Università degli Studi di Cagliari

PHD DEGREE

Civil Engineering and Architecture

Cycle XXXII

Construction and estimation of discrete choice models for assessing and forecasting the effects of sustainable mobility strategies

Scientific Disciplinary Sector

ICAR/05

PhD Student: Francesco Piras

Coordinator of the PhD Programme Prof. Ivan Blečić

Supervisor Prof. Italo Meloni

Final exam. Academic Year 2018 – 2019

Thesis defence: January-February 2020 Session

ABSTRACT

The current transport system, based on the use of vehicles fueled almost exclusively by petroleum products, generates non-negligible social, environmental and economic impacts. In Italy, the increase in the levels of PM_{10} and $PM_{2.5}$ between 2014 and 2018 has been attributed primarily to an increase in private car ownership and vehicle miles traveled by car. To overcome the problems associated with the current system, in recent years, particular attention has been paid to sustainable urban mobility at different levels. Governmental bodies, policy makers, transportation operators and academic researchers have focused on strategies and policies to reduce automobile travel and induce a modal shift in travel behavior. In particular, there has been increasing interest in soft measures which use information and communication tools to heighten people's awareness and encourage them to change their travel behavior.

The primary objective of this thesis is to explore the methodological processes able to guarantee that sustainable mobility policies seeking to reduce car use are accompanied by a quantitative assessment of the effects that this generates in the transport context. In particular, we focused on the role and weight of psycho-attitudinal variables (attitudes, perceptions, habits, *etc.*) in the process of choosing a sustainable mode of transport.

The first part of the thesis presents the findings of a study focusing on unraveling the linkage among psycho-attitudinal factors related to bike use and the choice to cycle. In doing so, we constructed and estimated different econometric models (Integrated Choice Latent Variable model, GHDM model, Multivariate ordered probit) each aimed at shedding light on those aspects overlooked in the research. The context of the study were the urban areas of Cagliari and Sassari, main cities in Sardinia (Italy), where, despite the implementation of policies supporting bike use, cycling levels for commuting trips are still low. Our modeling estimation results reinforce the idea that promoting cycling through the implementation of awareness campaigns and educational programs, intended to improve peoples' perceptions of the bike mode, can persuade them to consider the bike as an alternative means of transport to private motorized vehicles. Further, investments aimed at supporting use of the bike for leisure (*e.g.* cycle routes) may increase the number of people who choose to use the bike as an alternative means of transport for commuting or shopping.

The second part of the work attempted to assess the short-term effect on travel mode choice of introducing a new sustainable form of transport into the choice set (hard measure) when implementing a VTBC program (soft measure). The transport context chosen for this experiment is a corridor linking the city center of Cagliari (Italy) to a university/hospital complex, where a new light rail route went into service in February 2015. For assessing changes in travel behavior in the short term, the modal shares observed in the first and second wave surveys were compared. Our results show that the combination of hard and soft measures achieved a change in travel behavior of 34%, when the measure is not personalized, and 46% with the VTBC program.

Finally, we evaluated the long-term effects of these measures and we investigated if any changes in the psycho-attitudinal factors and/or in socio-economic characteristics exist after implementation of those measures. In particular, the objective of the study is to analyze whether these changes in individual characteristics are able to affect mode choice from a modeling perspective, through the specification and estimation of Integrated Choice Latent Variable models that use, for the same sample, the data collected for these two moments in time. Our results indicate that psycho-attitudinal variables were not significantly different. over waves, showing that the impact of the psychological construct remained stable over time, even after the introduction of the new light rail. Additionally, we found some evidence that the variables that explain the psycho-attitudinal variables could change over time.

AKNOWLEDGMENTS

I would like to express my deepest gratitude to my thesis supervisor Prof. Italo Meloni. I am grateful to him for his precious advice in various situations and for his constant encouragement.

I also want to thank Dr. Eleonora Sottile. I enjoyed a lot working with her and I would like to thank her for her advice and her constructive criticisms which helped me to identify and focus the problems.

I am grateful to many staff members of the department, especially to Catherine Mann for her friendly assistance.

I would also like to thank Prof. Elisabetta Cherchi and Dr. Chandra R. Bhat for their precious support during my visiting period at Newcastle University and UT Austin.

Next, I would like to thank my family who supported me during the three last years.

Finally, I want to thank ARST Sardegna and Regione Autonoma della Sardegna for funding this dissertation.

INDEX

INTRODUCTION

The demand for urban mobility has grown substantially over the last few decades, to the extent that projections based on data for 2000 predict a three- to fourfold increase in passenger miles traveled by the year 2050 (Un-Habitat, 2013). The reason for this is the rapid population growth, especially the urban population as a consequence of urbanization. This, in turn, has led to the haphazard and uncontrolled development of cities and towns with major negative impacts on the environment. The proliferation of road infrastructure, the increase in the distance traveled and number of trips have created a strong dependence on the car. In Italy, 65.3% of journeys are made by car (mostly as driver), an increase of nearly eight points over the last 15 years (ISFORT 2018).

