment characteristics.

However, as can be seen Table 4 almost no differences were found among cycling frequency categories since the majority of individuals live in urban areas and fewer than half have access to a bike lane within 400m from home.

Table 4. Built environment characteristics

Variables	**Cycling frequency**					
	Never	*1-10 times per year*	*1-5 times per month*	*More than once a week*	*Every day*	***Total***
Total	1,065	310	303	311	139	2,128
Residential location						
Urban	842 (79.1%)	224 (72.2%)	225 (74.3%)	247 (79.6%)	116 (83.0%)	1,654
Suburban and rural	223 (20.9%)	86 (27.8%)	78 (25.7%)	64 (20.4%)	23 (17.0%)	474
Presence of bike paths within 400m from home						
Yes	519 (48.7%)	151 (48.7%)	137 (45.1%)	146 (47.1%)	82 (59.2%)	1,035
No	546 (51.3%)	159 (51.7%)	166 (54.9%)	165 (52.9%)	57 (40.8%)	1,093
Average % of green areas within 400m from home per individual	5.2%	5.3%	5.1%	4.9%	5.1%	n/a
Average distance from home to the nearest bus stop [m]	260 m	258 m	268 m	300 m	237 m	n/a

n/a not applicable

3.2. Analysis of psycho-attitudinal characteristics

Psycho-attitudinal characteristics were measured by means of the questions with the 5-point Likert scale (1 = Totally disagree to 5 = Totally agree). As explained in paragraph 3, the survey included questions aimed at measuring three latent variables:

1. perception of the bike as a means of transport (items A1-A12);
2. perception of safety of bike lanes and paths (items B1-B4);
3. perceived importance of context characteristics (items C1-C8).

To detect any differences in psycho-attitudinal variables among cycling frequency groups, we computed the means of each item for all the subsamples and then tested whether subsample means were statistically different conducting a z-test.

Table 5 shows the results of the statistical test analysis for individuals with different levels of cycling experience with regard to the attitudinal factors. To avoid clutter, Table 5 only shows the results of the test conducted between two consecutive groups (*e.g.* "never" group *vs* "1-10 times per year" group).

There are items that are not statistically significantly different among categories. For example, no differences were detected for the item regarding the benefits of cycling in terms of cost, accessibility and reduced level of pollution, the majority of the sample recognizing the positive aspects of cycling. Interestingly, there were no significant differences between the groups as far as the implementation of certain measures was concerned, such as a greater extension of limited traffic zones and presence of end-of-trip facilities. The latter result suggests that all the respondents consider the existence of such facilities important, regardless of how frequently they cycle.

However, in general, the z-test analyses revealed that more experienced cyclists have more positive perceptions of cycling than the less experienced, as found in other works (Heinen *et al.*, 2011; Namgung and Jun, 2019). More specifically, frequent cyclists have a greater *perception of cycling comfort* than occasional cyclists and non-cyclists (items A8, A10 and A12). We also found differences in the perception of safety (items A2 and B1). In particular, the results suggest that more experienced cyclists tend to be less bothered by mixed traffic situations. Nevertheless, it should be noted that, in general, cyclists and non-cyclists agree about the inadequacy, in terms of safety, of the current cycling network. The z-test analyses also showed that the willingness to cycle in the case of integration with the public transit service (items C4 and C7) is greater among non-cyclists and occasional cyclists.

Table 5. Psycho-attitudinal characteristics for different levels of cycling (0 = never, 1=1-10 times per year, 2=1-5 times per month, 3= more than once a week, 4 = every day).

	Avg tot	Avg 1	Avg 2	Avg 3	Avg 4	Avg 5	**Z-stat 1-2**	**Z-stat 2-3**	**Z-stat 3-4**	**Z-stat 4-5**
PERCEPTION OF THE BIKE AS A MEANS OF TRANSPORT										
A1. It is a rapid means of transport	4.12	3.99	3.90	4.15	4.46	4.83	1.12	-2.79**	-3.72**	-5.31**
A2. Cycling in traffic is not dangerous	1.75	1.72	1.56	1.75	1.86	2.22	2.47**	-2.38**	-1.32	-2.98**
A3. The bike is not likely to be stolen and there are no adequate parking areas	2.22	2.32	2.25	2.12	2.03	2.11	0.80	1.35	0.86	-0.60
A4. It is not expensive	4.69	4.66	4.65	4.72	4.73	4.84	0.16	-1.01	-0.26	-1.78*
A5. It does not imply exposure to bad weather and air pollution	2.44	2.24	2.49	2.65	2.69	2.78	-3.31**	-1.79*	-0.41	-0.64
A6. It avoids wasting time looking for parking	4.38	4.30	4.24	4.46	4.55	4.78	0.79	-2.61**	-1.24	-2.70**
A7. It is healthy	4.69	4.57	4.67	4.85	4.85	4.86	-1.93*	-3.53**	-0.04	-0.16
A8. It is easy to carry heavy items	1.88	1.70	1.83	2.01	2.18	2.41	-1.91*	-1.99**	-1.81*	-1.91*
A9. It allows one to appreciate historic centers and increases accessibility to city services	4.36	4.31	4.25	4.38	4.50	4.63	0.96	-1.54	-1.59	-1.50
A10. No need for cycling gear	2.96	2.83	2.83	2.95	3.18	3.70	0.10	-1.35	-2.36**	-3.86**
A11. It contributes to reducing polluting emissions	4.84	4.83	4.79	4.86	4.86	4.93	0.78	-1.31	0.04	-1.78*
A12. It does not hamper daily activity patterns	3.09	2.89	3.08	3.08	3.39	3.91	-2.35**	-0.05	-3.18**	-4.24**
PERCEPTION OF SAFETY										
B1. Existing bike lanes are not useful for traveling	3.41	3.57	3.42	3.22	3.14	3.16	1.73*	1.77*	0.73	-0.10
B2. Existing bike lanes and crossings are safe, comfortable and well-marked	2.11	2.02	2.16	2.21	2.22	2.16	-1.98**	-0.55	-0.07	0.47
B3. It is better to ride in traffic than use the existing bike paths	1.95	1.97	1.86	1.81	1.98	2.17	1.29	0.56	-1.76*	-1.34
B4. Car drivers do not respect dedicated bike lanes and often invade them	4.00	3.88	3.93	4.13	4.19	4.30	-0.63	-2.08**	-0.78	-0.95
PERCEIVED IMPORTANCE OF THE CONTEXT. I would cycle more with:										
C1. An extensive network of dedicated bike lanes in urban area	4.39	4.23	4.44	4.61	4.61	4.61	-3.19**	-2.51**	0.06	0.01
C2. The presence of racks and secure parking for bikes	4.26	4.15	4.23	4.42	4.43	4.46	-1.11	-2.49**	-0.16	-0.26
C3. A greater extension of the LTZ or pedestrian zones	3.77	3.67	3.78	3.81	3.94	4.12	-1.28	-0.28	-1.36	-1.50
C4. A bike-sharing station close to home or at public transport stops	3.73	3.87	3.72	3.53	3.57	3.42	1.71*	1.71*	-0.30	1.01
C5. If other people use it	2.95	2.99	2.83	2.75	3.08	3.13	1.79*	0.69	-2.81**	-0.29
C6. Dedicated services at work / study (parking, showers, lockers for equipment, etc.)	3.95	3.86	3.93	4.04	4.17	4.01	-0.91	-1.25	-1.32	1.19
C7. An integrated ticket for bike-sharing and public transport services	3.84	3.92	3.57	3.81	3.90	3.78	4.21**	-2.22**	-0.85	0.90
C8. A Combination with public transport services	3.97	3.98	3.72	3.97	4.15	4.04	3.19**	-2.50**	-1.76*	0.88
C9. An Increase of car parking fees	2.27	2.24	2.20	2.16	2.40	2.53	0.50	0.36	-2.17**	-0.87

* Significant at 90% confidence. ** Significant at 95% confidence.

3.3. Factor analysis

Prior to the modeling phase, an explorative factor analysis (Bollen, 1989) was performed to identify the latent constructs underpinning the set of our attitudinal statements. The factor loadings are estimated using Principal Axis Factoring (PAF) with varimax rotation. To establish if the dataset is suitable for exploratory factor analysis, sample adequacy and strength of the intercorrelation of items must be examined. The Kaiser-Meyer-Olkin (KMO) measure is used for sample adequacy: KMO values between 0.7 and 1 indicate the sampling is adequate. The Bartlett test of sphericity is used to test the hypothesis that the correlation matrix is an identity matrix, which would indicate that variables are unrelated and therefore unsuitable for structure detection. Furthermore, to examine reliability Cronbach's alpha value is used. A Cronbach's alpha value higher than 0.6 indicates that the dataset is reliable and acceptable.

