Active Transportation, Environment, and Health

Center for Transportation, Environment, and Community Health
Final Report

by

Ricardo Daziano and So-Yeon Yoon

December, 2019

DISCLAIMER

TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No.	2.Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Active Transportation, Environment, and Health		5. Report Date February 16, 2020
		6. Performing Organization Code
7. Author(s) Ricardo A Daziano (ORCID ID # 0000-0002-5613-429X)		8. Performing Organization Report No.
9. Performing Organization Name and Address School of Civil and Environmental Engineering Cornell University Ithaca, NY 14853		10. Work Unit No.
		11. Contract or Grant No. 69A3551747119
12. Sponsoring Agency Name and Address U.S. Department of Transportation 1200 New Jersey Avenue, SE Washington, DC 20590		13. Type of Report and Period Covered Final Report 11/30/2016 – 12/31/2019
		14. Sponsoring Agency Code US-DOT
15. Supplementary Notes		

16. Abstract

Active transportation –cycling and biking– not only are sustainable travel modes with zero environmental impact, but also have associated health benefits. However, in comparison with motorized transportation, the motives underlying demand for active transportation –especially beyond recreational purposes– is poorly understood, especially because the standard tradeoff between travel time and cost does not apply to active modes (as it is virtually free and usually takes longer). In this project, we investigate the factors that explain demand for active transportation, including non-instrumental attributes, non-standard observed attributes, and extended decision rules. To integrate non-instrumental attributes (attitudes and perceptions) we propose an extension to the hybrid choice model (HCM) that considers data coming from virtual-reality environments. In fact, we designed and implemented virtual, immersive city blocks to analyze valuation of cycling infrastructure. These scenarios will be exploited in a follow-up project.

17. Key Words	18. Distribution Statement
Active transportation; cycling; choice models; experiment design	In addition to presentations in Workshops and Seminars, one publication in Transport Policy (Wang et al., 2018) showcased technical contributions developed as part of this project.

19. Security Classif (of this report)	20. Security Classif. (of this page)	21. No of Pages	22. Price
Unclassified	Unclassified	76	

Form DOT F 1700.7 (8-69)

Chapter 1 Introduction

1.1. Motivation and Statement of the Problem

Active transportation —cycling and biking— not only are sustainable travel modes with zero environmental impact, but also have associated health benefits. However, in comparison with motorized transportation, the motives underlying demand for active transportation —especially beyond recreational purposes— is poorly understood, especially because the standard tradeoff between travel time and cost does not apply to active modes (as it is virtually free and usually takes longer). As a result, travel demand models that cities use to evaluate projects almost exclusively focus on motorized modes.

Taking demand for bicycle for both utilitarian (commuting) and recreational (leisure) purposes as dependent variable, previous and current research has been exploring a wide range of independent variables, from attributes of both the natural and built environment to socio-economic and household characteristics. Literature surveys on this topic include Saelens, Sallis, and Frank (2003) Parkin, Ryley, and Jones (2007); Lorenc, Brunton, Oliver, Oliver, and Oakley (2008); Sirard and Slater (2008); Reynolds, Harris, Teschke, Cripton, and Winters (2009); Panter and Jones (2010); Willis, Manaugh, and El-Geneid (2014); Heinen, van Wee, and Maat (2010). For instance, built environment variables that have been tested in previous research include origin and destination locations; area of green spaces, forests, and public services; land use (single family residential, commercial, educational, industrial, small commercial) and land-use mix and diversity; accessibility (distance to transit, services, and recreational areas including green spaces); street intersection density; and road network type. Regarding transportation infrastructure the role of traffic flow and related externalities (noise, emissions) has been explored at a relative aggregate level.

1.2. Research Goals

This research has three main research goals:

1. As a first line of research we propose to further investigate the factors that explain demand for active transportation, including non-instrumental attributes (c.f. Motoaki and Daziano, 2015), and non-standard observed attributes. To integrate non-instrumental attributes (attitudes and perceptions) the hybrid choice modeling (HCM) approach is applied (Ben-Akiva et al., 1999; Ben-Akiva et al., 2002a; Ben-Akiva et al., 2002b; Walker and Ben-Akiva, 2002) to a case study in the city of Vitoria-Gasteiz in Spain with co-authors Begoña Muñoz and Andrés Monzón, from the Polytechnical University of Madrid. This Spanish city was selected to conduct the case study as bicycle use in Vitoria-Gasteiz has increased sharply during the last years, from 3.3% in 2006 to 6.9% in 2011, and to 12.3% in 2014 (Council of Vitoria-Gasteiz, 2015). Nowadays, Vitoria-Gasteiz is the Spanish city with the highest bicycle share of trips, while its nonmotorized share of trips is the highest of any European medium–size city. Using stated preferences that supplement the revealed-preference data, we build a hybrid chocie model that incorporates the following subjective attributes: safety and comfort (SC), direct advantages (DA), awareness (A), external facilities (EF), individual capacities (IC), and subjective norm (SN). A full paper with this case study is included in Appendix A of this report.
2. Improve the specification of discrete choice models to better accommodate the

representation of unobserved preference heterogeneity. This modeling extension is summarized in Chapter 2.

3. Finally, we build a set of highly realistic immersive scenarios that can be used for in-lab experiments to determine the effect of characteristics of the built environment to either encourage or deter demand for cycling. These scenarios are discussed in Chapter 3.

Chapter 2 Methodological Contributions

2.1. Motivation

Discrete choice analysis has a long tradition in transportation science to model mode and route decisions, including cycling behavior as reviewed in Appendix A. Discrete choice models represent economic decisions made by consumers among a finite set of mutually exclusive and exhaustive alternatives (Train, 2003), usually assuming a utility-maximizing decision rule. In random utility maximization (RUM) models, at a given choice occasion an agent n makes one choice among J alternatives by choosing the alternative that maximizes the indirect conditional utility U_{in}, which is evaluated at each alternative i. In the context of travel decisions that are hard to explain with the traditional time/cost tradeoff (as is the case of cycling choices), the hybrid choice modelling (HCM) framework (Fig. 2-1) has become the main empirical strategy to introduce qualitative attributes of the alternatives.

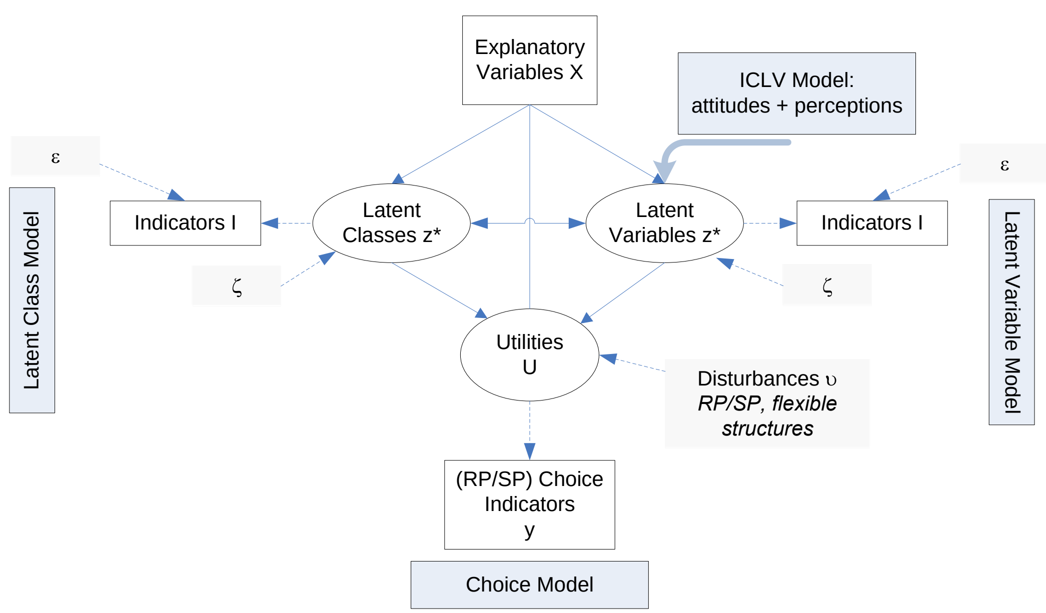


Figure 2-1. General representation of the Hybrid Choice Model

Figure 2-1 represents the general framework of hybrid choice models, with the possibility of including latent variables within an integrated choice and latent variable (ICLV) model. Another component of the framework is clustering agents according to a latent class logit model.

Within the HCM framework, latent attributes representing **attitudes**, **perceptions**, and **multidimensional attributes** (such as comfort, convenience, or health-impact) are explicitly modeled using structural equation modeling (SEM) concepts from psychometrics. In SEM, latent variables are explained by a structural equation and identification is provided by a measurement equation.

A general hybrid choice model is based on the following simultaneous system of latent variables (Daziano, 2015).

Structural equations

$$\underset{(L\times1)}{\mathbf{z}_n^*} = \underset{(L\times L)}{\mathbf{\Pi}}\mathbf{z}_n^* + \underset{(L\times M)}{\mathbf{B}}\underset{(M\times1)}{\mathbf{w}_n} + \underset{(L\times1)}{\zeta_n}, \quad \zeta_n \sim \mathcal{N}(0, \mathbf{H}_\Psi^{-1}) \tag{1}$$

$$\underset{(J\times1)}{\mathbf{U}_{tn}^*} = \underset{(J\times K)}{\mathbf{X}_{tn}}\underset{(K\times1)}{\boldsymbol{\beta}} + \underset{(J\times Q)}{\mathbf{Y}_{tn}^*(\mathbf{X}_{tn}, \mathbf{z}_n^*)}\underset{(Q\times1)}{\boldsymbol{\varrho}} + \underset{(J\times L)}{\mathbf{\Gamma}}\underset{(L\times1)}{\mathbf{z}_n^*} + \underset{(J\times1)}{\boldsymbol{\nu}_{tn}}, \quad \boldsymbol{\nu}_{tn} \sim \mathcal{N}(0, \mathbf{H}_\Sigma^{-1}) \tag{2}$$

$$\underset{(R\times1)}{\mathbf{I}_n^*} = \underset{(R\times1)}{\boldsymbol{\alpha}} + \underset{(R\times L)}{\mathbf{\Lambda}}\underset{(L\times1)}{\mathbf{z}_n^*} + \underset{(R\times1)}{\boldsymbol{\varepsilon}_n}, \quad \boldsymbol{\varepsilon}_n \sim \mathcal{N}(0, \mathbf{H}_\Theta^{-1}) \tag{3}$$

Measurement equations

$$\underset{(1\times1)}{I_{rn}} = \begin{cases} 1 & \text{if } \mu_{0r} < I_{rn}^* \leqslant \mu_{1r} \\ 2 & \text{if } \mu_{1r} < I_{rn}^* \leqslant \mu_{2r} \\ \vdots & \\ M_r & \text{if } \mu_{M_r-1} < I_{rn}^* \leqslant \mu_{M_r}, \end{cases} \tag{4}$$

$$\underset{(1\times1)}{y_{tn}} = i \in C_n \text{ iff } U_{itn} - U_{jtn} \geqslant 0, \ \forall j \in C_n, \ j \neq i, \forall n \in N. \tag{5}$$

where $\boldsymbol{z}_n^*$ is a vector of individual-specific latent explanatory variables that explains utility; the matrix $\mathbf{\Pi}$ allows for interactions among the latent variables; $\boldsymbol{B}$ is a matrix of unknown parameters that measure the global effect of sociodemographics $\boldsymbol{w}_n$ on the latent variables; and Ψ is a full covariance matrix associated with the normally distributed error ζ.

In the choice model in equation (2), $\boldsymbol{U}_{tn}$ is a vector of indirect utility functions with stacked alternatives for individual n; $\boldsymbol{X}_{tn}$ is a design matrix with $\boldsymbol{x}'_{itn}$ designating its i-th row; and $\boldsymbol{\beta}$ is a vector of unknown preference parameters or marginal utilities. $\boldsymbol{Y}_{tn}$ is a matrix of Q interactions between the observed attributes $\boldsymbol{X}_{tn}$ and the latent attributes $\boldsymbol{z}^*_n$ as well as interactions within the latent attributes; $\boldsymbol{\varrho}$ is a vector of unknown parameters associated with these interactions. $\boldsymbol{\Gamma}$ is a matrix of unknown parameters associated with the latent attributes, with $\boldsymbol{\gamma}'_i$ designating its i-th row of matrix. Although the system of equations is displaying a normally-distributed error term ν that leads to a multinomial probit model, it is possible to assume a type 1 extreme-value distributed kernel for a logit-based hybrid choice model.

Equations (3) and (4) represent a system of independent ordered probit models for identification and measurement of the attributes $\boldsymbol{z}^*_n$. In equation (3), $\boldsymbol{I}^*_n$ is a latent continuous vector of manifestations of the latent attributes; $\boldsymbol{\alpha}$ is an intercept and $\boldsymbol{\Lambda}$ is a matrix of unknown factor loadings. Finally, $\boldsymbol{\varepsilon}_n$ is a normally-distributed vector of measurement error terms with covariance matrix $\boldsymbol{\Theta}$. We assume that there are R measurement elements in $\boldsymbol{I}^*_n$.