The transport sector is the primary cause of the observed deterioration in urban air quality, road transportation having increased significantly. For example, in Italy, the increase in the levels of PM_{10} and $PM_{2.5}$ between 2014 and 2018 has been attributed primarily to an increase in private car ownership and vehicle miles traveled by car (ISFORT, 2018); (PM_{10} and $PM_{2.5}$ in 2018 in Italy were far higher than standards set by both the European Union and the World Health Organization, as per ISFORT, 2018). In the U.S., road transportation is estimated to contribute about 84% of the metric tons of CO_2 equivalent of total transportation-based GHG emissions (US EPA, 2017), and the transportation sector as a whole is estimated to be responsible for about 29% of energy-related CO_2 emissions (the next highest contributing sector, the industrial sector, is estimated to contribute about 22% of energy-related CO_2 emissions; see US EPA, 2017). In addition to the substantial impacts on air quality and greenhouse gas (GHG) emissions, the transportation sector also contributes to several other externalities that impact the environment and urban quality of life, including noise pollution, public health and safety issues.

Therefore, the current transport system, based on the use of vehicles fueled almost exclusively by petroleum products, generates non-negligible social, environmental and economic impacts. To overcome the problems associated with the current system, in recent years, particular attention has been paid to sustainable urban mobility at different levels. Governmental bodies, policy makers, transportation operators and academic researchers have focused on strategies and policies to reduce automobile travel and induce a modal shift in travel behavior. Many recent transportation plans have included Travel Demand Management (TDM) strategies as a means to create a more sustainable transportation system. Traditionally, TDM has been defined as any initiative endeavoring to influence traveler choices about whether to travel, which mode to use, and where to travel. Some examples include incentives to use higher occupancy modes (e.g., carpool, vanpool, transit) or non-motorized modes (cycling, walking), programs to encourage working or shopping from home (addressing whether to travel) or to reduce trip lengths (by combining trips or shopping or working closer to home). In particular, there has been increasing interest in soft measures which use information and communication tools to enhance people's knowledge (for example of transport alternatives), heighten people's awareness and encourage them to change their travel behavior.

However, the success of Travel Demand Management measures requires a deep knowledge of the determinants of mode choice. Understanding and predicting individual behavior is a complex task and a misrepresentation of the factors affecting individuals' decisions can lead to an inaccurate prediction of the transportation demand, which may have an adverse effect on the cost-effective analyses of transportation plans and policies. A quantitative characterization of the impact of different behaviors on demand can be obtained through the construction of mathematical models. Traditional Discrete Choice Models (DCM), also called Random Utility models, are a widely used tool adopted by transportation and urban planners to assess and forecast traffic volumes, transit ridership, walking and cycling market share across transportation networks. Based on the microeconomic theory of utility, the particularity of these models is their ability to predict decisions at the individual level, based on objective personal characteristics (age, gender, income, etc.) and attributes of the choice alternatives (mainly time and cost).

Different works have criticized this kind of approach and hold it responsible for the failure of policies in favor of sustainable mobility (Paulssen *et al.*, 2014), as they simulate in an incomplete manner the cognitive process that leads to the formation of individuals' tastes and preferences (Bamberg et al., 2003; Anable, 2005; Steg, 2005) and treat the unobserved psychological preliminaries of choice as contained in a "black box". Several studies have actually shown that, in addition to the objective characteristics of transport alternatives, cognitive aspects, such as habits, emotions, values, attitudes, perceptions, that are not directly observable, can affect individual behavior. For example, environmental concern has been found to have a positive effect on environmentally-friendly alternatives in several studies. Furthermore, specific attitudes towards the car alternative have been found to have an effect in mode choice studies (Anable, 2005).

Only in the last two decades have models of disaggregate decision-making started to include latent constructs for capturing the impact of subjective factors on choice process. These kinds of models, called hybrid choice models (HCM) or integrated choice and latent variable (ICLV) models, proposed in the 1980's by McFadden (1986) and Train *et al.* (1987), only became popular in 2002 following the work of Ben-Akiva *et al.* (2002). ICLVs appear to be a powerful and useful method for improving existing representations of decision making (Vij and Walker, 2016) and providing recommendations for travel demand management policies.

However, recent work in transportation research (Chorus and Kroesen, 2014) raised the question about the ability of ICLVs to derive policy implications that aim to change travel behavior, pointing out two potential issues. Firstly, latent attitudes and perceptions are partly endogenous with respect to travel behavior, and the analyst cannot be sure about the direction of causality in attitudes, perceptions, and choices. Moreover, latent variables and the observed choice could be influenced by the same underlying, unmeasured factors. Secondly, the data collected are typically cross-sectional, measured at a single moment in time. This means only between-person comparisons based on differences in latent variables can be evaluated, and not within-person comparisons and how latent variables may change over time as a result of variation in socio-economic (SE) characteristics or the implementation of travel demand management policies.