Only two of the three psycho-attitudinal constructs turned out to be suitable for factor analysis, *perception of the bike as a means of transport* (items A1-A12) and *perceived importance of the context* (items C1-C9), whose KMO was 0.765 and 0.806 respectively. Instead, for the construct *perception of safety of bike lanes and paths* (items B1-B4) we obtained a KMO of 0.572, which is below the reliability threshold of 0.7. Thus, we proceeded with factor analysis only for the two constructs with KMO higher than 0.7.

Factor analysis generated two factors for the latent construct *perception of the bike as a means of transport* (factor LV1 and factor LV2) and one factor for the latent construct *perceived importance of the context* (factor LV3). Table 6 gives the results of factor analysis showing the loadings of the survey items on each of the three identified factors.

The latent variable LV1, *Perception of cycling benefits,* expresses the agreement related to generally recognized positive features of bikes, while the LV2, *Perception of cycling comfort,* expresses the agreement related to generally recognized negative features of bikes (exposure to bad weather, carrying heavy items, limitations in daily activity patterns, fatigue). The latent variable LV3, *Perceived importance of bike infrastructure,* uses indicators capturing the appeal of some factors that would facilitate bike use. Most of the Cronbach's alpha values are above 0.6, except for LV2 that is just acceptable since it is around the "criterion-in-use" of 0.6.

Table 6. Factor scores of the psycho-attitudinal factors towards the bike mode (values below 0.4 are not reported)

Factor	Variables	Loading	Cronbach's alpha
LV1	A1. It is a rapid means of transport (avoids queues and traffic)	0.582	0.690
	A4. It is not expensive	0.540	
	A6. It avoids wasting time looking for parking	0.594	
	A7. It is healthy	0.744	
	A9. It allows one to appreciate historic centers and increases accessibility to city services	0.699	
	A11. It contributes to reducing polluting emissions	0.621	
LV2	A5. It does not imply exposure to bad weather and air pollution	0.554	0.597
	A8. It is easy to carry heavy items	0.629	
	A10. No need for cycling gear	0.629	
	A12. It does not hamper daily activity patterns	0.693	
LV3	C1. An extensive network of dedicated bike lanes in urban area	0.907	0.778
	C2. The presence of racks and secure parking for bikes	0.861	
	C3. A greater extension of the LTZ or pedestrian zones	0.721	

4. METHODOLOGICAL FRAMEWORK

To perform our analysis we employ a hybrid choice modeling (HCM) approach (Vij and Walker, 2016) with a generalized ordered probit choice (Greene and Hensher, 2010) kernel (Figure 1). The HCM provides a framework for incorporating psycho-attitudinal variables into our model of the decision-making process.

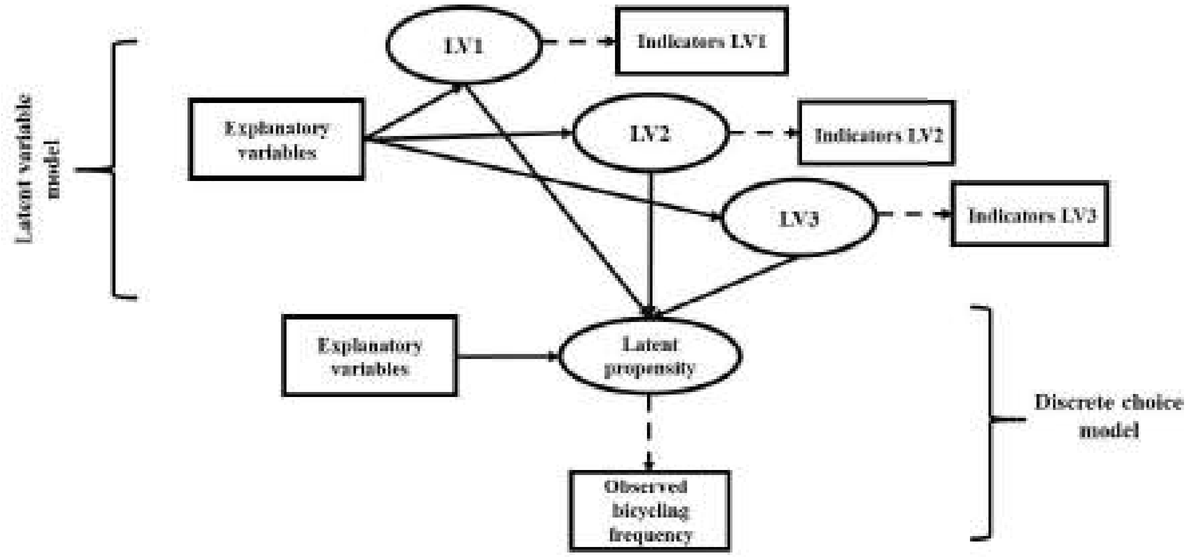

Figure 1. Conceptual model

Following the framework of hybrid choice models, each latent variable is assumed to be determined by a structural equation, in turn assumed to be a linear function of respondents' individual and household characteristics. Hence, for the person q we get:

$$\boldsymbol{LV}_q = \boldsymbol{\kappa} + \boldsymbol{\lambda}\boldsymbol{x}_q + \boldsymbol{\omega}_q, \boldsymbol{\omega}_q \sim N(0, \Sigma_\omega\ diagonal) \quad (1)$$

where $\boldsymbol{LV}_q$ is the $(N \times 1)$ vector of latent variables, $\boldsymbol{\kappa}$ is the $(N \times 1)$ vector of constants, $\boldsymbol{x}_q$ is the $(M \times 1)$ vector of individual background characteristics; $\boldsymbol{\lambda}$ is the $(N \times M)$ matrix of coefficients associated with these characteristics, and $\boldsymbol{\omega}_q$ is a normal distributed error term with a diagonal covariance matrix Σ_ω.

The R items reported in Table 6 are used as indicators of latent variables and are linked to them through measurement equations. As in the typical HCM theory, indicators are modeled using the following set of equations:

$$\boldsymbol{I}_q^* = \boldsymbol{\delta} + \boldsymbol{\zeta}\boldsymbol{LV}_q + \boldsymbol{\upsilon}_q,\ \ \boldsymbol{\upsilon}_q \sim Logistic(0, \Sigma_\upsilon\ diagonal) \quad (2)$$

$$I_{rq} = \begin{cases} 1 & if\ \rho_{0r} < I_{rq}^* \ \leq \rho_{1r} \\ 2 & if\ \rho_{1r} < I_{rq}^* \ \leq \rho_{2r} \\ \vdots & \\ S_r & if\ \rho_{S_r - 1} < I_{rq}^* \ \leq \rho_{S_r} \end{cases} \quad (3)$$

Eqs. (2), (3) represent a system of ordered logit models for measuring $\boldsymbol{LV}_q$, where $\boldsymbol{I}_q^*$ is the $(R \times 1)$ vector of continuous measurement indicators of the latent variables, with elements $I_{rq}^* = \delta_r + \boldsymbol{\zeta}_r\boldsymbol{LV}_q + \upsilon_{rq}$ (we assume that $r \in \{1, \dots, R\}$ measurement elements); $\boldsymbol{\delta}$ is the $(R \times 1)$ vector of constants, $\boldsymbol{\zeta}$ is the $(R \times N)$

matrix of parameters denoting the estimated effect of $\boldsymbol{LV}_q$ on the indicators and υ_{rq} is a logistic distributed disturbance with a diagonal covariance matrix $\boldsymbol{\Upsilon}$ (δ_r and ζ_r are normalized to 0 and 1 respectively for one of the indicators of each latent variable for identification purposes, as suggested by Ben-Akiva *et al.*, 2002). I_{rq} is a categorical indicator with S_r categories and $\boldsymbol{\rho}_r = (\rho_{0r}, \dots, \rho_{S_r})$ is a vector of threshold parameters $\left(\rho_{0r} = -\infty \text{ and } \rho_{S_r} = \infty;\ \rho_{1r} < \rho_{2r} < \cdots < \rho_{S_r-1}\right)$. For identification purposes we impose the condition $\rho_{1r} = 0$. The probability for a certain response *s* to the indicator *r* is thus given by:

$$P(I_{rq} = 1) = \frac{e^{[-\delta_r - \zeta_r \boldsymbol{LV}_q]}}{1 + e^{[-\delta_r - \zeta_r \boldsymbol{LV}_q]}}$$

$$P(1 < I_{rq} < S_r) = \frac{e^{\left[\delta_r + \zeta_r \boldsymbol{LV}_q - \rho_{I_{rq}-1}\right]}}{1 + e^{\left[\delta_r + \zeta_r \boldsymbol{LV}_q - \rho_{I_{rq}-1}\right]}} - \frac{e^{\left[\delta_r + \zeta_r \boldsymbol{LV}_q - \rho_{I_{rq}}\right]}}{1 + e^{\left[\delta_r + \zeta_r \boldsymbol{LV}_q - \rho_{I_{rq}}\right]}} \quad (4)$$

$$P(I_{rq} = S_r) = \frac{e^{\left[\delta_r + \zeta_r \boldsymbol{LV}_q - \rho_{S_r-1}\right]}}{1 + e^{\left[\delta_r + \zeta_r \boldsymbol{LV}_q - \rho_{S_r-1}\right]}}$$