Because latent variables are not directly observed, the choice probability for the hybrid choice model is derived by integrating out the latent attributes over the whole space of Z^*_n . The unconditional choice probability is the joint probability of the choice and latent indicators y_n and I_n:

$$P\left(y_n, I_n \middle| X_n; \alpha, \beta, \gamma, \Sigma_\varepsilon, \Sigma_\eta, \Sigma_\nu\right) = \int_{Z_n^*} P(y_n|X_n, Z_n^*; \beta, \Sigma_\varepsilon) f(I_n|X_n, Z_n^*; \alpha, \Sigma_\nu) g\left(Z_n^* \middle| X_n; \gamma, \Sigma_\eta\right) dZ_n^*$$

2.2. Previous Related Research (Daziano and Motoaki, 2014[1]; Motoaki and Daziano, 2015[2]).

In our previous, related research (Motoaki and Daziano, 2015), we extended the hybrid choice modeling framework to analyze cycling demand by incorporating a latent class logit model with latent attributes that informed assignment to underlying latent classes. Using data from a choice experiment among members (students and staff) of Cornell University, we analyzed the effects of weather (temperature, rain, and snow), cycling time, slope, cycling facilities (bike lanes), and traffic on cycling decisions. Our results showed that cyclists can be separated into two segments based on a latent attribute that summarized cycling skills: those cyclists with more skills and experience were less affected by adverse weather conditions. From the median of the ratio of the marginal rate of substitution for the two classes, we showed that rain deters cyclists with lower skills from bicycling 2.5 times more strongly than those with better cycling skills. The median of the effects also showed that snow is almost 4 times more deterrent to the class of less experienced cyclists.

2.3. Further modeling extensions

Estimation of the model parameters can be accomplished by derivation of the maximum likelihood estimator (MLE). Equation (6) presents the expression of the likelihood of observing both the choice indicator vector $\boldsymbol{y}$ as well as the indicators $\boldsymbol{I}$ that measure the latent attributes.

[1] Daziano, RA and Y Motoaki, 2014. Data Collection and Econometric Analysis of the Demand for Nonmotorized Transportation. Final Report, UTRC-RF (University Transportation Research Center Region 2) Project No: 49997-35-24 http://www.utrc2.org/sites/default/files/pubs/Final-Data-Collection-Econometric-Analysis.pdf

[2] **Motoaki, Y** and RA Daziano. 2015. A hybrid-choice latent-class model for the analysis of the effects of weather on cycling demand. *Transportation Research Part A* 75, 217-250.

$$\ell(\mathbf{y},\mathbf{I};\boldsymbol{\delta}) = \prod_{n=1}^{N}\int_{z_n^*}\prod_{t=1}^{T}P_{tn}(i_{tn}\left|\mathbf{z}_n^*,\mathbf{X}_n,\boldsymbol{\theta},\mathbf{H}_{\Sigma}^{-1}\right.)\prod_{r=1}^{R}f(I_{rn}|\mathbf{z}_n^*,\boldsymbol{\Lambda},\boldsymbol{\mu}_r,\mathbf{H}_{\Theta}^{-1})g(\mathbf{z}_n^*\left|\mathbf{w}_n,\tilde{\mathbf{B}},\mathbf{H}_{\tilde{\Psi}}^{-1}\right.)dz_n^*, \tag{6}$$

Evaluation of the loglikelihood requires values for the conditional choice probability $P_{tn}(i_{tn}|\mathbf{z}_n^*)$ as well as for the probability density functions of both the indicators and the latent attributes. In the case of multinomial indicators, it is possible to write:

$$f(I_{rn}=m) = \Phi\left(\frac{\mu_{m_r}-\boldsymbol{\lambda}_r'\mathbf{z}_n^*}{[\mathbf{H}_{\Theta}^{-1}]_{rr}}\right) - \Phi\left(\frac{\mu_{m-1_r}-\boldsymbol{\lambda}_r'\mathbf{z}_n^*}{[\mathbf{H}_{\Theta}^{-1}]_{rr}}\right) \tag{7}$$

,

whereas for binary indicators we have:

$$f(I_{rn}) = \Phi\left(\frac{\alpha_r+\boldsymbol{\lambda}_r'\mathbf{z}_n^*}{[\mathbf{H}_{\Theta}^{-1}]_{rr}}\right)^{I_{rn}}\left(1-\Phi\left(\frac{\alpha_r+\boldsymbol{\lambda}_r'\mathbf{z}_n^*}{[\mathbf{H}_{\Theta}^{-1}]_{rr}}\right)\right)^{(1-I_{rn})}. \tag{9}$$

And for continuous indicators:

$$f(I_{rn}) = \frac{1}{[\mathbf{H}_{\Theta}^{-1}]_{rr}}\phi\left(\frac{I_{rn}-\alpha_r-\boldsymbol{\lambda}_r'\mathbf{z}_n^*}{[\mathbf{H}_{\Theta}^{-1}]_{rr}}\right), \tag{10}$$

We have worked on an extension of the hybrid choice model for the consideration of unobserved preference heterogeneity. To do so, consider random preference parameters which can be written as individual-specific parameters $\boldsymbol{\beta}_n$:

$$\underset{(J\times 1)}{\mathbf{U}_{tn}^*} = \underset{(J\times K)}{\mathbf{X}_{tn}}\underset{(K\times 1)}{\boldsymbol{\beta}_n} + \underset{(J\times Q)}{\mathbf{Y}_{tn}^*(\mathbf{X}_{tn},\mathbf{z}_n^*)}\underset{(Q\times 1)}{\boldsymbol{\varrho}} + \underset{(J\times L)}{\boldsymbol{\Gamma}}\underset{(L\times 1)}{\mathbf{z}_n^*} + \underset{(J\times 1)}{\boldsymbol{\nu}_{tn}}, \boldsymbol{\nu}_{tn}\sim\mathcal{N}(0,\mathbf{H}_{\Sigma}^{-1}) \tag{11}$$

On the one hand, if for the random preference parameter $\boldsymbol{\beta}_n$ a **discrete** heterogeneity distribution is assumed, then a latent-class-logit kernel can be hypothesized, as we did in Motoaki and Daziano (2015). On the other hand, for the consideration of a more flexible representation of unobserved taste variation, a discrete-continuous heterogeneity distribution was derived as part of this project and implemented in Wang, Sun, Russell, and Daziano (2018).

References

Daziano, RA. 2015. Inference on mode preferences, vehicle purchases, and the energy paradox using a Bayesian structural choice model. *Transportation Research Part B*, 76, 1-26.

Motoaki, Y and RA Daziano. 2015. A hybrid-choice latent-class model for the analysis of the effects of weather on cycling demand. *Transportation Research Part A* 75, 217-250.

Wang, C, **J Sun**, **R Russell** and RA Daziano. 2018. Analyzing willingness to improve the resilience of New York City's transportation system. *Transport Policy* 69, 10-19.

Chapter 3 Toward Highly Realistic Virtual Environments to Understand Demand for Cycling

3.1. Design of a Virtual Reality Cycling Experiment

As a supplement to standard discrete choice analysis of cycling behavior, a virtual city environment was developed as part of this project. The goal is to simulate a cyclist's traveling experience across variations of cycling infrastructure. The virtual cycling environment will be implemented as part of an integrated lab and survey-based choice experiment to analyze behavioral response to built-environment features that could encourage broader adoption of cycling as a utilitarian and recreational transportation mode.

The urban environment was designed to appear to be believable, busy city blocks with high-rise buildings and sidewalks, along with two to three vehicle lanes. The three-dimensional model of the environment was created in Autodesk 3D Studio Max, and then converted into Twinmotion to program moving vehicles, pedestrians, and cyclists. Twinmotion is an Unreal Gaming Engine based real-time immersive 3D visualization platform that allows for automated objects to follow a programmed path, speed, and density.

3.2. Stimuli Development

After a series of focus groups and our previous research, the experimental conditions for the virtual cycling environment were designed around the following cycling infrastructure and general conditions attributes:

1. Dedicated versus shared routes
2. Painted demarcations
3. Buffer protection

4. Presence of car parking
5. One-way vs two-way cycling traffic

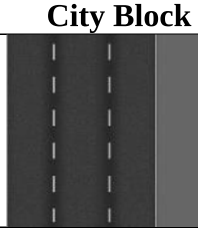
Following a statistical design for choice experiments, 14 different bike lane design options on New York City streets, in two traffic scenarios (light and heavy traffic), were created. The 14 conditions are summarized in the table below.

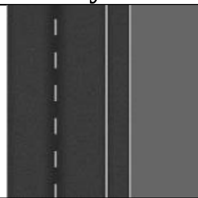
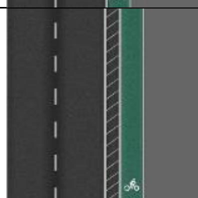
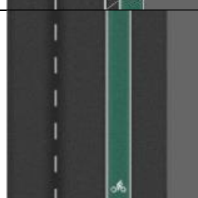
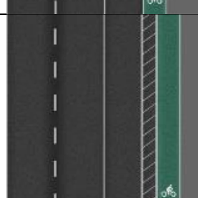
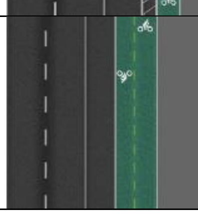
Table: Experimental cycling conditions

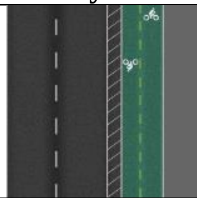
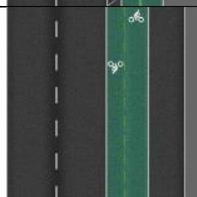
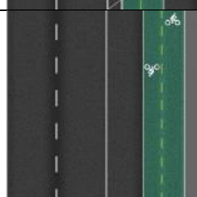
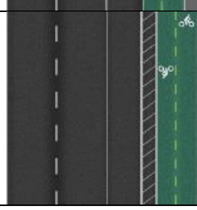
Condition	**Description**
1	Standard lane, shared with cars
2	Marked one-way cycle path, unprotected
3	Painted one-way cycle path, unprotected
4	Protected one-way cycle path
5	Unprotected one-way cycle path, car parking on the right
6	Protected one-way cycle path, car parking on the right
7	Unprotected one-way cycle path, car parking on the left
8	Protected one-way cycle path, car parking on the left
9	Unprotected two-way cycle path
10	Protected two-way cycle path
11	Unprotected two-way cycle path, car parking on the right
12	Protected two-way cycle path, car parking on the right
13	Unprotected two-way cycle path, car parking on the left
14	Protected two-way cycle path, car parking on the left

These 14 cycling-path designs were implemented in a VR environment following the procedure stated above. The following table presents bird's eye-views of the virtual city blocks.

Table: Virtual City Blocks – Infrastructure

Condition	**City Block**	**Description**
1		Standard lane, shared with cars

Condition	City Block	Description
2		Marked one-way cycle path, unprotected
3		Painted one-way cycle path, unprotected
4		Protected one-way cycle path
5		Unprotected one-way cycle path, car parking on the right
6		Protected one-way cycle path, car parking on the right
7		Unprotected one-way cycle path, car parking on the left
8		Protected one-way cycle path, car parking on the left
9		Unprotected two-way cycle path

Condition	City Block	Description
10		Protected two-way cycle path
11		Unprotected two-way cycle path, car parking on the right
12		Protected two-way cycle path, car parking on the right
13		Unprotected two-way cycle path, car parking on the left
14		Protected two-way cycle path, car parking on the left

After the basic 2D design of the city blocks was implemented in Autodesk 3D Studio Max, the three-dimensional version was converted to Twinmotion to add moving vehicles, pedestrians, and cyclists. The figure below presents bird's-eye views of the three-dimensional city blocks.

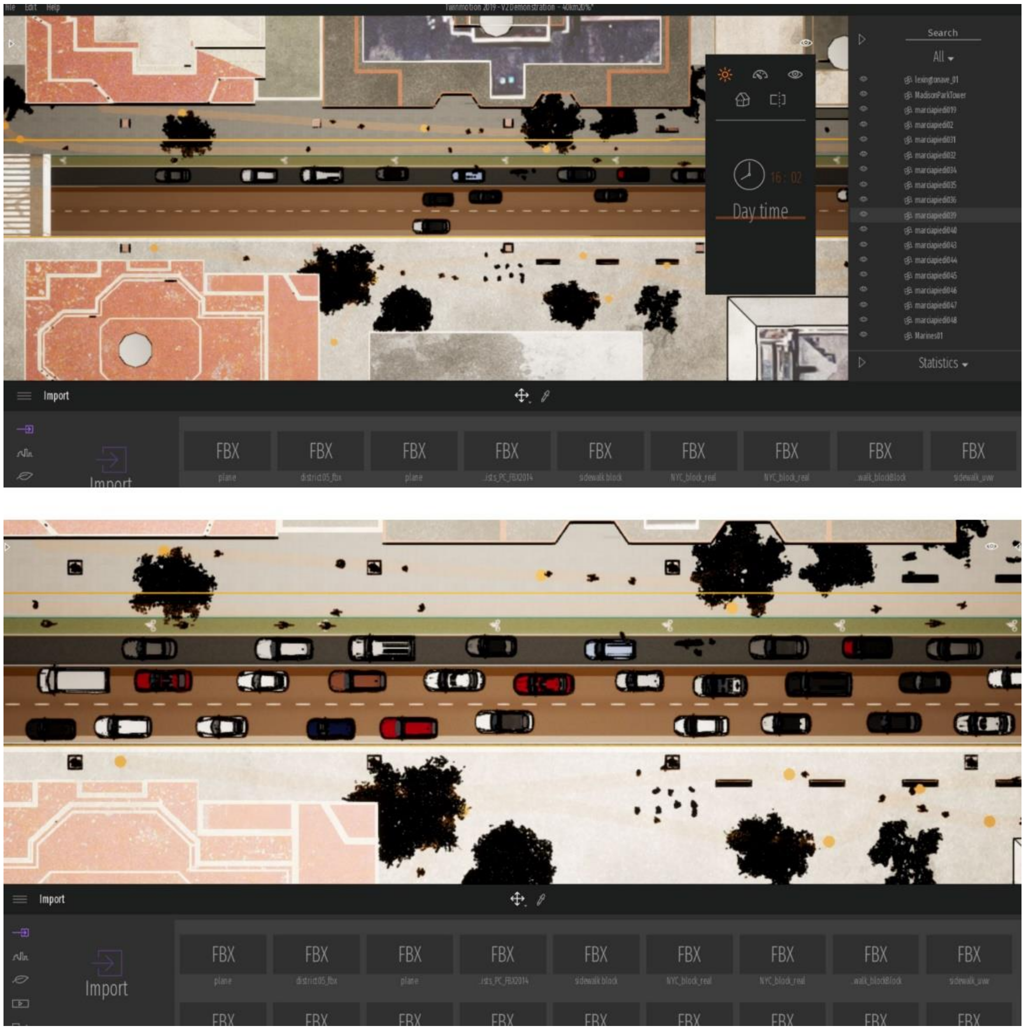


Figure. 3D implementation of the virtual cycling environments.

Two traffic scenarios were implemented, namely: 1) traffic congestion with more vehicles moving very slowly (12 km/hr; 7mph); and 2) fewer cars moving at the NYC speed limit (40 km/hr; 25mph, 2014). The consideration of the two traffic scenarios led to a total of 28 experimental conditions.

Table: Virtual City Blocks – Light Traffic

	City Block	**Descriptio n**

	City Block	Description
1		Standard lane, shared with cars
2		Marked one-way cycle path, un-protected

	City Block	Description
3		Painted one-way cycle path, unprotected
4		Protected one-way cycle path

	City Block	**Description**
5		Un-protected one-way cycle path, car parking on the right
6		Protected one-way cycle path, car parking on the right

	City Block	Description
7		Un-protected one-way cycle path, car parking on the left
8		Protected one-way cycle path, car parking on the left

	City Block	Description
9		Un-protected two-way cycle path
10		Protected two-way cycle path

	City Block	Description
11		Un-protected two-way cycle path, car parking on the right
12		Protected two-way cycle path, car parking on the right

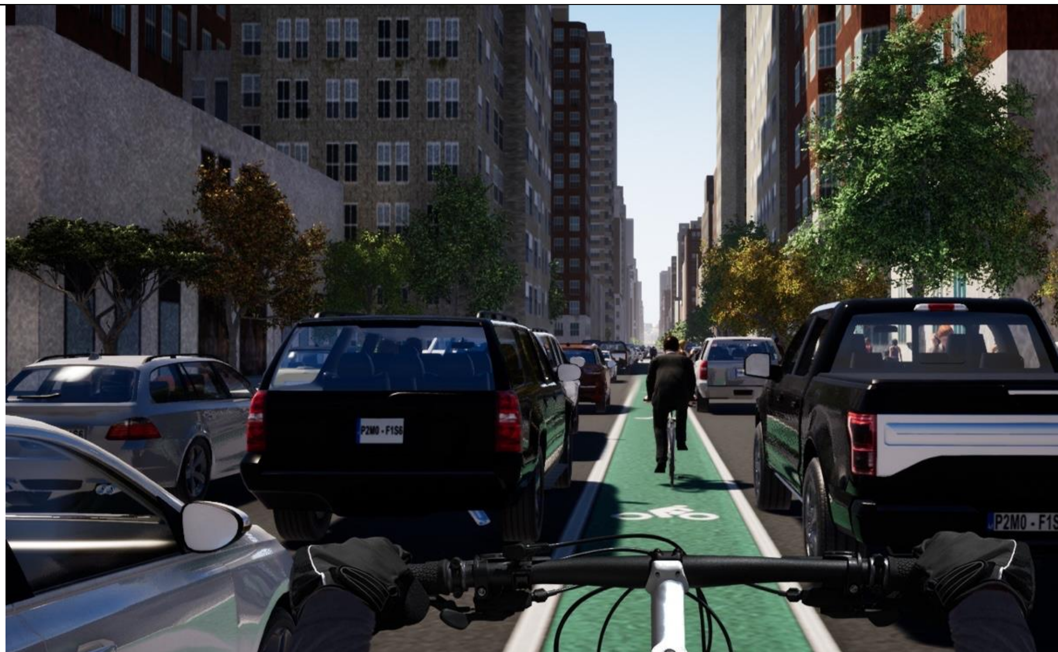
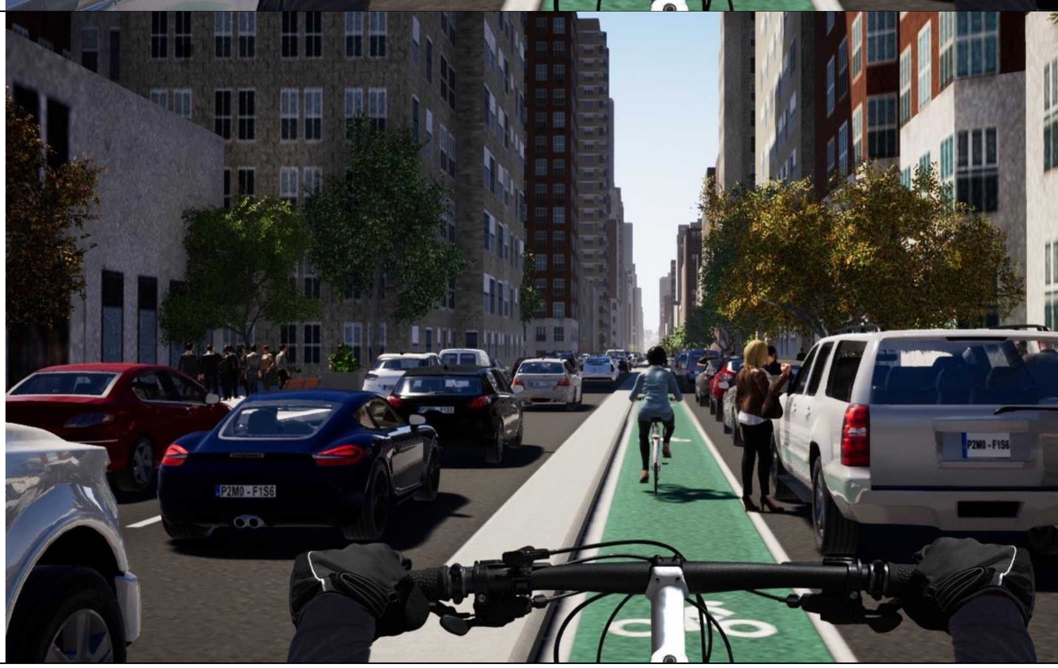
	City Block	**Description**
13		Un-protected two-way cycle path, car parking on the left
14		Protectedt wo-way cycle path, car parking on the left

Table: Virtual City Blocks – Heavy Traffic

	City Block	**Description**

1		Standard lane, shared with cars
2		Marked one-way cycle path, un-protected

3		Painted one-way cycle path, un-protected
4		Protected one-way cycle path

5		Un-protected one-way cycle path, car parking on the right
6		Protected one-way cycle path, car parking on the right

7		Un-protected one-way cycle path, car parking on the left
8		Protected one-way cycle path, car parking on the left

9			Un-protected two-way cycle path
10			Protected two-way cycle path

11		Un-protected two-way cycle path, car parking on the right
12		Protected two-way cycle path, car parking on the right

13			Un-protected two-way cycle path, car parking on the left
14			Protected two-way cycle path, car parking on the left

With the virtual stimuli ready, the final stage for implementation of the experiment is the lab set-up shown in the figure below.

Figure. DUET lab experimental setup

These VR cycling conditions and lab setup will be used for an integrated in-lab choice experiment with an online survey in a follow-up study also funded by CTECH.

APPENDIX A: Case Study in the City of Vitoria-Gasteiz

Modelling the effect of policy measures for improving cycling for urban transport

Begoña Muñoz, Andrés Monzón, and Ricardo A Daziano

1 Introduction

Bicycle as a non-motorized transport alternative is one of the possible solutions to the problems of congestion and environmental nuisance produced by an increasing number of motorized urban trips. Bicycle has actually captured the attention of transport policies aimed at sustainable development, worldwide and especially in Europe (Alegre and Carbonell, 2015; Pucher et al., 2010; Yang et al., 2010).

Many cities in Spain developed bicycle mobility plans during the last decade and started implementing measures to promote the use of the bicycle for urban travel. These policies mainly consisted in building bicycle lanes and safe bicycle parking, facilitating intermodal bicycle-public transport and promoting public bike-share systems. Although these measures have significantly improved shares of bicycle use, Spanish cycling levels are still far from the high levels that are observed in some North European countries, especially when focusing on commuting trips.

Some of the partial failures in encouraging bicycle for commuting are due to poor understanding of the effect of cycling policies on users. Poor characterization of cycling demand stems from the use of a traditional modelling approach to the objective travel cost and time trade-off. Modelling cycling choices requires additional variables – including psychological factors – to take into consideration the subjective value attached to bicycle. In fact, these psychological factors may be the main determinants of demand rather than standard measures of level of service. Psychological factors motivating bicycle choice can be quantitatively modelled as latent variables that cannot be measured directly and thus have to be inferred from observable variables called indicators (Bartholomew et al., 2011). Explicit inclusion of psychological latent variables needs advanced choice models (Muñoz et al., 2016).

As shown in the review by Muñoz et al. (2016), the number of studies using extended discrete choice models –following the hybrid choice model (HCM) framework– for analyzing bicycle choice is relatively limited[3]. Furthermore, these extended models are not used to forecast. Only Maldonado-Hinarejos et al. (2014) developed a forecasting exercise; however, inefficient sequential estimators were used, and the latent variables were constructed using Principal Component Analysis (PCA) – a method that lacks causal variables that can be used to properly evaluate hypothetical scenarios.

Considering the research gaps mentioned above, this paper aims at investigating the effect of several bicycle latent variables on mode choice behavior throughout the development of a jointly estimated integrated choice and latent variable (ICLV) model with revealed preference data,

[3] These studies either followed a sequential estimation approach (Maldonado-Hinarejos et al., 2014; Fernández-Heredia, et al., 2016), or a simultaneous estimation approach (Kamargianni and Polydoropoulou, 2013; Habib et al., 2014; Kamargianni et al., 2015)

paying special attention to the forecasting issue. The proposed ICLV model is used to test several potential transport policies, including soft measures related to bicycle experience that hard to properly analyze with standard discrete choice models.

Data for the model building stems from the 2014 household travel survey (HTS) of Vitoria-Gasteiz (Spain). Data specifically included perceptual indicators towards bicycle use for urban mobility that were used to identify six bicycle latent variables.

In the final specification, two latent variables entered directly into the utility of bicycle. Forecasting results show that a proposed urban toll to cars produces the highest increase in the share of cycling (coming from a large decrease in the use of car). The effectiveness of soft measures (related to the latent variables) was more limited than that related to time and cost, probably due to weak structural relationships in the MIMIC model. Future research might focus on finding better-supported SEM and nonlinearities in the utility specifications to improve forecasting power when latent variables are introduced.

The rest of the paper is structured as follows. The next section summarizes the modelling framework. The third section describes the empirical context and data used. The fourth section describes the model specifications and implements a forecasting exercise. The final section provides a summary of the findings, policy recommendations, and identifies avenues for further research.

2 Methodological framework

A discrete choice model (DCM) properly considering subjective or psychological factors in the form of latent variables, requires following the hybrid choice model (HCM) framework (Ben-Akiva et al., 1999; Ben-Akiva et al., 2002a; Ben-Akiva et al., 2002b; Walker and Ben-Akiva, 2002). The resulting model, belonging to this new generation of discrete choice models, is called "integrated choice and latent variable (ICLV) model" and will be used herein following the latest trends in bicycle mode choice modelling (Muñoz et al., 2016). Because latent variables are not observed, observable indicators and exogenous variables provide identification of the latent constructs, which are integrated[4] into the discrete choice model.