The latent propensity underlying the ordered response observation, that is the cycling frequency reported for each individual q, has been specified as a function of observed and latent variables:

$$y_q^* = \boldsymbol{\beta x}_q + \boldsymbol{\beta}^* \boldsymbol{LV}_q + \varepsilon_q, \quad \varepsilon_q \sim N(0, \Sigma_\varepsilon) \quad (5)$$

where $\boldsymbol{x}_q$ is the $(M \times 1)$ vector of explanatory variables, $\boldsymbol{LV}_q$ is the $(N \times 1)$ vector of individual specific latent variables, $\boldsymbol{\beta}$ and $\boldsymbol{\beta}^*$ are the $(1 \times M)$ and $(1 \times N)$ vectors of unknown parameters to be estimated and ε_q is the error term capturing the effects of unobserved factors on the latent propensity. We assume that ε_q is normally distributed across observations with mean = 0 and variance =1. Because the distribution of ε_q is univariate, the covariance matrix Σ_ε contains only one term, equal to 1. In the usual ordered-response fashion, the latent propensity y_q^* is linked to the observed level y_q through a set of threshold parameters $\boldsymbol{\mu}_q\left(\mu_{0q} = -\infty \text{ and } \mu_{Kq} = +\infty;\ \mu_{1q} < \mu_{2q} < \cdots < \mu_{K-1q} \forall q\right)$ as follows:

$$y_q = \begin{cases} 1 & if\ \mu_{0q} < y_q^* \le \mu_{1q} \\ 2 & if\ \mu_{1q} < y_q^* \le \mu_{2q} \\ \vdots & \\ K & if\ \mu_{K-1q} < y_q^* \le \mu_{Kq} \end{cases} \quad (6)$$

In our specific case K = 5. To allow for heterogeneity (across observations) in the thresholds, they are parametrized as a function of both socio-demographic and latent variables as in Eluru *et al.* (2008):

$$\mu_{kq} = \mu_{k-1q} + exp(\alpha_k + \boldsymbol{\vartheta}_k \boldsymbol{x}_q + \boldsymbol{\gamma}_k \boldsymbol{LV}_q) \quad (7)$$

where α_k is a scalar, $\boldsymbol{\vartheta}_k$ and $\boldsymbol{\gamma}_k$ are $(1 \times M)$ and $(1 \times N)$ vectors of coefficients associated with level k=1, 2, …, K-1. For identifications purposes, we impose the normalization $\mu_{1q} = const$ for all q. The formulation of the thresholds in Eq. (6), where a higher threshold is specified as the sum of its preceding threshold μ_{k-1q} and a non-negative term $exp(\alpha_k + \boldsymbol{\vartheta}_k x_q + \boldsymbol{\gamma}_k LV_q)$, guarantees the increasing order of the thresholds. Note that if none of the dependent variables influence the value of the thresholds, the generalized ordered probit collapses to the standard ordered probit formulation.

Because we assumed that ε_q is normal distributed, the unconditional probability that decision-maker q belongs to category k is given by:

$$\text{Prob}\,(y_q = k) = \Phi(\mu_{kq} - \boldsymbol{\beta x}_q - \boldsymbol{\beta}^* \boldsymbol{LV}_q) - \Phi(\mu_{k-1q} - \boldsymbol{\beta x}_q - \boldsymbol{\beta}^* \boldsymbol{LV}_q) \quad (8)$$

where $\Phi(\cdot)$ is the cumulative distribution function of the error term ε_q.

Assuming the error components $(\boldsymbol{\omega}_q, \boldsymbol{\upsilon}_q, \varepsilon_q)$ are independent, the joint likelihood function for individual q may be written as follows:

$$LL_q(y, \boldsymbol{I}|\boldsymbol{x}; \boldsymbol{\zeta}, \boldsymbol{\beta}, \boldsymbol{\beta}^*, \boldsymbol{\lambda}, \Sigma_\varepsilon, \Sigma_\upsilon, \Sigma_\omega) = \int_{\boldsymbol{LV}} P(y|\boldsymbol{x}, LV; \boldsymbol{\beta}, \boldsymbol{\beta}^*, \mu_k, \Sigma_\varepsilon) f(\boldsymbol{I}|\boldsymbol{x}, \boldsymbol{LV}; \boldsymbol{\zeta}, \Sigma_\upsilon)\, g(\boldsymbol{LV}|\boldsymbol{x}; \boldsymbol{\lambda}, \Sigma_\omega) d\boldsymbol{LV} \quad (9)$$

where P denotes the probability function of observing that the individual belongs to the frequency category k, the density function f for the indicators of the latent variables corresponds to the measurement equation of the latent variable model, and the density function g of the latent variables corresponds to the structural equation of the latent variable model.

Simulation techniques are applied to approximate the multidimensional integral in the likelihood function, and the resulting simulated log-likelihood function is maximized. The models were estimated using Python Biogeme software (Bierlaire, 2016).

5. MODEL RESULTS

Table 7 presents the results of the structural equation component of the model. It emerged that the *Perception of cycling benefits* construct is positively affected by the number of bikes in the household. Further, individuals with a lower level of education have a better perception of the positive aspects of cycling. Note that we found a high value of the constant, which suggests that this latent variable is truly latent and only a few explanatory variables are able to explain it.

The second construct is the *Perception of cycling comfort*. As found in other studies (Akar *et al.*, 2013; Habib *et al.*, 2014), males are more comfortable with cycling than females. The positive sign associated with the variable *Age* indicates that younger people place less importance on the limitations of the bike mode. In line with other studies (Handy *et al.*, 2010; Noland *et al.*, 2011), we also found that the presence of children at home makes people less inclined to cycle. Not surprisingly, the presence of children in the household negatively affects the latent variable, suggesting the existence of other kinds of barriers for this segment of individuals in their choice to cycle.

The number of cars per household negatively influences the latent construct. This effect could reflect the fact that car-addicted individuals, since they are used to the car's comfort, tend to stigmatize the disadvantages of cycling. On the other hand, the number of bikes per household has a positive effect on the latent variable. Furthermore, people with a high level of education are more likely to recognize the negative aspects of cycling.

The structural model related to the *Perceived importance of bike infrastructure* shows that women consider proper infrastructure more important than men. One possible interpretation of this effect could be the fact that women are more concerned about traffic and safety conditions, a finding consistent with the literature (Akar *et al.*, 2013; Bhat *et al.*, 2015; Manton *et al.*, 2016). Furthermore, people with a lower education (high school or lower) are more likely to consider as important the presence of facilitators encouraging cycling than those with a higher level of education.

Table 7. Determinants of latent constructs.

Explanatory Variables	**LV1 – Perception of cycling benefits**		**LV2 - Perception of cycling comfort**		**LV3 - Perceived importance of bike infrastructure**	
	Coeff	R T-stat	Coeff	R T-stat	Coeff	R T-stat
Age	--	--	-0.037	-1.85	--	--
Gender (male=1, female=0)	--	--	0.045	1.29	-0.158	-2.38
Bachelor's degree or higher (yes=1, no=0)	-0.332	-3.02	-0.093	-2.56	-0.173	-2.50
# of bikes per household	0.271	5.20	0.128	6.71	0.132	4.24
# of cars per household	--	--	-0.081	-3.20	--	--
Presence of children (yes=1, no=0)	--	--	-0.134	-3.43	-0.078	-1.09
Constant	6.010	33.33	1.670	14.97	2.660	28.77
Variance	1.790	10.71	0.540	13.29	1.220	20.24

"--"in a cell indicates that the variable in the corresponding row does not have a significant impact on the utility of the alternative in the corresponding column.