The **latent variable model** is based on a multiple indicators multiple causes (MIMIC) structure. The MIMIC structural equations (1) link observable exogenous variables, such as individuals' characteristics ($X2_n$), with the latent variables (Z_n^*) as dependent variables. The MIMIC measurement equations (2) represent how the latent variables are manifested through observable indicators (I_n) –usually responses to attitudinal or perceptual survey questions. The MIMIC system of equations, for individual *n*, are (Ben-Akiva et al., 2002b):

Structural: $Z_n^* = X2_n\gamma + \eta_n, \eta_n \sim N(0, \Sigma_\eta)$ (1)

Measurement: $I_n = Z_n^*\alpha + \upsilon_n, \upsilon_n \sim N(0, \Sigma_\upsilon),$ (2)

where

Z_n^* = (L × 1) vector of latent variables, where L indicates the number of latent variables;

I_n = (R × 1) vector of R indicators associated with individual n;

$X2_n$ = (1 × M) vector of M regressors;

[4] Integration at the same time (simultaneous estimation) or afterwards (sequential estimation)

γ = (M × 1) vector of unknown structural parameters;

α = (M × 1) vector of unknown measurement parameters;

η_n, υ_n= vectors of error terms;

N = multivariate normal distribution;

Σ_η = L_{diag} variance–covariance matrix (assuming independent latent variables); and

Σ_υ = R_{diag} variance–covariance matrix (assuming independent indicators).

Following the random utility maximization framework, the equations describing the **choice model including the latent variables** are:

Structural: $U_n = X1_n\beta_1 + Z_n^*\beta_2 + \ \varepsilon_n$ (3)

Measurement: $y_{in} = \begin{cases} 1 & if\ U_{in} = max\,\{U_{jn}\} \\ 0 & otherwise \end{cases}$ (4)

where

U_n = (J × 1) vector of indirect utilities, where J indicates the number of alternatives;

$X1_n$ = (J × K) matrix, where K is the number of explanatory variables (attributes of the alternatives and characteristics of the decision maker);

β_1 = (K × 1) vector of unknown parameters;

β_2 = (J × L) matrix of unknown parameters associated with the latent variables;

ε_n = (J × 1) error term vector; and

y_{in} = choice indicator (whether alternative *i* is chosen by individual *n* or not).

Because latent variables are not directly observed, the choice probability (5) is obtained by integrating the conditional choice probability (given the latent variables) over the whole space of Z_n^* . The unconditional choice probability is the joint probability of the indicators y_n and I_n:

$$P\left(y_n, I_n \middle| X_n; \alpha, \beta, \gamma, \Sigma_\varepsilon, \Sigma_\eta, \Sigma_\nu\right) = \int_{Z_n^*} P(y_n | X_n, Z_n^*; \beta, \Sigma_\varepsilon) f(I_n | X_n, Z_n^*; \alpha, \Sigma_\nu) g\left(Z_n^* \middle| X_n; \gamma, \Sigma_\eta\right) dZ_n^* \quad (5)$$

where $X_n = (X1_n, X2_n)$, and f and g are the density functions of the indicators and latent variables, respectively.

We use the full-information maximum likelihood estimator and assume a logit probability kernel ($\varepsilon_n \overset{iid}{\sim} EV1$), which avoids severe bias when transport policies are forecasted or evaluated, as pointed out by Raveau et al. (2010).

3 Context and description of the data

3.1 Vitoria-Gasteiz

Vitoria-Gasteiz is a municipality in northern Spain with a population of 242,924 inhabitants in 2014. The main urban area is the city center which is compact and has an almost completely flat topography. The climate of the area is moderately cold, with damp winters and cool summers.

Local authorities have been implementing sustainable transport policies for about a decade, following the Mobility and Public Space Plan (Council of Vitoria-Gasteiz, 2007) and the Cycling Mobility Master Plan (Council of Vitoria-Gasteiz, 2010). Some of the notable measures include a new joint traffic regulation for cars, bicycles and pedestrians, pedestrianization of streets in the city center, building of bicycle parking and a bicycle lane network, promotion of safe cycling courses, traffic calming in 47 streets of the city center, an increase in the regulated parking area fees, and camera control to car access to the city center.

Modal share of bicycle in Vitoria-Gasteiz has increased sharply, from 3.3% in 2006 to 6.9% in 2011, and to 12.3% in 2014 (Council of Vitoria-Gasteiz, 2015). Nowadays, Vitoria-Gasteiz is the Spanish city with the highest bicycle share, while the non-motorized share (66.7%) is the highest of any European medium–size city (around 250.000 inhabitants), according to EPOMM data (http://www.epomm.eu).

3.2 Household travel survey

3.2.1. General characteristics

Data was collected from a mobility survey that was especially designed to fulfil two requirements:

1. Municipal household travel survey (HTS) to know about the population's trips in a workday of 2014, and comparable with previous municipal HST (2006 and 2011), and
2. Ad-hoc to the TRANSBICI project[5], to collect a large and representative survey with both revealed preference (RP) travel data and perceptions towards utilitarian bicycle use.

Therefore, the survey followed the same methodology than previous municipal HTS and included the specific new questions from the TRANSBICI project. The methodological

[5] Spanish Research Project 'TRANSBICI-Travel behavior analysis for modelling the potential use of bicycle: transition to a cycling city'

characteristics were: telephone household survey; stratified random sampling representative of the municipal residents' mobility (older than 9) by transport zone, age group and gender; from Tuesdays to Fridays, during the spring of 2014[6].

The final database consisted in a sample of 4,192 individuals, representing 218,515 people in the population[7] older than 9. 96.7% of that population travelled the previous day, representing a total of 911,326 trips (4.3 trips per person per day). The modal share of the population trips was: 12.3% by bicycle, 54.4% by walking, 7.6% by public transport (bus and tram), 24.7% by car and motorbike, and 1.1% by other modes (taxi, van or truck, school/company bus or coach, railway, public transport on request, and other individual and collective modes).

3.2.2. Survey questions

Innovation in the questionnaire stems from the inclusion of bicycle perceptual questions in the usual HTS. The bicycle perceptual questions were designed based on the results of the first part of the TRANSBICI research project. TRANSBICI consisted of a panel survey for commuting trips and included multidisciplinary points of view from transport planners, psychologists, and geographers. The panel survey was conducted in Vitoria-Gasteiz in 2012, 2013, and 2014. The survey included and tested indicators –perceptions of cycling factors– following the attitudinal Theory of Planned Behavior (TPB; Ajzen, 1991). The bicycle indicators of the survey were measured using a 7-point Likert scale ranging from completely-disagree/unimportant (+1) to

[6] A 9% of the sample, especially teenagers, could not be surveyed during that period and they were finally collected in October, after the new beginning of the classes

[7] Planned quotes for the stratums (by transport zones, age groups and gender) were not completely reached and weights were required for the analysis of the population results

completely-agree/important (+7), and were of the following four types (see Table 1 for detailed indicators of each type):

I1. Degree of agreement or disagreement towards several factors related to the (possible) bicycle trip for urban mobility.

I2. Degree of limitation provoked by several factors related to the (possible) bicycle trip for urban mobility.

I3. Subjective norm (SN) is one of the TPB predictors of intention, defined as 'the perceived social pressure to engage or not to engage in a behavior'. SN is calculated by multiplying the beliefs linking the behavior (bicycle use for urban mobility in our case) by their corresponding importance.

I4. Global perception of and intentions related to bicycle use.

The usual HTS questions included in the survey were divided in three blocks:

Q1. Individual and household socioeconomic characteristics; availability of transport modes; type of parking at home (see Table 2).

Q2. Trips made the previous day: origin, destination, purpose, mode(s), line(s) and ticket(s) (if public transport used), parking at destination (if car driver, motorbike or van/truck used), infrastructure(s) and parking at destination (if bicycle used). For commuting trips, availability of showers and/or lockers at the work/study place was also included.

Q3. Frequency of bicycle use for different purposes [work, study, non-commuting transport purposes (visiting, going out, going shopping, or going to the doctor/hospital), and doing sport]; and experience riding a bicycle for different purposes.

3.2.3. Modelling sample

Only urban trips (within the continuous populated area composed by the city of Vitoria-Gasteiz and the nearby nucleus) were considered. The main[8] modes considered were bicycle (B), walking (W), public transport (PT) –bus and tram– and car (C) –driver and passenger–, dismissing trips by minor modes such as motorbike, coaches, taxi, and van (2.1% of urban transport trips). None of the geographically excluded trips had chosen bicycle as the main mode. All trip purposes with the exception of 'without destination or going for a walk[9]' were contemplated.

The final sample consisted of 14,406 trips distributed as follows: 13.7% (B); 56.2% (W); 9.3% (PT); 20.8% (C). Weighting those trips to the population (730,044 trips) the distribution is similar: 12.4% (B); 58.8% (W); 8.8% (PT); 20.0% (C). Therefore, and to avoid problems related to choice-based sampling, the sample was used as representative of transport urban trips in Vitoria-Gasteiz.

3.2.4. Descriptive statistics

Table 1 summarizes the average responses to the perceptual questions (types I1, I2, I3 and I4), both for 'cyclists' –people who used bicycle as main mode the previous day– and 'non-cyclists'. Most of the questions were used as indicator variables for the construction of the latent variables. From the first type, cyclists give similar or more positive values to most indicators than other users, except to indicators related to uncomfortable issues such as weather, sweat and traffic stress, or risk issues (theft or accident). Regarding the second type –limitations– cyclists perceive

[8] When more than one mode was used for a trip, a hierarchical organization was applied in order to determine the main mode
[9] Bicycle trips for doing sports were included in this trip purpose

fewer limitations than non-cyclists, but for the use of helmet. For all users, the highest influence (I3) stems from family, followed by friends, and co-workers/study colleagues at the last place.

Table 1 Average values of the perceptual questions about the bicycle

	Cyclists	**Non-cyclists**
(I1) Degree of agreement or disagreement towards: the bicycle use for urban mobility is...		
Environmentally friendly	6.8	6.8
Cheap	6.6	6.6
Healthy	6.2	6.5
Time reliable	6.1	5.3
Independent	6.0	5.6
Flexible	5.9	5.7
Relaxing and fun	5.9	4.8
Quick	5.7	5.2
Theft risky	5.4	5.8
Weather dependent	5.2	6.1
Traffic stressful	5.0	5.7
Accident risky	4.3	5.1
Conflicts with pedestrians	4.2	4.5
Sweat	3.9	4.8

(I2) Degree of limitation provoked by…		
Ride in the traffic	4.9	5.7
No safe parking	4.4	5.2
No cycleways	4.4	4.9
Fix a puncture	3.9	4.9
Maneuvering	3.9	4.6
Long distances	3.7	4.7
Helmet use	3.7	3.4
Hilliness	3.4	4.6
No showers/ranks at destination	3.1	4.2
Physical condition	2.7	4.0
(I3) Subjective norm		
Family	5.0	4.3
Friends	4.6	3.9
Co-workers/Study colleagues	4.1	3.8
(I4) Global perception and intentions		
Global perception about the bicycle use in Vitoria-Gasteiz	5.4	4.9
Intention to increase/start using the bicycle for urban mobility	4.6	3.0
Intention to use a bike-sharing system	3.5	3.1

According to I4, the global perception about bicycle use in Vitoria-Gasteiz is positive (around five), although it is slightly higher for cyclists. Intention to start using the bicycle or increase its use for urban mobility is under the average (3.0) for non-cyclists whereas it increases up to moderately positive (4.6) for current cyclists. Intention to start using a public bike-sharing system is also under the average (around three), with a slightly higher value for cyclists.

The distribution of questions from block Q1 (individual and household socioeconomics, availability of transport modes and type of parking at home), some characteristics related to trips (block Q2) and questions from block Q3 (frequency and experience riding a bicycle for different purposes) are summarized in Table 2. Values are shown for the total sample, and for each of the main transport modes considered: bicycle, walking, public transport and car. Note the following:

1. There were slightly more male bicycle trips (54%) than female bicycle trips (46%). Trips by the rest of modes were mainly made by women. Almost half of bicycle trips (48%) were made by workers.
2. The highest percentage of trips by people in households with children aged below 6 was made by car (26%), whereas the highest percentage of trips by people in households with elderly (> 64) was walking (32%).
3. Household income was similar among households with bicycle trips and car trips.
4. Most non-car trips were made by people with car available, especially among bicycle trips (87%). Only 8% of the total trips were made by people who did not know how to ride a bicycle, especially among people older than 64. The majority of non-bicycle trips were made by people who had a bicycle available, with a percentage that is even higher among car trips.

5. The majority of trips had a non-commuting purpose, especially among the walking trips (81%). Among the commuting trip purposes, going to work produced higher percentages of trips than going to studies, for all modes except for bicycle. Note that most bicycle trips (58%) were made by people younger than 30. Whereas half of the walking trips were made by people older than 49; car and public transport trips were mainly made by people aged between 30 and 49.
6. Bicycle is the second fastest mode (10.9 min) after car (6.7 min). However, the calculated car time does not take into account parking times.
7. The average car distance was 3.7 km, which is reasonable for cycling and actually was ridden by about 18% of cyclists. Therefore, there is a high potential among car users to switch to bicycle.
8. Bicycle use for doing sports are similar between cyclists and non-cyclists, regarding all time frameworks: recently (last week or last month) or generally in the past. However, biclycle use for travel purposes is considerably higher for cyclists. On the one hand, only around 6% and 11% of non-cyclists have recent experience in cycling for commuting (up to 23%-36% regarding generally in the past). On the other hand, about 39% and 55% of non-cyclists have recent experience cycling for non-commuting purposes.

Table 2 Characteristics of the modelling sample (14,406 trips)

Block	Variable		Bicycle (14%)	Walking (56%)	Public transport (9%)	Car (21%)	Total
	Gender	Male	54%	34%	21%	42%	37%
		Female	46%	66%	79%	58%	63%
	Age	<30	58%	18%	28%	16%	24%
		30–49	31%	33%	39%	59%	39%
		>49	11%	49%	34%	24%	37%
	Nationality	Spanish	94%	97%	93%	97%	97%
		Foreigner	6%	3%	7%	3%	3%
	University studies	Yes	27%	29%	24%	33%	29%
		No	73%	71%	76%	67%	71%
Q1	Professional situation	Worker	48%	14%	23%	7%	18%
		Student	35%	36%	37%	64%	42%
		Other	17%	51%	40%	29%	41%
	Type of schedule	Part–time	16%	11%	13%	18%	13%
		Continuous	54%	28%	34%	39%	34%
		Other	30%	61%	53%	43%	52%
	Household situation	Father/Mother	26%	38%	39%	53%	40%
		Son/Daughter	58%	22%	28%	18%	27%
		Other	16%	40%	33%	29%	34%
	People in household	Mean	3.4	2.9	3.0	3.1	3.0

Children < 6 in household	Yes	9%	10%	14%	23%	13%
	No	91%	90%	86%	77%	87%
Elderly > 64 in household	Yes	8%	32%	21%	16%	24%
	No	92%	68%	79%	84%	76%
Household income	< 2.000€	44%	54%	59%	45%	51%
	> 2.000€	56%	46%	41%	55%	49%
	N/A	30%	20%	23%	14%	20%
Car license	Yes	53%	69%	59%	94%	71%
	No	47%	31%	41%	6%	29%
Availability of car	Yes	87%	85%	83%	100%	88%
	No	13%	15%	17%	0%	12%
Household's cars	Mean	1.4	1.2	1.2	1.7	1.3
Car parking at home	Inside a building's parking	73%	77%	73%	74%	75%
	Outside (in the street)	27%	23%	27%	26%	25%
Availability of motorbike	Yes	10%	5%	7%	13%	8%
	No	90%	95%	93%	87%	92%
Motorbike parking at home	Inside a building's parking	88%	92%	95%	96%	93%
	Outside (in the street)	12%	8%	5%	4%	7%
Know to ride a bicycle	Yes	100%	90%	89%	95%	92%
	No	0%	10%	11%	5%	8%
Availability of	Yes	100%	66%	67%	76%	73%

	bicycle	No	0%	34%	33%	24%	27%
	Bicycle parking at home	Inside the household	36%	25%	24%	20%	26%
		Inside the household's building	62%	74%	75%	79%	73%
		Outside (in the street)	2%	1%	0%	0%	1%
Q2	Purpose	Work	23%	11%	19%	32%	18%
		Study	28%	8%	18%	3%	11%
		Non–commuting	49%	81%	63%	65%	72%
	Calculated travel time	Mean (min)	10.9	13.1	19.9	6.7	12.1
	Calculated travel distance	Mean (Km)	2.5	1.0	3.2	3.7	2.0
Q3	Bike use last week (yes)	Commuting	69%	7%	7%	6%	15%
		Non-commuting	82%	50%	49%	39%	52%
		Sport	51%	53%	51%	45%	51%
	Bike use last month (yes)	Commuting	71%	9%	11%	11%	18%
		Non-commuting	87%	55%	55%	46%	58%
		Sport	36%	39%	40%	41%	39%
	Bike use in the past (yes)	Commuting	79%	23%	30%	36%	34%
		Non-commuting	97%	74%	75%	73%	77%
		Sport	95%	96%	96%	97%	96%

4 Empirical analysis

4.1 Modeling

This subsection describes both the construction of the choice set and variables to include in the model –time, cost and latent variables– and the specifications to develop the integrated choice and latent variable (ICLV) model.

4.1.1. Travel times and costs for alternatives modes

Travel times and distances for each trip were calculated for the four alternatives in the model –bicycle, walking, public transport and car–, based on origin and destination addresses from the survey. Since declared travel times are subjective and a preliminary examination showed an important number of errors, calculated times were used both for the chosen and for the alternative modes.

For **car trips** the software Microsoft® MapPoint 2013 and its complement MPMileage 2.4 were used to calculate the quickest point-to-point routes. The complement allowed batching-calculating thousands of routes. Error calculations in a small number of car routes were solved individually using Google Maps. *Fuel costs* were derived using an average cost of fuel per unit of distance (0.00013€/m). This average was estimated considering the distribution of car types in the region –gasoline/gas-oil– (Spanish Directorate-General of Traffic, 2013), average fuel costs in the region in May 2014 (Spanish Ministry of Industry, Energy and Tourism, 2014), and average fuel consumption factors in urban context for Spain, calculated based on the EMEP/EEA guidebook to European emissions (European Monitoring and Evaluation Programme, European Environment Agency, 2013). If trip purpose was different from coming back home and if trip destination was located inside the OTA zone –the regulated parking area in the city center– an approximation of average *parking cost* was assigned (1.35€). This approximation was estimated considering the percentage of car users declaring to pay when they park inside the OTA zone and the average rent cost for a garage in the area.

Public transport, walking, and bicycle travel times and distances were calculated using the web application ‘GEO Vitoria-Gasteiz’, based on the open source platform OpenTripPlanner. Error calculations in a small number of routes by public transport and walking were solved individually using Google Maps. The corresponding errors for some bicycle routes were solved individually with GEO Vitoria-Gasteiz but doing little changes in origin or destination addresses. *Public transport fare* (0.65€) was calculated as an average between the bus and tram trip fare using the BAT card (general transport ticket for a regular public transport use).

4.1.2. Choice set

To define the availability of the four modes for each trip, several rules were established:

1. Bicycle availability: a calculated travel distance less than 9Km; that the individual knows how to ride a bicycle and has a bicycle available; and a moderate (+4) or positive (+5, +6 or +7) intention to start using the bicycle (in a scale from no intention at all (+1) to completely positive intention (+7)).
2. Walking availability: a calculated travel distance less than 7Km.
3. Public transport availability: when there is a possible public tranport route and therfore there is a calculated travel distance.
4. Car availability: a calculated travel distance and that the individual has a car available.

Limits in calculated distances were established considering the corresponding maximum distances done by pedestrians (6.7Km) and cyclists (8.4Km) in the sample. The indicator of intention to start using the bicycle for urban mobility was used as an approximation to discard non-cyclists that might never use the bicycle. According to this, there are 45% of potential cyclists among public transport users, 41% among car users and 37% among pedestrians. The current bicycle share of 13.7% in the sample might thus increase up to 42.0% of urban trips.

4.1.3. Latent variables

Most answers to the perceptual questions about bicycle were used as indicator variables for the construction of latent variables. Several analyses were applied in order get a preliminary structure, namely: explanatory factor analysis (EFA) with SPSS®v20, and a multiple indicators multiple causes (MIMIC) model with AMOS (SPSS® Amos 22.0). Since the focus of this research is ICLV, only the structure of the latent variables is reported but not the details of these analyses. The names of the latent variables were assigned according to their constituent indicators (see Table 3), and considering existing literature.

Table 3 Relationship between latent variables and indicators

Latent variables	**Indicators**	**Survey questions (see 3.2.2.)**
SC - Safety and comfort*	'Low accident risk', 'No traffic stress', 'No conflicts with pedestrian',	**Type I1**: degree of agreement or

Latent variables	Indicators	Survey questions (see 3.2.2.)
	'Weather independent', 'No theft risk' and 'No sweat'	disagreement towards several factors related to the (possible) trip by bicycle for urban mobility
DA - Direct advantages	'Relaxing', 'Quick', 'Time reliable', 'Independent' and 'Flexible'	
A - Awareness	'Cheap', 'Environmentally beneficial' and 'Healthy'	
EF - External facilities*	'Ride separate from traffic', 'Short distance', 'Cycleways', 'Safe parking' and 'Shower at destination'	**Type I2**: degree of limitation provoked by several factors related to the (possible) trip by bicycle for urban mobility
IC - Individual capacities*[1]	'Hilliness', 'Physical condition', 'Fixing a puncture' and 'Manoeuvring'	
SN - Subjective norm	'Family', 'Friends' and 'Co-workers/Study colleagues'	**Type I4**: perceived social pressure to cycle for urban mobility or not

*: The name and scale values of the original indicators were reversed

[1]: The indicator 'Wear a bike helmet' had to be removed.

4.1.4. Model specification

In order to find the best specification that fits the data and explains mode choice, especially bicycle choice, a testing procedure was used: several specifications for the mode-choice model[10], different types of coefficients (generic and alternative-specific), interactions between variables, and progressively including different types of variables [socioeconomic, alternative-varying and bicycle-specific (such as the latent variables)]. This paper presents the results for the best models of each case.

In the discrete choice kernel, utility regressors include alternative-varying variables (travel time and cost, and the interaction between time and age groups), trip-specific variables (purpose), socioeconomics and alternative-specific constants (ASCs) for bicycle, walking and public transport. The joint ICLV model was developed by progressively including the six bicycle latent variables as bicycle-specific attributes. These bicycle latent variables are expected to increase the choice probability of cycling. The initial part of the process consisted in calculating the ICLV models sequentially to minimize estimation times. Sequential estimates[11] were used as starting values for the simultaneous ICLV models.

4.2 Results of model estimation

This subsection presents the estimation results of the choice models. Python Biogeme (Bierlaire, 2003) was used for estimation. When needed, the number of random draws was set to 1,000.

[10] Apart from the multinomial-logit, different specifications for mixed logit were also tried. The only coefficient that showed a statistically significant random variation was the time coefficient for walkers. Therefore, the mixed logit was dismissed and the multinomial-logit was the specification finally chosen

[11] Results of the sequential models are available upon request

Table 4 presents estimates of a multinomial-logit model without latent variables. ASCs show a clear preference for walking trips over car, followed by bicycle and public transport trips. As expected, the generic coefficient for travel cost and the specific ones for travel times –for bicycle (*bicycle-time*), walking (*walking-time*) and motorized modes (*motorized-time*)– are all negative and statistically significant ($p<0.05$). The statistically significant (at least $p<0.10$) interaction[12] coefficients between travel time and age groups were used to derive estimates of the value of time (VOT) for each subgroup. A commuting purpose shows statistically significant coefficients for all modes, positive for bicycle and public transport and negative for walking.

In relation to socioeconomic variables we see that only being a ***man*** and ***having children younger than 6 in the household*** have a significant effect (positive and negative, respectively) on choosing bicycle. Being *foreigner*, *having university studies, having elderly people (older than 64) in the household* or *having a household income less than 2.000€/month* do not appear to be statistically significant for choice of bicycle. Regarding results from the variable age group, the bicycle coefficient for individuals aged 30-49 (compared to older than 49) is negative (-0.96), contradicting what is observed in reality. Other socioeconomic variables from Table 2 were not statistically significant and were not included in the final model.