Table 8 presents the results of the measurement model. Several indicators were considered in the latent variable measurement model, which linked the latent variables to the responses to the qualitative attitudinal survey questions. The ζ parameters that indicate the associations between the responses to the items and the associated latent variable all have the expected signs. For example, a more positive *perception of cycling benefits* will lead respondents to agree more with the statements about the bike as a healthy and environmentally friendly mode of transport.

Table 8. Impact of latent variables on non-nominal dependent variables.

Latent Variable	Indicators	Const. δ	R T-stat	Coef. ζ	R T-stat
Perception of cycling benefits	A1. It is a rapid means of transport	-0.629	-1.43	0.801	8.75
	A4. It is not expensive	0.152	0.37	0.750	8.66
	A6. It avoids wasting time looking for parking	-0.702	-1.66	0.788	8.80
	A7. It is healthy	-1.780	-2.75	1.380	8.83
	A9. It allows one to appreciate historic centers and increases accessibility to city services	-0.651	-1.71	0.853	10.61
	A11. It contributes to reducing polluting emissions	0.000	n/a	1.000	n/a
Perception of cycling comfort	A5. It does not involve exposure to bad weather and air pollution	-2.080	-5.85	2.260	8.77
	A8. It is easy to carry heavy items	-3.970	-8.58	2.680	8.50
	A10. No need for cycling gear	0.000	n/a	1.000	n/a
	A12. It does not limit daily activity patterns	-0.576	-2.55	2.130	12.59
Perceived importance of bike infrastructure	C1. Presence of an extensive network of dedicated bike lanes	-0.765	-1.90	3.430	8.99
	C2. Presence of racks and secure parking for bikes	-0.864	-2.11	3.550	9.83
	C3. Greater extension of the LTZ or pedestrian zones	0.000	n/a	1.000	n/a

"n/a" not applicable

The estimation results of the discrete part of the model are reported in Table 9. For reasons of identification, there is no constant in the latent propensity to cycle and the first threshold is constant, with no socio-demographic attributes or latent variables in its specification.

Table 9 contains three columns. The first column corresponds to the estimate of the parameters characterizing the latent propensity to cycle. The second column corresponds to the estimates of constant and the parameters for the μ_2 threshold (threshold between the frequency levels "1-10 times per year" and "1-5 times per month"). The third and fourth columns correspond to the constant and the parameters linked to the third threshold, delimiting the "1-5 times per month" and "more than once a week" categories, and to the fourth threshold, demarcating the "more than once a week" and "every day" cycling categories.

Some socio-economic variables were found to have a significant effect on the propensity to cycle. In agreement with several previous studies (Stinson and Bhat, 2004; Noland *et al.*, 2011; Fu and Farber, 2017), males are more likely to cycle. In addition to the effect on the propensity to cycle, the gender variable also impacts the value of the thresholds. If the coefficient associated with a variable in the vector μ_k is negative, the corresponding threshold shifts to the left. In this case, the pattern of threshold effects indicates that men are more likely than women to belong to the category "1-5 times per month", which differs from the result obtainable with a standard ordered probit model. Regarding age, older individuals are less inclined to cycle, a finding consistent with other studies (Ma *et al.*, 2014; Bhat *et al.*, 2017). Interestingly, the Body Mass Index (a biometric datum used as an indicator of ideal weight) negatively affects the propensity to bike, indicating that people with BMI higher are less inclined to cycle. The causal relationship underlying this correlation could go in either or both directions: healthier people are more likely to use the bike or people who cycle are likely to be healthier due to the benefits of physical activity. The level of education also impacts the propensity to cycle. In contrast with other works (Handy *et al.*, 2010; Bhat *et al.*, 2017) respondents with a bachelor's degree are less likely to cycle. One possible explanation could be that university graduates usually have a job where a dress code is often required, incompatible with bike usage in a workplace not equipped with locker rooms, as in our context.

Different household demographics play a crucial role in the propensity to cycle frequently: number of cars, number of bikes, number of household members. As expected, the number of bikes per family positively affects the latent propensity. By contrast, the number of household members negatively impacts the propensity to cycle. This might be because people with children often have to chain trips and do drop offs/pick ups, which might be difficult if traveling by bike. Also, the number of cars per household has a significant negative impact on bike use.

Among the built environment variables, living in an urban neighborhood and the presence of bike paths within 400m from home positively affect cycling frequency. The effect on the threshold here indicates that, compared to the standard probit, the generalized probit predicts a lower probability of individuals living in urban areas being occasional cyclists. Instead, no association was found with the percentage of green areas within 400m from home. It is possible that the variable *urban neighborhood* captures this factor, as in our context the majority of parks and recreational areas are located in urban areas. We also tested the effect of the variable *distance from home to the nearest bus stop*, given the importance of the integration between public transport and cycling (Martens, 2007, La Paix *et al.*, 2020), but the variable was not found to be statistically relevant. This effect can be explained by the fact that in our context it is not possible to take bikes on buses or park bikes safely close to bus stops or train stations

All the attitudinal factors had a positive impact on the propensity to cycle. The positive influence of the latent variable *Perceived importance of bike infrastructure* emphasizes the importance of providing bike facilities such as bike paths or bike lanes, rental bikes, and bike parking. People who are aware of cycling benefits, such as protecting the environment as well as staying healthy, (latent variable *Perception of cycling benefits*) were more likely to be cyclists. This finding suggests that if the benefits of cycling are better understood, then more people would be likely to cycle more frequently. The latent variable *Perception of cycling comfort* is also positively associated with the propensity to cycle. This may be because the more people feel that cycling is easier and more accessible, the more they cycle. Interestingly, the effects on the thresholds show that the generalized probit predicts a lower probability, with higher values of LV2 and LV3, of being occasional cyclists, validating the idea of using this modeling structure for our analysis.

Table 9. Estimation results of the hybrid generalized ordered probit.

Variables	**Latent propensity to cycle**		**Threshold between "1-10 times per year" and "1-5 times per month"**		**Threshold between "1-5 times per month" and "More than once a week"**		**Threshold between "More than once a week" and "every day"**	
	Estimate	*R t-stat*	*Estimate*	*R t-stat*	*Estimate*	*R t-stat*	*Estimate*	*R t-stat*
Threshold between not being a cyclist and cycling **"**1**-**10 times per year**"** μ_{1q}	1.140	3.31	n/a	n/a	n/a	n/a	n/a	n/a
Threshold constants α_k	n/a	n/a	0.413	1.53	-0.246	-2.14	-0.005	-0.10
Age	-0.099	-3.07	--	--	--	--	--	--
Gender (male=1, female=0)	0.543	8.17	-0.359	-3.42	-0.125	-1.22	--	--
Bachelor's degree or higher (yes =1, no=0)	-0.130	-2.19	--	--	--	--	--	--
Body Mass Index	-0.030	-3.23	--	--	--	--	--	--
# of bikes per household	0.811	25.26	--	--	--	--	--	--
# of cars per household	-0.063	-1.33	--	--	--	--	--	--
# of household members	-0.281	-9.82	--	--	--	--	--	--
Residential location (urban = 1, suburban and rural = 0)	0.082	1.09	-0.212	-1.83	-0.210	-1.82	--	--
Presence of bike paths within 400m from home (yes = 1, no = 0)	0.125	2.28	--	--	--	--	--	--
LV1 **–** Perception of cycling benefits	0.114	4.80	--	--	--	--	--	--
LV2 **-** Perception of cycling comfort	0.473	5.30	-0.269	-2.03	--	--	--	--
LV3 **-** Perceived importance of bike infrastructures	0.086	2.70	-0.099	-1.85	--	--	--	--

"**--**"in a cell indicates that the variable in the corresponding row does not have a significant impact on the utility of the alternative in the corresponding column. "n/a" not applicable

5.1. Pseudo-elasticity effects

The coefficients in Table 9 do not provide a sense of the magnitude and direction of the effects of each variable on each cycling frequency category. But we can compute aggregate-level "pseudo-elasticity effects" of socio-demographic variables (Bhat *et al.*, 2017; Hirk *et al.*, 2017), that can be calculated as:

$$\Delta P(y = k|\boldsymbol{x}, \widetilde{\boldsymbol{x}}) = \sum_{q=1}^{Q} \frac{1}{Q} \left[P\left(y_q = k|\widetilde{\boldsymbol{x}}_q\right) - P\left(y_q = k|\boldsymbol{x}_q\right)\right] \quad (10)$$

where all elements of $\widetilde{\boldsymbol{x}}$ are equal to $\boldsymbol{x}$ except for the socio-demographic variable *v*, which is equal to $\tilde{x}_v = x_v + \Delta x_v$ because of the discrete change Δx_{qv} in the variable x_v. Note that Q is the dimension of our sample. For the dummy variable, we first set the value of the dummy variables to zero (note that all other exogenous variables keep their original values) and we then compute the probability $P\left(y_q = k|\boldsymbol{x}_q\right)$ for each frequency category of each independent variable. Next, we change the value of the dummy variable from zero to one for each individual and we calculate the new probability $\left(y_q = k|\widetilde{\boldsymbol{x}}_q\right)$. To obtain the disaggregate pseudo-elasticity effect we compute the difference $\left[P\left(y_q = k|\widetilde{\boldsymbol{x}}_q\right) - P\left(y_q = k|\boldsymbol{x}_q\right)\right]$. For the count variable (*e.g.* the number of cars per household) we change the value of the count variable by the value of one and then we calculate the percentage of change for each category. For the continuous variable, we decrease the value of the variable by 20% for each individual. It is important to note here that, since the latent variables influencing each level of cycling frequency are a function of socio-demographic variables, as socio-demographic variables change so too does the value of the latent variables.

Table 10 shows the pseudo-elasticity effects. The numbers in the table can be interpreted as the percentage change in the probability of each cycling frequency category after a change in the socio-demographic variable. For example, the first entry in the table indicates that, *ceteris paribus*, the probability of a male not being a cyclist is 14.7% lower than for a female. For the variable bike lanes, the probability of an individual not being a cyclist decreases by 3.27% if bike lanes are provided in the individual's residential neighborhood. The directions of the elasticity effects of the model are consistent with the discussion of the results presented in the previous section.

Table 10. Pseudo-elasticity effects

Variables	I never cycle	1-10 times per year	1-5 times per month	More than once a week	Every day
Gender (male)	-14.70%	-2.96%	+2.26%	+7.96%	+7.44%
Age	+2.91%	-0.28%	-0.56%	-1.04%	-1.03%
Body Mass Index	+3.68%	-0.49%	-0.66%	-1.26%	-1.26%
Education (bachelor's degree or higher level of education)	+5.95%	-0.15%	-1.14%	-2.27%	-2.38%
# of cars per household	+2.66%	-0.08%	-0.55%	-1.02%	-1.01%
# of bikes in the household	-22.73%	-0.07%	+2.29%	+7.61%	+12.91%
# of household members	7.34%	-1.07%	-1.39%	-2.51%	-2.36%
Presence of children in the household	+1.84%	+0.33%	-0.46%	-0.85%	-0.86%
Residence location choice (urban)	-0.81%	-1.72%	-1.13%	+1.95%	+1.71%
Presence of bike lanes within 400m of home	-3.27%	+0.39%	+0.56%	+1.12%	+1.19%

5.2. Impact of the latent variables on choice probability

It is not possible to analyze the direct impact of the latent variables on the latent propensity by computing the pseudo-elasticity effects. In fact, it would be meaningless to artificially increase by some percentage points the value of the latent variables, since they were estimated as a function of socio-economic variables. Instead, as suggested by Hess *et al.* (2018), we tested what would happen if all the individuals were assigned a value of the latent variables equal to that of a given segment of population. No changes were made to the discrete part of the model, *i.e.* the direct impact of socio-demographics on latent propensity and thresholds.

The results are reported in Table 11. We ran different test scenarios for each of the three latent variables. The scenarios concern gender (two possible values), education (two possible values), household composition (two possible values), car ownership (one possible value) and bike ownership (one possible value). Note that for car and bike ownership it is possible to develop more than one scenario, *e.g.* one for each different ownership level. But, to keep the presentation simple, we only focused on those scenarios that assign to everyone the perception of an individual not owning a car or bike.

We found that changes in the latent variable LV3, *Perceived importance of bike infrastructure,* only have a minor impact on choice probability, because of the low value of the parameters on the latent propensity to cycle. The greatest changes in choice probability can be observed in the latent variable LV2, *Perception of cycling comfort,* either adopting the psycho-attitudinal variable of the individuals owning no cars or bikes. Finally, for the latent variable LV1, *Perception of cycling benefits,* a shift in bike ownership level produces the greatest impact in choice probability.

Table 11. Test Scenarios

	I never cycle	1-10 times per year	1-5 times per month	More than once a week	Every day
Latent variable LV1: Perception of cycling benefits					
Graduate degree	+0.41%	-0.04%	-0.06%	-0.14%	-0.17%
No graduate degree	-0.58%	+0.08%	+0.11%	+0.20%	+0.19%
No bikes	+1.30%	+0.00%	-0.13%	-0.49%	-0.67%
Latent variable LV2: Perception of cycling comfort					
	I never cycle	**1-10 times per year**	**1-5 times per month**	**More than once a week**	**Every day**
Graduate degree	+0.48%	+0.08%	-0.10%	-0.21%	-0.24%
No graduate degree	-0.67%	-0.10%	+0.17%	+0.31%	+0.28%
No bikes	+2.57%	+0.80%	-0.47%	-1.31%	-1.59%
Male	-0.29%	-0.02%	+0.10%	+0.13%	+0.08%
Female	+0.27%	+0.06%	-0.04%	-0.12%	-0.17%
No car	-1.71%	-0.29%	+0.40%	+0.79%	+0.81%
Children	+0.78%	+0.08%	-0.21%	-0.34%	-0.31%
No children	-0.41%	-0.09%	+0.09%	+0.20%	+0.20%
Latent variable LV3: Perceived importance of bike infrastructure					
	I never cycle	**1-10 times per year**	**1-5 times per month**	**More than once a week**	**Every day**
Graduate degree	+0.16%	+0.07%	-0.05%	-0.09%	-0.10%
No graduate degree	-0.23%	-0.10%	+0.08%	+0.13%	+0.11%
No bikes	+0.47%	+0.31%	-0.10%	-0.30%	-0.38%
Male	+0.19%	+0.07%	-0.08%	-0.11%	-0.06%
Female	-0.17%	-0.09%	+0.03%	+0.10%	+0.13%
Children	+0.08%	+0.03%	-0.03%	-0.05%	-0.04%
No children	+0.38%	+0.03%	-0.08%	-0.16%	-0.19%

It is also interesting to explore how the value of the latent variables varies among the different frequency categories. In doing this, we segment the sample based on the value of the *N* latent variables. In particular, for each latent variable *n* we split the sample into tertiles comprised of the top, middle and bottom third:

- the bottom third includes individuals with the lowest values of the *n-th* latent variable;
- the middle third consists of individuals with medium values of the *n-th* latent variable;
- the top third includes individuals with the highest values of the *n-th* latent variable.

Figures 2, 3 and 4 depict the predicted share of individuals in each cycling frequency category by the value of latent variables. Note that the sum of the values reported at the top of each bar with the same color is 100%. The graphs in the figures are consistent with the results reported in the previous sections, indicating that as the value of the perception of cycling benefits and cycling comfort increases, so too does the probability of using the bike more frequently. For example, the probability of not being a cyclist for individuals with a low perception of cycling comfort is 78.6%. Conversely, for individuals with a high level of perception of cycling comfort the model indicated that the probability of being a frequent cyclist is 42.3%

(24.5% + 17.8%). Further, we found that for the segment with a medium perception of cycling comfort the probability of choosing never to cycle is 49.9%, while the distribution of the choice probability for the remaining cycling categories is almost flat.

We observed a similar trend when looking at the choice probability for each cycling frequency for the latent variable 3 class. However, note that, for individuals who perceive bike infrastructure as less important, the probability of choosing not to cycle is significantly lower than for individuals with a lower value of the latent variables 1 and 2. One possible interpretation of this outcome may be that some experienced cyclists are more aware of the presence of barriers to cycling.