[12] The interaction term between cost and household income was also tested but resulted statistically insignificant

Table 4 Results for the mode choice model without latent variables

	β	Reference level	MNL Value	P-value
Alternative-specific constants	ASC-bicycle	car	**1.25**	**0.00**
	ASC-walking		**3.49**	**0.00**
	ASC-public transport		**0.34**	**0.03**
Alternative-varying variables	Cost	**-**	**-0.59**	**0.00**
	Bicycle-time	**-**	**-0.17**	**0.00**
	Walking-time	**-**	**-0.13**	**0.00**
	Motorized-time	**-**	**-0.06**	**0.00**
	Bicycle-time*(<30)	>49	**0.07**	**0.00**
	Bicycle-time*(30–49)		**0.04**	**0.02**
	Walking-time*(<30)		**-0.02**	**0.03**
	Walking-time*(30–49)		-0.01	0.05
	Motorized-time*(<30)		**0.04**	**0.00**
	Motorized-time*(30–49)		**-0.03**	**0.03**
Trip-specific variables	Bicycle-commuting	non-commuting	**0.59**	**0.00**
	Walking-commuting		**-0.15**	**0.02**
	Public transport-commuting		**0.34**	**0.00**
Socioeconomic variables	Bicycle-man	woman	**0.30**	**0.00**
	Walking-man		**-0.25**	**0.00**
	Public transport-man		**-1.01**	**0.00**
	Bicycle-(<30)	>49	0.06	0.80
	Walking-(<30)		0.27	0.13
	Public transport-(<30)		**-0.42**	**0.03**
	Bicycle-(30-49)		**-0.96**	**0.00**
	Walking-(30-49)		**-0.53**	**0.00**
	Public transport-(30-49)		-0.14	0.41
	Bicycle-foreigner	Spanish	0.18	0.40
	Walking-foreigner		0.18	0.31
	Public transport-foreigner		**0.53**	**0.01**
	Bicycle-university studies	without university studies	0.06	0.49
	Walking-university studies		-0.10	0.10
	Public transport-university studies		**-0.32**	**0.00**
	Bicycle-children<6	without children<6 in the household	**-0.56**	**0.00**
	Walking-children<6		**-0.43**	**0.00**
	Public transport-children<6		-0.11	0.29
	Bicycle-elderly>64	without elderly>64 in the household	-0.06	0.66
	Walking-elderly>64		**0.21**	**0.01**

	Public transport-elderly>64		0.18	0.07
	Bicycle-(<2,000€)		-0.02	0.80
	Walking-(<2,000€)		**0.13**	**0.04**
	Public transport -(<2.000€)	>2,000€/month in the household	**0.30**	**0.00**
	Bicycle-(N/A)		**0.53**	**0.00**
	Walking-(N/A)		**0.37**	**0.00**
	Public transport-(N/A)		**0.50**	**0.00**
	Final log-likelihood	-9,901.85		
	Likelihood ratio index (ρ)	0.33		

Several **ICLV models** were developed based on the benchmark logit model (but omitting the socioeconomic variables due to low explanatory power for bicycle choice, and because sociodemographics enter the MIMIC model) and testing each of the bicycle latent variables individually. The latent variables 'Awareness' and 'Subjective norm' did not show statistically significant coefficients. The other four latent variables –'Safety and comfort', 'Direct advantages', 'External facilities' and 'Individual capacities'– showed positive and statistically significant coefficients and were tested in pairs in six models. The number of latent variables in the final models –only two– was conditioned by convergence difficulties and very lengthy estimation times.

The six developed models were compared with the consistent Akaike information criterion (CAIC). The model with '***Direct advantages***' (DA) and '***Individual capacities***' (IC) showed the lowest CAIC (516,951), indicating the best fit and it was chosen as the final ICLV model (ICLV-1, see Table 5 and upper part of Figure 1). Considering the high importance of safety issues related to bicycle in the literature, results from a second ICLV including the latent variable 'Safety and Comfort' are also shown. Specifically, the one with the lowest CAIC (570,528), that is, the model with the latent variables '***Safety and Comfort***' (SC) and '***Individual capacities***' (IC) (ICLV-2, see Table 5 and the lower part of Figure 1).

ASCs from Table 5 show a clear preference for walking trips over those by car and a preference for car trips over bicycle trips, reflecting reality. The specific constant for public transport appears to be the only statistically non-significant coefficient. The other coefficients are quite similar to those of the logit model, except the interaction between bicycle-time and (30-49) age group which now inverts its sign.

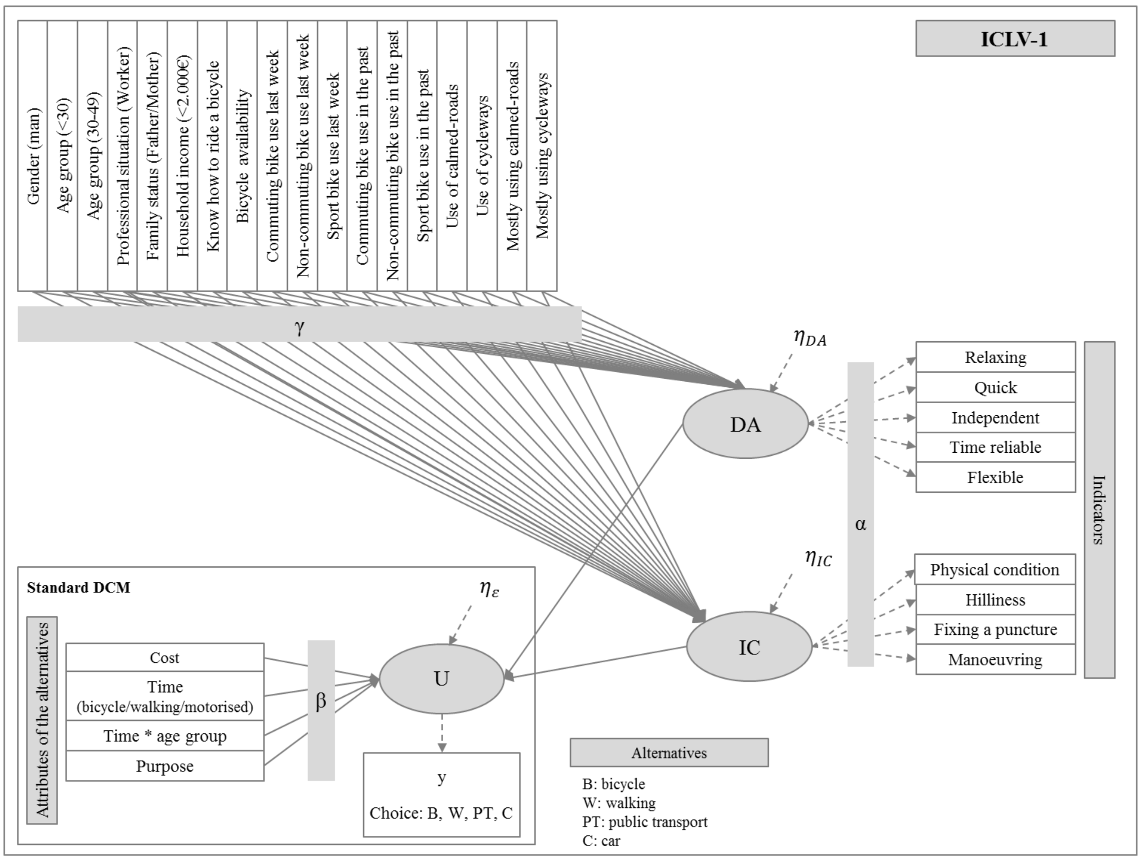
In terms of estimates of the value of time, an aggregated VOT for the motorized modes is around 6€/hour, which is in line with most recent studies reporting VOT in Spanish contexts: values for car in the range from 5.69€/h (dell'Olio et al., 2011) to 9.00€/hour (Salas et al., 2009; Shires and De Jong, 2009); and for bus from 4.75€/h (dell'Olio et al., 2011) to 7.59€/hour (Shires and De Jong, 2009). The interaction coefficients allow disaggregating VOT as shown at the bottom of Table 5. Bicycle and car users follow the same pattern: the lowest willingness to pay is held by the youngest (<30), then the oldest (>49) and finally the highest willingness to pay is held by the medium age group (30-49). Among pedestrians, the highest willingness to pay is also held by the medium age group (30-49)[13], whereas the lowest VOT is held by the oldest (>49). VOTs for bicycle are in line with previous research[14] (Börjesson and Eliasson, 2012; Rodríguez and Joo, 2004; Wardman et al., 1997; Wardman et al., 2007) in that non-motorized users would pay more to reduce their travel time than motorized users. Bicycle VOTs go from 1.9 times (for the 30-49 age group) to 2.7 times (for the >49 age group) the values for car/public transport.

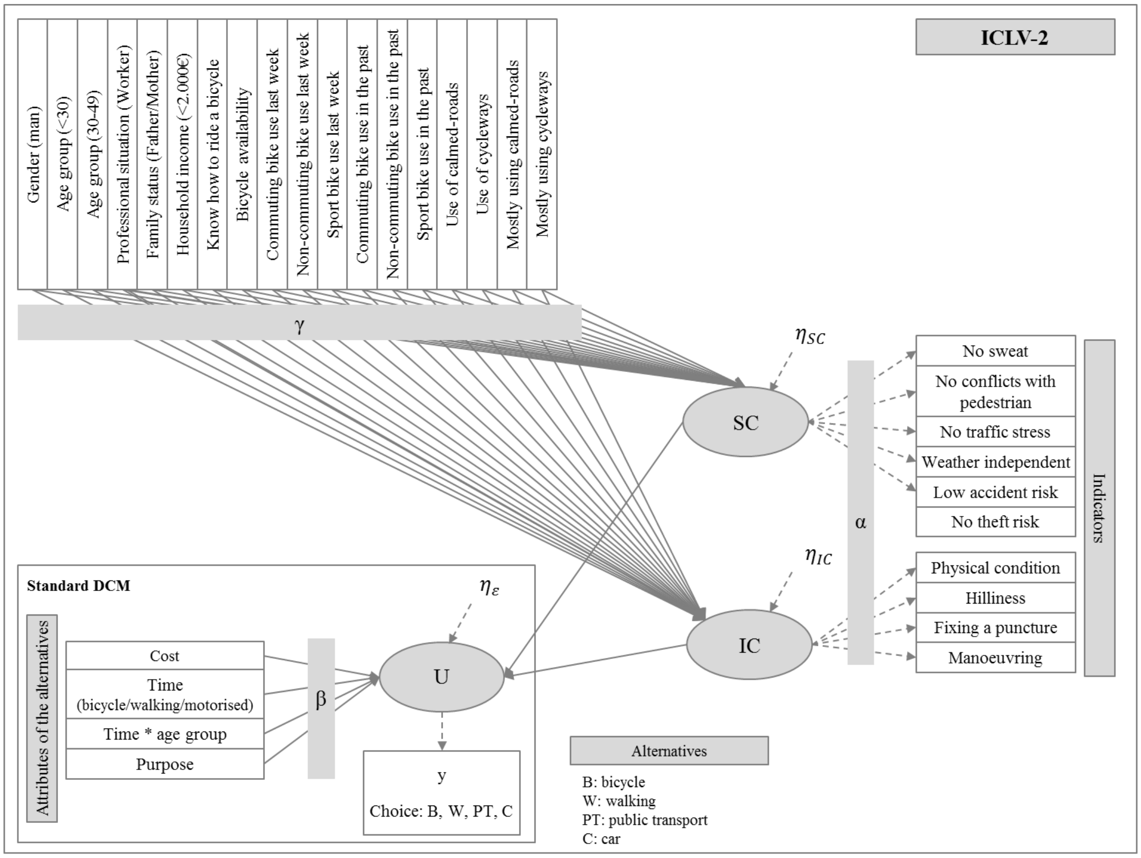
Regarding the **latent variables**, all of them enter significantly into the choice kernel with a positive impact in the bicycle choice probability. In the first model (ICLV-1) the highest

[13] This group is mainly composed by workers and therefore it is the most sensitive group to transport policies implemented in the city
[14] Only in (dell'Olio et al., 2011) bicycle VOTs are lower than the corresponding for motorized modes

influence appears from the latent variable ***Direct advantages*** of cycling –in terms of quickness, time reliability, relax, independency and flexibility– (β=0.93); whereas ***Individual capacities*** of the rider –in terms of overcoming hills, maneuvering, his/her physical condition and the capacity to fix a puncture– appears to play a secondary latent role (β=0.69); however there might be an issue of different scale in both latent constructs.

Figure 1 Modelling frameworks of the two models proposed (ICLV-1 and ICLV-2)





In the second model (ICLV-2), both ***Individual capacities*** and ***Safety and comfort*** –in terms of low accident and theft risk, no traffic stress, no conflicts with pedestrians, no sweat and weather independent– appear to be influencing in a similar magnitude (β=0.57 and β=0.59, respectively). The model results show that the latent variables are the key drivers of bicycle choice, although travel time is still statistically significant.

Table 5 Mode choice model results

β		Reference level	ICLV-1 Value	ICLV-1 P-value	ICLV-2 Value	ICLV-2 P-value
Alternative-specific constants	ASC-bicycle	car	**-7.25**	**0.00**	**-3.65**	**0.00**
	ASC-walking		**3.27**	**0.00**	**3.28**	**0.00**
	ASC-public transport		0.05	0.48	0.05	0.52
Alternative-varying variables	Cost	-	**-0.63**	**0.00**	**-0.62**	**0.00**
	Bicycle-time	-	**-0.13**	**0.00**	**-0.12**	**0.00**
	Walking-time	-	**-0.12**	**0.00**	**-0.12**	**0.00**
	Motorized-time	-	**-0.05**	**0.00**	**-0.05**	**0.00**
	Bicycle-time*(<30)	>49	**0.06**	**0.00**	**0.04**	**0.00**
	Bicycle-time*(30-49)		**-0.05**	**0.00**	**-0.05**	**0.00**
	Walking-time*(<30)		**-0.01**	**0.00**	**-0.01**	**0.00**
	Walking-time*(30-49)		**-0.04**	**0.00**	**-0.04**	**0.00**
	Motorized-time*(<30)		**0.02**	**0.00**	**0.02**	**0.00**
	Motorized-time*(30-49)		**-0.04**	**0.00**	**-0.04**	**0.00**
Trip-specific variables	Bicycle-commuting	non-commuting	**0.80**	**0.00**	**0.61**	**0.00**
	Walking-commuting		**-0.14**	**0.03**	**-0.14**	**0.03**
	Public transport-commuting		**0.28**	**0.00**	**0.28**	**0.00**
Bicycle-specific latent variables	Direct advantages	-	**0.93**	**0.00**	-	-
	Direct advantages - variance	-	**0.91**	**0.00**	-	-
	Individual capacities	-	**0.69**	**0.00**	**0.57**	**0.00**
	Individual capacities - variance	-	**1.00**	**0.00**	**1.00**	**0.00**
	Safety and comfort	-	-	-	**0.59**	**0.00**
	Safety and comfort - variance	-	-	-	**0.87**	**0.00**
Number of iterations			118		104	
Final log–likelihood			-258,094.66		-284,872.93	
Likelihood ratio index (ρ)			0.90		0.74	

Values of time (€/h)	**Age groups < 30**	**30–49**	**> 49**	**Age groups < 30**	**30–49**	**> 49**
Bicycle	6.66	16.77	12.15	6.92	15.54	11.17
Walking	12.75	14.96	11.39	12.75	15.19	11.56
Motorized	2.77	7.93	4.58	2.67	7.98	4.60

The introduction of latent variables provided apparent better goodness of fit; however, we note the discussion in Vij and Walker (2016) about model fit. Moreover, Table 6 and Table 7 also show reliability measures of the latent variable model part –referred to the corresponding sequential models– throughout the Squared Multiple Correlation (R^2).