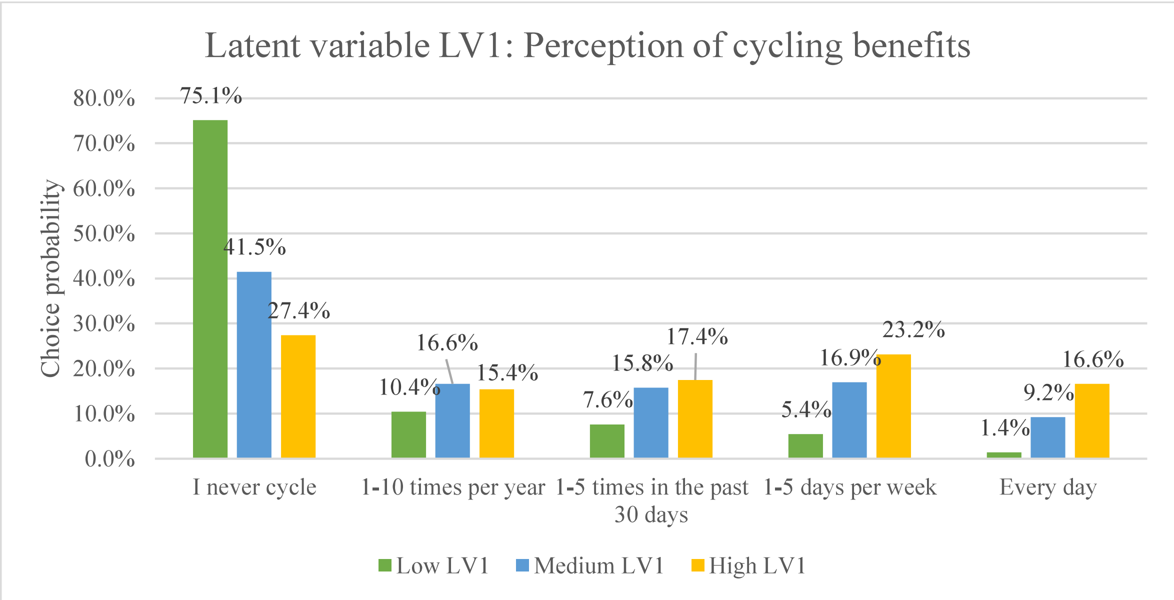

Figure 2. Choice probability for each cycling frequency for LV1 class

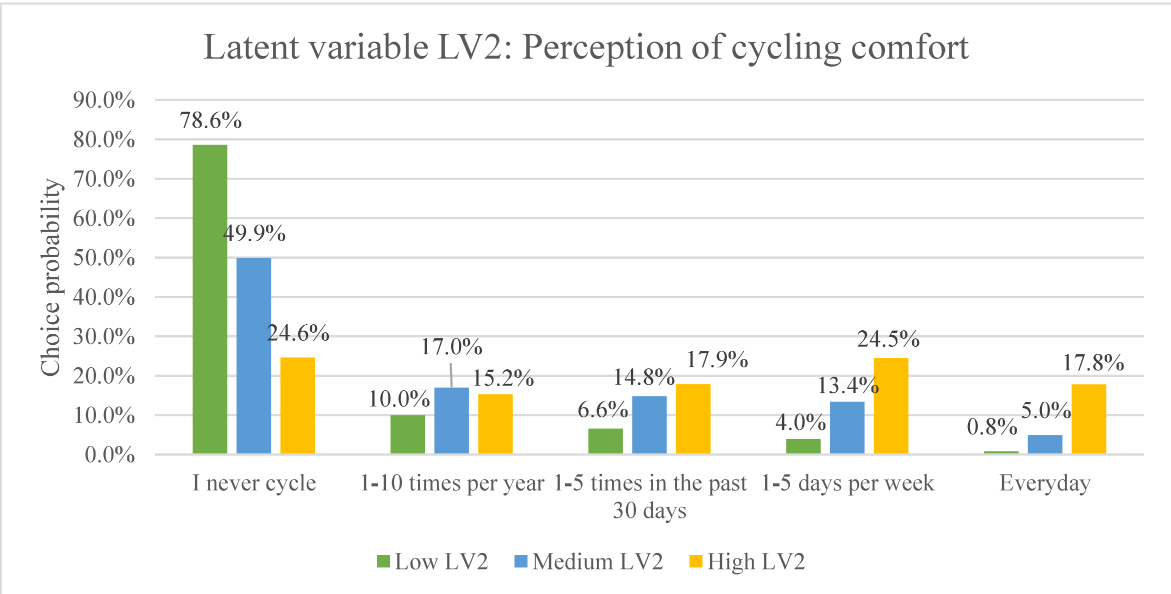

Figure 3. Choice probability for each cycling frequency for LV2 class

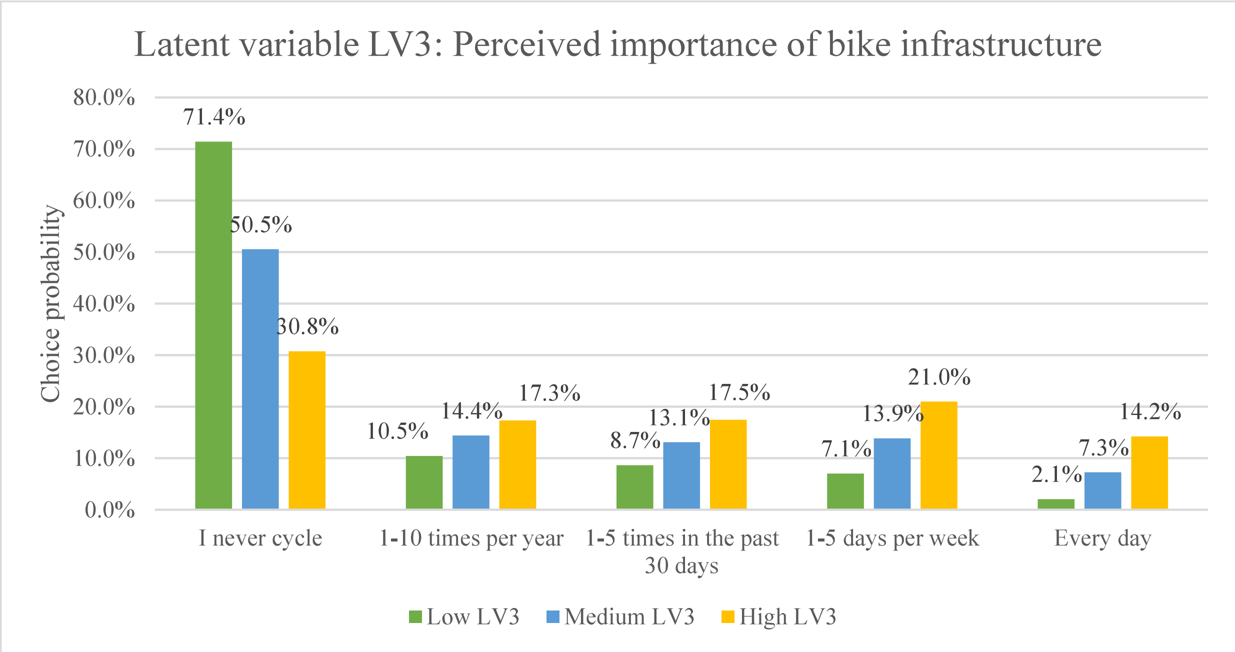

Figure 4. Choice probability for each cycling frequency by LV3 class

5.3. Data fit

Evaluating the data fit of hybrid choice models is not an easy task. Following the approach suggested by Walker (2001), we determine the log-likelihood value of the null model, which considers each choice option equally likely, and the log-likelihood value of the discrete model obtained by removing the measurement model of the latent variables. We examine the data fit of the estimated hybrid generalized ordered probit with its restrictive version, a hybrid standard ordered probit. Then, as proposed by Bhat *et al.* (2017), we calculate the percentage of share of each category for each model and compare the predicted shares with the observed shares.

The results of the data fit comparisons are shown in Table 12. As can be observed, the hybrid generalized ordered probit exhibits a slightly higher value of the likelihood ratio index than the standard model, suggesting that the generalized model offers a superior goodness-of-fit. Instead, in terms of probabilities, the generalized model tends to overestimate choice probabilities compared to the standard one.

The two models can also be compared with a likelihood ratio test. The χ^2 test statistic of the likelihood ratio test between the generalized model and the standard model is statistically significant at any degree of confidence.