Table 6 presents the estimation results from the structural models. Being the family head (father/mother) or having low household income recognize the **direct benefits** of urban cycling. Furthermore, men, medium-aged people and workers do not recognize those direct benefits. With respect to **individual capacities**, men and people younger than 49 (especially the youngest) recognize a higher individual cycling capacity. However, a low household income shows the opposite result. Finally, **safety and comfort** characteristics are positively recognized by men, workers and parents. Knowing how to ride a bicycle and having one available significantly explain the three latent variables.

The use of cycleways only appears to statistically explain the individuals' capacity for cycling; and the majority of use of calmed-roads is explaining the individuals' perception of bicycle safety and comfort.

Both the most recent (last week) and past (some time in life) cycling experience, for different purposes, are explaining the latent variables. Specifically, commuting bike use last week for safety and comfort; commuting bike use in the past for direct advantages and individual capacities; and non-commuting and sport bike use last week or in the past for all three latent variables.

Table 6 Structural models results

Block	Latent variables / Variables	ICLV-1: Direct advantages, Estimates (γ)	P-value	ICLV-1: Individual capacities, Estimates (γ)	P-value	ICLV-2: Individual capacities, Estimates (γ)	P-value	ICLV-2: Safety and comfort, Estimates (γ)	P-value
Q1	Gender (Male)	-0.06	0.00	0.76	0.00	0.75	0.00	0.28	0.00
	Age < 30 / > 49	-0.00	1.00	0.32	0.00	0.33	0.00	-0.16	0.00
	Age 30-49 / > 49	-0.11	0.00	0.27	0.00	0.27	0.00	-0.20	0.00
	Professional situation (Worker)	-0.20	0.00	-0.04	0.35	-0.07	0.15	0.36	0.00
	Family status (Father/Mother)	0.26	0.00	-0.04	0.15	-0.04	0.13	0.06	0.01
	Income <2,000€/month in the household	0.33	0.00	-0.10	0.00	-0.10	0.00	-0.08	0.00
	Know how to ride a bike (Yes)	1.42	0.00	0.84	0.00	0.85	0.00	0.93	0.00
	Bicycle availability (Yes)	1.36	0.00	1.04	0.00	1.04	0.00	0.94	0.00
Q2	Use of calmed-roads (Yes)	-0.08	0.10	0.11	0.13	0.11	0.16	-0.11	0.13
	Use of cycleways (Yes)	0.00	0.96	0.28	0.00	0.26	0.00	0.10	0.16
	Mostly using calmed-roads (Yes)	0.10	0.15	0.10	0.42	0.09	0.46	0.23	0.03
	Mostly using cycleways (Yes)	0.01	0.92	-0.04	0.64	-0.04	0.59	0.08	0.25
Q3	Commuting bike use last week (Yes)	0.03	0.34	0.05	0.18	0.04	0.28	0.16	0.00
	Non-commuting bike use last week (Yes)	0.68	0.00	0.44	0.00	0.44	0.00	0.49	0.00
	Sport bike use last week (Yes)	0.39	0.00	0.22	0.00	0.22	0.00	0.12	0.00
	Commuting bike use in the past (Yes)	0.05	0.03	0.16	0.00	0.17	0.00	-0.02	0.45
	Non-commuting bike use in the past (Yes)	0.59	0.00	0.37	0.00	0.38	0.00	0.37	0.00
	Sport bike use in the past (Yes)	1.81	0.00	0.97	0.00	0.97	0.00	1.04	0.00
Squared Multiple Correlation (R^2)*		0.15		0.34		0.34		0.23	

* Refers to the sequential model.

Table 7 Measurement models results

Indicators	Latent variable	Estimates (α)	P-value	R^2	Variances	P-value
Quick	Direct advantages (ICLV-1)	1.00	na	0.29	1.35	0.00
Time reliable		1.03	0.00	0.22	1.26	0.00
Relaxing		0.96	0.00	0.41	1.53	0.00
Independent		1.08	0.00	0.20	1.50	0.00
Flexible		1.08	0.00	0.14	1.36	0.00
Hilliness	Individual capacities (ICLV-1) & (ICLV-2)	1.00	na	0.50	1.38	0.00
Maneuvering		0.95	0.00	0.29	1.50	0.00
Physical condition		1.16	0.00	0.43	1.51	0.00
Fixing a puncture		0.90	0.00	0.32	1.77	0.00
No sweat	Safety and comfort (ICLV-2)	1.00	na	0.15	1.63	0.00
No conflicts with pedestrian		1.07	0.00	0.21	1.74	0.00
No traffic stress		0.76	0.00	0.29	1.55	0.00
Weather independent		0.63	0.00	0.19	1.2	0.00
Low accident risk		0.94	0.00	0.40	1.28	0.00
No theft risk		0.69	0.00	0.15	1.39	0.00

* Refers to the sequential model.

With regard to the measurement part (table 7), normalization was set by fixing the coefficient for the indicators 'Quick', 'Hilliness' and 'No sweat' to 1. The other coefficients and variances from the measurement equations are shown in Table 7. First, a higher perception of the **direct benefits** of cycling lead to recognize mostly the independence, flexibility, time reliability and quickness that bicycle provides, and finally its relaxing character. Second, the latent variable **Individual capacities** is mainly explaining the perceptions of the own physical condition and of the ability to overcome hills. To a lesser extent this latent variable is explaining the perception of maneuvering when riding the bicycle and the ability to fix punctures. Third, the latent variable **Safety and comfort** mainly explains the safety perceptions of the bicycle related to pedestrians or motorized traffic. The lowest coefficient refers to weather effects.

4.3 Forecasting the impact of policy measures

The final ICLV model (ICLV-1)[15], intended to properly explain cycling demand, was used as a tool to predict the impact on choice behavior of different types of potential transport policies. Four policies focused on time and cost –some fostering bicycle use and others punishing car use– and on the choice set. Some of these measures continue the initiated path of the local transport policy, and others have been successfully implemented in other cities. Three other measures were related to the bicycle latent variables, trying to mimic change in perceptions[16]. These three soft measures generalize the already implemented measure of promotion of safe cycling courses.

[15] This was chosen as the final model because it showed the lowest CAIC and therefore it had the best fit

[16] For forecasting there is no need for the latent variable measurement model

Moreover, the combined effect of measures fostering bicycle use and others punishing car use were also tested. The individual measures were the following:

- **Related to time and cost and to the choice set (hard measures)**

  1. **Cycling network extension**. Provision of cycleways and improvement of the bicycle network connectivity so that bicycle travel times in the mode choice model are reduced by 10% in all trips. This would mean the continuation of a measure that has been progressively implemented since 2007.

  2. **Extension of the traffic calming area**, so that car travel times in the mode choice model increase by 10% in all trips. This would mean the extension of the current traffic calming area (transport zones 1-6 and 9) to the entire city center (transport zones 1-9, 15-21 and 25-27).

  3. **Urban toll to cars** travelling within the central part of the city. This would increase travel cost for car trips with origin or destination inside the charging area in 1.80€[17]. This measure has been implemented in several European cities to reduce traffic congestion and to improve air quality and reduce noise.

  4. **Public bike-share system**, so that the whole population would have access to bicycle. This scenario is not a true forecasting exercise, since alternatives' attributes do not change. However, it is an approximation exercise to test a popular cycling measure in Spain, which was planned to be implemented in Vitoria-Gasteiz in 2012, but was finally rejected because of financial problems.

[17] According to an average of charging schemes in non-capital cities in Europe

- **Related to the bicycle latent variables (soft measures)**

5. **Commuting cycling programs in companies** to experience bicycle for commuting ('Bike to work days'). This would change 'Commuting bike use in the past' in the structural model to Yes in all trips made by workers.
6. **Non-commuting cycling programs for the general public** to experience urban cycling. This would change 'Non-commuting bike use in the past' in the structural model to Yes in all trips.
7. **Sport cycling programs for the general public** to promote recreational cycling. This would change 'Sport bike use in the past' in the structural model to Yes in all trips.

Table 8 presents the observed market shares, the estimated baseline scenario (simulated market shares) and the market share changes predicted by the final ICLV-1 model for each of the transport measures and some of their combinations. Obviously, the results show that under all initiatives the bicycle mode share increases. Moreover, car mode share decreases in all scenarios.

The proposed **urban toll to cars** produces the highest increases in bicycle share and the highest decreases in car share –without taking into account the public bike-share system– with the additional positive effect of increasing walking and public transport shares. The **cycling network extension** would be the second most preferred measure, according to the bicycle share increase. The **extension of the traffic calming** would produce a similar car share decrease than the aforementioned measure, although the impact on bicycle share would be very limited.

The effectiveness of the measures related to the latent variables is more limited than the previous ones, especially the one related to programs such as the 'Bike to work days'. **Non-commuting**

cycling programs for the general public would be most effective for bicycle share, drawing mainly from public transport and car trips. The combined effect of the measure fostering bicycle use –**cycling network extension (1)**–and the preferred measure punishing car use –**urban toll to cars (3)**– shows a moderately increase in bicycle share and a moderately decrease in the rest of mode shares with respect to scenario (8). The addition of measure (6) tinges previous results in the same direction.

Table 8 Policy scenarios and market shares with ICLV-1 model: predicted values and changes

Policy measures			**Bicycle**	**Walking**	**Public transport**	**Car**
Observed market share		value	13.69%	56.21%	9.27%	20.84%
0. Baseline scenario		**value**	**14.48%**	**55.80%**	**9.14%**	**20.58%**
Related to time and cost and to the choice set	1. Cycling network extension (↓ 10% bicycle travel time)	**value**	**15.25%**	**55.52%**	**9.01%**	**20.22%**
		change$_{0-1}$	5.33%	-0.49%	-1.52%	-1.74%
	2. Extension of the traffic calming (↑ 10% car travel time)	**value**	**14.56%**	**55.99%**	**9.27%**	**20.18%**
		change$_{0-2}$	0.52%	0.35%	1.37%	-1.92%
	3. Urban toll to cars	**value**	**15.52%**	**59.58%**	**11.08%**	**13.82%**
		change$_{0-3}$	7.14%	6.79%	21.21%	-32.86%
	4. Public bike-share system	**value**	**25.86%**	**50.25%**	**7.32%**	**16.57%**
		change$_{0-4}$	78.55%	-9.93%	-20.00%	-19.46%
Related to the bicycle latent variables	5. Bike to work days	**value**	**14.54%**	**55.76%**	**9.14%**	**20.57%**
		change$_{0-5}$	0.41%	-0.07%	-0.09%	-0.06%
	6. Non-commuting cycling programs for the general public	**value**	**15.04%**	**55.49%**	**9.06%**	**20.40%**
		change$_{0-6}$	3.88%	-0.54%	-0.88%	-0.87%
	7. Sport cycling programs for the general public	**value**	**14.96%**	**55.52%**	**9.06%**	**20.46%**
		change$_{0-7}$	3.30%	-0.50%	-0.89%	-0.58%
Combined scenarios	Cycling network extension (1) & Urban toll to cars (3)	**value**	**16.31%**	**59.25%**	**10.90%**	**13.54%**
		change$_{0-13}$	12.60%	6.20%	19.15%	-34.19%
	Cycling network extension (1) & Urban toll to cars (3) & Non-commuting cycling programs for the general public (6)	**value**	**16.93%**	**58.89%**	**10.78%**	**13.40%**
		change$_{0-136}$	16.92%	5.550%	17.84%	-34.87%

5 Conclusions

The empirical output of this paper is to present a jointly estimated ICLV model focused on urban cycling, including several bicycle latent variables (LVs) and considering forecasting issues to test the effects of several potential transport measures. Unlike other ICLV models with bicycle latent variables, data for the model building comes from an important RP sample (14,406 trips) representative of the urban mobility in the city of Vitoria-Gasteiz (Spain) in 2014. Perceptual indicators towards the bicycle use for urban mobility were used to define six bicycle LVs that were introduced individually and directly into the bicycle utility function.