Table 12. Model Goodness of fit

	Hybrid Generalized ordered probit	**Hybrid Standard Ordered probit**	-
Log-likelihood value at convergence	-31,451.15	-31,465.17	-
Number of parameters	90	84	-
Log-likelihood value of discrete model	-2,401.02	-2,415.92	-
Log-likelihood of null model	-3,424.88	-3,424.88	-
Log-likelihood of constant model	-2,902.34	-2,902.34	-
Likelihood ratio index	0.299	0.294	-
Likelihood ratio test	$\chi^2 = -2[-31{,}451.15 - (-31{,}465.17)] = 28.04$, 6 df, p = 0.0001		
Cycling frequency	Predicted percentage	Predicted percentage	Observed percentage
I never do	51.0%	50.8%	50.0%
1-10 times per year	14.1%	14.0%	14.6%
1-5 times in the past 30 days	13.1%	13.3%	14.2%
1-5 days per week	14.0%	14.1%	14.6%
Every day	7.9%	7.7%	6.5%

6. CONCLUSIONS

The purpose of the current research was to examine whether psycho-attitudinal factors toward bike use vary in a sample of local authority employees with different cycling frequency and to quantify the determinants influencing the propensity to cycle. Some papers have studied the impact of subjective variables on cycling frequency in different contexts. However, the majority of these treat the problem employing an inappropriate methodology from an econometric perspective, either adopting a two stage sequential approach or assuming the frequency variable as a continuous variable, despite it being measured in ordinal discrete categories. By contrast, we estimated a hybrid choice model with a generalized ordered probit choice kernel, where we defined the thresholds as a function of both objectives and psycho-attitudinal variables. This approach allowed us to overcome the limitations of previous studies and gain further insight into the decision-making process underlying bike usage and frequency.

According to the estimated parameters in latent propensity and threshold functions, several factors affect the frequency of using the bike, especially gender, education level, number of cars per household and number of household members. Residential location and presence of bike facilities variables were also found to influence the latent propensity, suggesting that a good urban design and the provision of bike infrastructure are essential for increasing cycling levels. However, computation of the pseudo-elasticity effects showed that the impact of some socio-demographic factors is much greater than the effect of the built environment characteristics. Therefore, strategies focusing simply on the physical part of the problem, such as expanding and improving bike infrastructure, might not be sufficient to encourage cycling.

The results also indicate that, besides objective characteristics, the latent propensity to cycle is positively influenced by psycho-attitudinal variables, supporting the idea of a relationship between attitudes and perceptions and cycling experience. These conclusions are consistent with the findings of previous studies indicating that a greater inclination to cycle leads individuals to use the bike more frequently. Further, our model specification, with the threshold function of psycho-attitudinal variables, suggests the presence of individual heterogeneity among people with different cycling frequency. In particular, it has been found that a significant variation of the probability of being an occasional cyclist depends on the value of the latent variables *perception of cycling comfort* and *perceived importance of bike infrastructure*.

In view of these results, considering people's psychological characteristics is essential when implementing policies and strategies aimed at increasing bike use. In fact, they reinforce the idea that promoting cycling, through the implementation of awareness campaigns and educational programs intended to improve individuals' perceptions of the bike mode, can persuade them to consider the bike as an alternative means of transport to private motorized vehicles. Further, our findings about the link between socio-economic characteristics and latent constructs revealed that some promotional strategies would have a stronger impact, depending on the target of the intervention. For example, a marketing campaign involving the distribution of material informing people of the location of safe bike paths and parking spots or that cycling involves greater health benefits than risks (Rojas-Rueda *et al.*, 2011) would have a greater effect among women, who are more concerned with safety issues.

The study, however, is not without limitations. An issue that was not addressed was that we analyzed the frequency of cycling for all purposes, mixing utilitarian and recreational trips. In fact, the determinants (both objective and subjective) affecting the choice to travel by bike may be different, depending on trip purpose (Félix *et al.*, 2017; Sottile *et al.*, 2020). Further, a relationship may exist between cycling for leisure and the choice to cycle for utilitarian purposes. Further work needs to be conducted, through the estimation of multivariate models, to establish the interplay between the use of the bike for different purposes.

In addition, the sample is composed predominantly of public sector employees, who are mainly aged between 41-60, are highly educated and live in households with a larger number of components. Understanding the decision to cycle of different categories of people is crucial for policy makers who intend to implement effective strategies for promoting bike use. In fact, it is possible that individuals with different socio-economic characteristics have different attitudes and propensities compared to our sample, so that our results cannot be generalized to the entire population.

Lastly, it should be pointed out that a major concern with hybrid choice models is the derivation of transport policies because of the endogenous nature of psycho-attitudinal variables (Chorus and Kroesen, 2014). Nevertheless, it is worth highlighting that, in some contexts, the explicability of a phenomenon is far more important than predictability. In our specific case, a black-box model may have been superior in terms of predicting probabilities in the short-term under different scenarios but it would have deprived us of the possibility of understanding which psycho-attitudinal factors come into play when an individual decides to cycle and their significance in the process of choosing to cycle.

REFERENCES

Akar, G., Fischer, N., & Namgung, M. (2013). Bicycling choice and gender case study: The Ohio State University. *International Journal of Sustainable Transportation*, 7(5), 347-365.

Allen, J., Eboli, L., Mazzulla, G., & Ortúzar, J.d.D. (2018). Effect of critical incidents on public transport satisfaction and loyalty: an Ordinal Probit SEM-MIMIC approach. *Transportation*, 1-37.

Arroyo, R., Ruiz, T., Mars, L., Rasouli, S., & Timmermans, H. (2020). Influence of values, attitudes towards transport modes and companions on travel behavior. *Transportation research part F: traffic psychology and behaviour*, 71, 8-22.

Balusu, S. K., Pinjari, A. R., Mannering, F. L., & Eluru, N. (2018). Non-decreasing threshold variances in mixed generalized ordered response models: A negative correlations approach to variance reduction. *Analytic methods in accident research*, 20, 46-67.

Ben-Akiva, M., Walker, J., Bernardino, A. T., Gopinath, D. A., Morikawa, T., & Polydoropoulou, A. (2002). Integration of choice and latent variable models. *Perpetual motion: Travel behaviour research opportunities and application challenges*, 431-470.

Bhat, C. R., Astroza, S., & Hamdi, A. S. (2017). A spatial generalized ordered-response model with skew normal kernel error terms with an application to bicycling frequency. *Transportation Research Part B: Methodological*, 95, 126-148.

Bierlaire, M. (2016). PythonBiogeme: a short introduction. Report TRANSP-OR 160706, Series on Biogeme. Transport and Mobility Laboratory, School of Architecture, Civil and Environmental Engineering, Ecole Polytechnique Fédérale de Lausanne, Switzerland.

Bollen, K. A. (1989). A new incremental fit index for general structural equation models. *Sociological Methods & Research*, 17(3), 303-316.

Cervero, R., & Duncan, M. (2003). Walking, bicycling, and urban landscapes: evidence from the San Francisco Bay Area. *American journal of public health*, 93(9), 1478-1483.

Chorus, C. G., & Kroesen, M. (2014). On the (im-) possibility of deriving transport policy implications from hybrid choice models. *Transport Policy*, 36, 217-222.

Dill, J., & McNeil, N. (2013). Four types of cyclists? Examination of typology for better understanding of bicycling behavior and potential. *Transportation Research Record*, 2387(1), 129-138.

Eluru, N., Bhat, C. R., & Hensher, D. A. (2008). A mixed generalized ordered response model for examining pedestrian and bicyclist injury severity level in traffic crashes. *Accident Analysis & Prevention*, 40(3), 1033-1054.

European Environmental Agency (2019). Air quality in Europe — 2019 report www.eea.europa.eu/publications/air-quality-in-europe-2019.

Ewing, R., & Cervero, R. (2010). Travel and the built environment: A meta-analysis. *Journal of the American planning association*, 76(3), 265-294.

Fu, L., & Farber, S. (2017). Bicycling frequency: A study of preferences and travel behavior in Salt Lake City, Utah. *Transportation research part A: policy and practice*, 101, 30-50.

Félix, R., Moura, F., & Clifton, K. J. (2017). Typologies of urban cyclists: review of market segmentation methods for planning practice. *Transportation research record*, 2662(1), 125-133.

Gatersleben, B., & Appleton, K. M. (2007). Contemplating cycling to work: Attitudes and perceptions in different stages of change. *Transportation Research Part A: Policy and Practice*, 41(4), 302-312.

Greene, W. H., & Hensher, D. A. (2010). Modeling ordered choices: A primer. Cambridge University Press.

Gutiérrez, M., Hurtubia, R., & Ortúzar, J.d.D. (2020). The role of habit and the built environment in the willingness to commute by bicycle. *Travel behaviour and society*, 20, 62-73.

Habib, K. N., Mann, J., Mahmoud, M., & Weiss, A. (2014). Synopsis of bicycle demand in the City of Toronto: Investigating the effects of perception, consciousness and comfortability on the purpose of biking and bike ownership. *Transportation research part A: policy and practice*, 70, 67-80.

Handy, S. L., Xing, Y., & Buehler, T. J. (2010). Factors associated with bicycle ownership and use: a study of six small US cities. *Transportation*, 37(6), 967-985.

Handy, S., Van Wee, B., & Kroesen, M. (2014). Promoting cycling for transport: research needs and challenges. *Transport reviews*, 34(1), 4-24.

Heinen, E., Maat, K., & Van Wee, B. (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(2), 102-109.

Heinen, E., Van Wee, B., & Maat, K. (2010). Commuting by bicycle: an overview of the literature. *Transport reviews*, 30(1), 59-96.

Hess, S., Spitz, G., Bradley, M., & Coogan, M. (2018). Analysis of mode choice for intercity travel: Application of a hybrid choice model to two distinct US corridors. *Transportation Research Part A: Policy and Practice*, 116, 547-567.

Hirk, R., Hornik, K., & Vana, L. (2017). mvord: an R package for fitting multivariate ordinal regression models. R package vignette, 2018a. URL https://cran. r-project. org/web/packages/mvord/vignettes/vignette_mvord. pdf.

ISFORT (2018). 15° Rapporto sulla mobilita` in Italia. www.isfort.it.

Kaplan, S., Wrzesinska, D. K., & Prato, C. G. (2019). Psychosocial benefits and positive mood related to habitual bicycle use. *Transportation research part F: traffic psychology and behaviour*, 64, 342-352.

Kelarestaghi, K. B., Ermagun, A., & Heaslip, K. P. (2019). Cycling usage and frequency determinants in college campuses. *Cities*, *90*, 216-228.

Kroesen, M., Handy, S., & Chorus, C. (2017). Do attitudes cause behavior or vice versa? An alternative conceptualization of the attitude-behavior relationship in travel behavior modeling. *Transportation Research Part A: Policy and Practice*, *101*, 190-202.

Ma, L., Dill, J., & Mohr, C. (2014). The objective versus the perceived environment: what matters for bicycling?. *Transportation*, 41(6), 1135-1152.

La Paix, L., Cherchi, E., & Geurs, K. (2020). Role of perception of bicycle infrastructure on the choice of the bicycle as a train feeder mode. *International Journal of Sustainable Transportation*, 1-14.

Manaugh, K., Boisjoly, G., & El-Geneidy, A. (2017). Overcoming barriers to cycling: understanding frequency of cycling in a University setting and the factors preventing commuters from cycling on a regular basis. *Transportation*, 44(4), 871-884.

Manton, R., Rau, H., Fahy, F., Sheahan, J., & Clifford, E. (2016). Using mental mapping to unpack perceived cycling risk. *Accident Analysis & Prevention*, 88, 138-149.

Martens, K. (2007). Promoting bike-and-ride: The Dutch experience. *Transportation Research Part A: Policy and Practice*, 41(4), 326-338.

Muñoz, B., Monzon, A., & Daziano, R. A. (2016). The Increasing Role of Latent Variables in Modelling Bicycle Mode Choice. *Transport Reviews*, 36(6), 737-771.

Namgung, M., & Jun, H. J. (2018). The influence of attitudes on university bicycle commuting: Considering bicycling experience levels. *International Journal of Sustainable Transportation*, 1-14.

Noland, R. B., Deka, D., & Walia, R. (2011). A statewide analysis of bicycling in New Jersey. *International Journal of Sustainable Transportation*, 5(5), 251-269.

Oliva, I., Galilea, P., & Hurtubia, R. (2018). Identifying cycling-inducing neighborhoods: A latent class approach. *International journal of sustainable transportation*, 12(10), 701-713.

Raveau, S., Álvarez-Daziano, R., Yáñez, M. F., Bolduc, D., & Ortúzar, J.d.D. (2010). Sequential and simultaneous estimation of hybrid discrete choice models: some new findings. *Transportation Research Record*, 2156(1), 131-139.

Rojas-Rueda, D., de Nazelle, A., Tainio, M., & Nieuwenhuijsen, M. J. (2011). The health risks and benefits of cycling in urban environments compared with car use: health impact assessment study. *Bmj*, 343, d4521.

Sallis, J. F., Conway, T. L., Dillon, L. I., Frank, L. D., Adams, M. A., Cain, K. L., & Saelens, B. E. (2013). Environmental and demographic correlates of bicycling. *Preventive medicine*, 57(5), 456-460.

Sardegna Statistiche (2018). Sardegna in cifre 2018 http://www.sardegnastatistiche.it/documenti/12_103_20181212133014.pdf

Sener, I. N., Eluru, N., & Bhat, C. R. (2009). Who are bicyclists? Why and how much are they bicycling?. *Transportation Research Record*, 2134(1), 63-72.

Sottile, E., Sanjust di Teulada, B., Meloni, I., & Cherchi, E. (2019). Estimation and validation of hybrid choice models to identify the role of perception in the choice to cycle. *International journal of sustainable transportation*, 13(8), 543-552.

Sottile, E., Diana, M., Piras, F., Meloni, I., & Pirra, M. (2020). To play but not for travel: Utilitarian, hedonic and non-cyclists in Cagliari, Italy. In *Mapping the Travel Behavior Genome* (pp. 209-228). Elsevier.

Stinson, M. A., & Bhat, C. R. (2004). Frequency of bicycle commuting: internet-based survey analysis. *Transportation Research Record*, 1878(1), 122-130.

Stinson, M. A., Porter, C. D., Proussaloglou, K. E., Calix, R., & Chu, C. (2014). Modeling the impacts of bicycle facilities on work and recreational bike trips in Los Angeles County, California. *Transportation Research Record*, 2468(1), 84-91.

Swiers, R., Pritchard, C., & Gee, I. (2017). A cross sectional survey of attitudes, behaviours, barriers and motivators to cycling in University students. *Journal of Transport & Health,* 6, 379-385.

Thigpen, C. (2019). Do bicycling experiences and exposure influence bicycling skills and attitudes? Evidence from a bicycle-friendly university. *Transportation research part A: policy and practice*, *123*, 68-79.

Ton, D., Duives, D. C., Cats, O., Hoogendoorn-Lanser, S., & Hoogendoorn, S. P. (2019). Cycling or walking? Determinants of mode choice in the Netherlands. *Transportation research part A: policy and practice*, 123, 7-23.

Vij, A., & Walker, J. L. (2016). How, when and why integrated choice and latent variable models are latently useful. *Transportation Research Part B: Methodological*, 90, 192-217.

Walker, J. L. (2001). Extended discrete choice models: integrated framework, flexible error structures, and latent variables (Doctoral dissertation, Massachusetts Institute of Technology).

Wang, Y., Chau, C. K., Ng, W. Y., & Leung, T. M. (2016). A review on the effects of physical built environment attributes on enhancing walking and cycling activity levels within residential neighborhoods. *Cities*, 50, 1-15.

Willis, D. P., Manaugh, K., & El-Geneidy, A. (2015). Cycling under influence: summarizing the influence of perceptions, attitudes, habits, and social environments on cycling for transportation. *International Journal of Sustainable Transportation*, 9(8), 565-579.

Xing, Y., Handy, S. L., & Mokhtarian, P. L. (2010). Factors associated with proportions and miles of bicycling for transportation and recreation in six small US cities. *Transportation research part D: Transport and Environment*, *15*(2), 73-81.

Yang, Y., Wu, X., Zhou, P., Gou, Z., & Lu, Y. (2019). Towards a cycling-friendly city: An updated review of the associations between built environment and cycling behaviors (2007–2017). *Journal of Transport & Health*, 14, 100613.

Zhang, C. Q., Zhang, R., Gan, Y., Li, D., & Rhodes, R. E. (2019). Predicting transport-related cycling in Chinese employees using an integration of perceived physical environment and social cognitive factors. *Transportation research part F: traffic psychology and behaviour*, 64, 424-439.