The study revealed the importance of '**Direct advantages**', '**Safety and comfort**', '**External facilities**' and '**Individual capacities**' as latent constructs that played a significant role in the bicycle choice process. Although with a lower number of indicators, these LVs were also identified and appeared to be significant in previous applications. However, latent 'Awareness' and 'Subjective norm' did not show statistically significant coefficients.

The final ICLV model including two[18] of the significant latent variables –'**Direct advantages**' and '**Individual capacities**'– was chosen to analyze the potential effect of differing transport policies. In particular, the inclusion of the bicycle latent variables allowed testing three real-world soft measures related to bicycle experience that are hard to analyze in standard discrete choice models. These soft measures were expected to change some of the causal variables

[18] The number of latent variables in the final models –only two– was conditioned by convergence difficulties and very lengthy estimation times, related with the simultaneous estimation.

defining the latent variables, which in turn would change the latent variable values –through the structural equations in the MIMIC model– and finally affect mode choice behavior. This method is different from the one referring to a certain change in a latent variable directly targeted by a policy intervention (which is done in Maldonado-Hinarejos et al., 2014) and a less problematic way of using ICLV for the derivation of transport policies, as pointed by Chorus and Kroesen (2014).

Forecasting results showed that urban toll to cars would be the preferred measure, not only because it produced the highest increase in bicycle share but also because this measure also decreased car share the most. Moreover, the effectiveness of the soft measures related to the latent variables was more limited than those related to time and cost. The tested combination of hard and soft measures –some fostering bicycle use and others punishing car use– shows an additive increase in bicycle share, drawing trips only from car (c.f. Maldonado-Hinarejos et al., 2014)

Although the soft measures tested were affected by estimates that were larger than those of time and cost, the limited effectiveness of the soft measures might be explained by weak structural relationships in the MIMIC model (see R^2 values in Table 6). Therefore, future research might focus on finding better-supported SEM to improve forecasting power with latent variables. In this regard, it is also worth mentioning the absence of a methodology to test soft measures affecting indicators from the latent variable measurement model, since for forecasting only structural equations in the MIMIC model are needed. Finally, although the simultaneous approach is preferred over the sequential approach, the very lengthy estimation times of the models associated to full information MLE, was an actual limitation in the present study.

Finally, it is also worth mentioning that future research with the present data might develop specific models for different trips purposes, focusing especially on commuting trips and comparing the strength of other latent variables, such as cycling habit for non-commuting trips, with those in previous bicycle commuting studies (see e.g. Heinen et al., 2011; Muñoz et al., 2013; Muñoz et al. 2016). In fact, future research will also look into ways of better addressing the scrutiny that the ICLV approach has received lately (Chorus and Kroesen, 2014; Mariel and Meyerhoff, 2016; Vij and Walker, 2016).

6 Acknowledgements

The authors gratefully acknowledge the funding awarded to Begoña Muñoz by the Spanish Science and Innovation Ministry under Grant BES–2011–049533 associated with the National Research Project: "TRANSBICI–Travel behavior analysis for modelling the potential use of bicycle: transition to a cycling city". Ricardo Daziano acknowledges support from the Center for Transportation, Environment, and Community Health (CTECH, US DOT-69A3551747119).

7 References

Ajzen, I. (1991) The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50, 179-211, DOI: 10.1016/0749-5978(91)90020-T.

Alegre, L. and Carbonell, F. (2015) Cycling and Pedestrian Policies in Cities and Metropolitan Areas Worldwide. *Routes/Roads-PIARC Magazine*, 76-83. Retrieved from http://www.piarc.org/en/order-library/22660-en-Cycling%20and%20pedestrian%20policies%20in%20cities%20and%20metropolitan%20areas%20worldwide.htm Last accessed: 27/10/2017.

Bartholomew, D.J., Knott, M., and Moustaki, I. (2011) *Latent Variable Models and Factor Analysis: A Unified Approach.* John Wiley & Sons, Chichester (England).

Ben-Akiva, M., McFadden, D., Gärling, T., Gopinath, D., Walker, J., Bolduc, D., Börsch-Supan, A., Delquié, P., Larichev, O., and Morikawa, T. (1999) Extended framework for modeling choice behavior. *Marketing Letters*, 10, 187-203, DOI: 10.1023/A:1008046730291.

Ben-Akiva, M., McFadden, D., Train, K., Walker, J., Bhat, C., Bierlaire, M., Bolduc, D., Börsch-Supan, A., Brownstone, D., Bunch, D.S., Daly, A., De Palma, A., Gopinath, D., Karlstrom, A., and Munizaga, M.A. (2002a) Hybrid choice models: Progress and challenges. *Marketing Letters*, 13, 163-175, DOI: 10.1023/A:1020254301302.

Ben-Akiva, M., Walker, J., Bernardino, A.T., Gopinath, D.A., Morikawa, T., and Polydoropoulou, A. (2002b) Integration of choice and latent variable models. ed H.S. Mahmassani, pp. 431-470. Pergamon, Amsterdam (The Netherlands), New York, NY (USA).

Bierlaire, M. (2003). BIOGEME: A free package for the estimation of discrete choice models, *Proceedings of the 3rd Swiss Transportation Research Conference*, Ascona (Switzerland).

Börjesson, M. and Eliasson, J. (2012) The value of time and external benefits in bicycle appraisal. *Transportation Research Part A: Policy and Practice*, 46, 673-683, DOI: 10.1016/j.tra.2012.01.006.

Chorus, C.G. and Kroesen, M. (2014) On the (im-) possibility of deriving transport policy implications from hybrid choice models. *Transport Policy*, 36, 217-222, DOI: 10.1016/j.tranpol.2014.09.001.

Council of Vitoria-Gasteiz (2007) *Sustainable Mobility and Public Space Plan (Plan de Movilidad y Espacio Público)*. Vitoria-Gasteiz (Spain). Retrieved from Council of Vitoria-Gasteiz website: http://www.vitoria-gasteiz.org/we001/was/we001Action.do?aplicacion=wb021&tabla=contenido&idioma=es&uid=_5e2b2877_120d224e518__7fe7 Last accessed: 27/10/2017.

Council of Vitoria-Gasteiz (2010) *Cycle Mobility Master Plan (Plan Director de Movilidad Ciclista)*. Vitoria-Gasteiz (Spain). Retrieved from Council of Vitoria-Gasteiz website: http://www.vitoria-gasteiz.org/wb021/http/contenidosEstaticos/adjuntos/es/45/63/34563.pdf Last accessed: 27/10/2017.

Council of Vitoria-Gasteiz (2015) *2014 Vitoria-Gasteiz Mobility Survey - Analysis of the Results.* Vitoria-Gasteiz (Spain). Retrieved from Council of Vitoria-Gasteiz website: http://www.vitoria-gasteiz.org/wb021/http/contenidosEstaticos/adjuntos/es/04/47/60447.pdf Last accessed: 27/10/2017.

dell'Olio, L., Ibeas, A., and Moura, J.L. (2011) Implementing bike-sharing systems. *Proceedings of the ICE-Municipal Engineer*, 164, 89-101, DOI: 10.1680/muen.2011.164.2.89.

European Monitoring and Evaluation Programme, European Environment Agency (2013) *EMEP/EEA Air Pollutant Emission Inventory Guidebook—2013.* Luxembourg. Retrieved from Office for Official Publications of the European Union website: http://www.eea.europa.eu/publications/emep-eea-guidebook-2013 Last accessed: 27/10/2017.

Fernández-Heredia, A., Jara-Díaz, S., and Monzon, A. (2016) Modelling bicycle use intention: the role of perceptions. *Transportation*, 43, 1-23, DOI: 10.1007/s11116-014-9559-9.

Habib, K.N., Mann, J., Mahmoud, M., and 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, DOI: 10.1016/j.tra.2014.09.012.

Heinen, E., Maat, K., and 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, 102-109, DOI: 10.1016/j.trd.2010.08.010.

Kamargianni, M., Bhat, C.R., Polydoropoulou, A., and Dubey, S. (2015) Investigating the Subjective and Objective Factors Influencing Teenagers' School Travel Mode Choice–An Integrated Choice and Latent Variable Model. *Transportation Research Part A: Policy and Practice*, 78, 473-488, DOI: 10.1016/j.tra.2015.06.011.

Kamargianni, M. and Polydoropoulou, A. (2013) Hybrid choice model to investigate effects of teenagers' attitudes toward walking and cycling on mode choice behavior. *Transportation Research Record: Journal of the Transportation Research Board*, 2382, 151-161, DOI: 10.3141/2382-17.

Maldonado-Hinarejos, M.R., Sivakumar, A., and Polak, J.W. (2014) Exploring the Role of Individual Attitudes and Perceptions in Predicting the Demand for Cycling: A Hybrid Choice Modelling Approach. *Transportation*, 41, 1287-1304, DOI: 10.1007/s11116-014-9551-4.

Mariel, P., and Meyerhoff, J. (2016). Hybrid discrete choice models: gained insights versus increasing effort. *Science of the Total Environment*, 568, 433-443.

Muñoz, B., Monzon, A., and Daziano, R.A. (2016) The Increasing Role of Latent Variables in Modelling Bicycle Mode Choice. *Transport Reviews. A Transnational Transdisciplinary Journal*, 36(6), 737-771, DOI: 10.1080/01441647.2016.1162874.

Muñoz, B., Monzon, A., and Lois, D. (2013) Cycling habits and other psychological variables affecting commuting by bicycle in Madrid, Spain. *Transportation Research Record: Journal of the Transportation Research Board,* 2382, 1-9, DOI: 10.3141/2382-01.

Muñoz, B., Monzon, A., and Lopez, E. (2016) Transition to a cyclable city: latent variables affecting bicycle commuting. *Transportation Research Part A: Policy and Practice,* 84, 4-17, DOI: 10.1016/j.tra.2015.10.006.

Pucher, J., Dill, J., and Handy, S. (2010) Infrastructure, programs, and policies to increase bicycling: an international review. *Preventive Medicine,* 50, S106-S125, DOI: 10.1016/j.ypmed.2009.07.028.

Raveau, S., Alvarez-Daziano, R., Yáñez, M.F., Bolduc, D., and de Dios Ortúzar, J. (2010) Sequential and simultaneous estimation of hybrid discrete choice models. *Transportation Research Record: Journal of the Transportation Research Board,* 2156, 131-139, DOI: 10.3141/2156-15.

Rodríguez, D.A. and Joo, J. (2004) The relationship between non-motorized mode choice and the local physical environment. *Transportation Research Part D: Transport and Environment,* 9, 151-173, DOI: 10.1016/j.trd.2003.11.001.

Salas, M., Robusté, F., and Saurí, S. (2009) Impact of two-part pricing scheme on social welfare for congested metropolitan expressway networks. *Paper presented at the meeting of Transportation Research Board (TRB) 88th Annual Meeting,* Washington, DC (USA).

Shires, J.D. and De Jong, G.C. (2009) An international meta-analysis of values of travel time savings. *Evaluation and Program Planning*, 32, 315-325, DOI: 10.1016/j.evalprogplan.2009.06.010.

Spanish Directorate-General of Traffic (2013) *Fleet of vehicles by typology, autonomous community, region and fuel type (Parque De Vehículos Por Tipos, Ccaa, Provincias y Carburantes).* Madrid (Spain). Retrieved from Spanish Directorate-General of Traffic website: www.dgt.es/Galerias/seguridad-vial/estadisticas-e-indicadores/parque-vehiculos/prov-y-tipo-carburante/12_Parque_Vehiculos-tipo-carburante_13.xls Last accessed: 27/10/2017.

Spanish Ministry of Industry, Energy and Tourism. (2014) *Fuel and Combustible costs. May 2014 Data (Precios De Carburantes y Combustibles. Datos De Mayo De 2014).* Madrid (Spain). Retrieved from Spanish Ministry of Industry, Energy and Tourism website: http://www.minetur.gob.es/energia/petroleo/Precios/Informes/InformesMensuales/2014/mayo2014.pdf Last accessed: 27/10/2017.

Vij, A., and 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. and Ben-Akiva, M. (2002) Generalized random utility model. *Mathematical Social Sciences*, 43, 303-343, DOI: 10.1016/S0165-4896(02)00023-9.

Wardman, M., Hatfield, R., and Page, M. (1997) The UK national cycling strategy: can improved facilities meet the targets? *Transport Policy*, 4, 123-133, DOI: 10.1016/S0967-070X(97)00011-5.

Wardman, M., Tight, M., and Page, M. (2007) Factors influencing the propensity to cycle to work. *Transportation Research Part A: Policy and Practice*, 41, 339-350, DOI: 10.1016/j.tra.2006.09.011.

Yang, L., Sahlqvist, S., McMinn, A., Griffin, S.J., and Ogilvie, D. (2010) Interventions to promote cycling: systematic review. *BMJ (Clinical Research Ed.)*, 341, c5293, DOI: 10.1136/bmj.c5293.