Therefore, the objective of this thesis work is to identify the methodological processes that can guarantee that sustainable mobility policies seeking to reduce car use are accompanied by a quantitative assessment of the effects that this generates in the transport context. In particular, the thesis focuses on the role of psycho-attitudinal variables (attitudes, perceptions, habits, *etc.*) in the process of choosing a mode alternative through the use of cross-sectional and panel data. An in-depth analysis is advantageous from both a policy point of view and for representing the decision-making process. In fact, the implementation of effective strategies for the promotion of sustainable mobility, such as information campaigns that focus on those factors that could persuade people not to use the car, can benefit from an improved understanding of this phenomenon and help to avoid wasting limited resources, as well as failures that would diminish public support.

The first part of the work focused on the analysis of the determinants influencing cycling behavior. Recently, there has been a surge of interest in cycling as a physically active transportation option able to bring immediate and multiple benefits at both the individual as well as the community level. However, the traditional random utility modeling approach failed in some cases to explain the process underlying the decision to cycle, making it necessary to incorporate those psychological factors attached to bicycle use and derived from its special characteristics. In particular, the aim was to accommodate psychological variables within the analytical framework and understand to what extent they affect the choice to use the bike. In doing so, we constructed and estimated different econometric models, each aimed at shedding light on those aspects overlooked in the research. The context of the study were the urban areas of Cagliari and Sassari, main cities in Sardinia (Italy), where, despite the implementation of policies supporting bike use, cycling levels for commuting trips are still low.

In the second part of the work we analyzed the short-term effect on travel mode choice of introducing a new sustainable form of transport into the choice set (hard measure) when implementing an information measure (Personalized Travel Plan). The transport context chosen for this experiment is a corridor linking the city center of Cagliari to a university/hospital complex, where a new light rail route went into service. In particular, an attempt was made to overcome the critical issue concerning evaluation of the measure by creating a control group, so as to disentangle the effect of the structural measure from that of personalized information provision.

Next, we evaluated the long-term effects of these measures, along with an analysis that aimed to detect any changes in the psycho-attitudinal factors and SE characteristics. In particular, through the specification and estimation of ICLV models that used for the same sample the data collected for these two moments in time, we analyzed whether changes in individual characteristics were able to affect mode choice, and so the criticism raised against ICLV models is valid.

The thesis is structured into three parts, each composed of different chapters and paragraphs. Part I reviews the state of the art in the research areas covered in the thesis work. In Part II we explore what determinants, objective and subjective, influence cycling behavior. In Part III we report the results, evaluated in the short- and long-term, of the combined implementation of a structural measure with an information measure.

PART 1 – LITERATURE REVIEW

1.1. INTRODUCTION

This chapter reviews the state of the art in the research areas covered in the present thesis work. The discussion of this chapter is organized as follows. The first section (1.2) treats Travel Demand Management strategies, their strengths and weaknesses and the best way to evaluate their effectiveness.

Then we describe the existing work in the area of discrete choice models (1.3). After a brief review of the determinants influencing travel behavior (1.3.1) and utility-based choice theory (1.3.2), we look at the mathematical formulation of logit and probit models from Section 1.3.3 to Section 1.3.7. We then turn to look at models of ordered discrete data in Section 1.3.8. Next, we present the ICLV model in Section 1.3.9 and the GHDM model in section 1.3.10. Finally, we describe goodness-of-fit measures (1.3.11) and aggregation and forecasting techniques (1.3.12).

1.2. TRAVEL DEMAND MANAGEMENT STRATEGIES

1.2.1. Travel demand management definition

Travel demand management (TDM) is defined as any initiative with the object of helping to alleviate the negative impact of car use (Taylor and Ampt, 2003; Loukopoulos, 2007). Other terms with similar meanings include transport system management (Pendyala *et al.*, 1997), transportation demand management (Litman, 2003) and mobility management (Kristensen and Marshall, 1999, Rye 2002).

These measures may vary in several attributes or dimensions, which are likely to affect the key outcome variables of effectiveness in reducing travel demand, political feasibility, and public attitudes (see Table 1).

A widely accepted distinction is that between structural strategies, aimed at changing the physical and/or legislative context in which choices are made, and informational strategies, aimed at changing prevalent motivations, perceptions, cognitions and norms and motivating voluntary changes in transportation choices (Graham-Rowe *et al.*, 2011).

Travel demand management measures may also be distinguished between push and pull measures (Steg and Vlek, 1997). Push measures attempt to discourage car use by making it less beneficial (*e.g.* implementation of a congestion pricing measure). Pull measures encourage the use of alternative modes to the car by making them more attractive. Cheaper public transport fares or new bike lanes are examples of pull measures.

Table 1. Travel Demand Management Measures (adapted from Steg (2005) and Cairns *et al.* (2008)

TDM Measures		Examples
Hard measures	Physical change measures	Improving public transport Improving infrastructure for walking and cycling Park & ride schemes Land use planning to encourage shorter travel times Technical changes to make cars more energy-efficient
	Legal policies	Prohibiting car traffic in city centers Parking control Decreasing speed limits
	Economic policies	Taxation of cars and fuel Road or congestion pricing Kilometer charging Reducing costs of public transport
Soft measures	Informative measures	Workplace/school travel plans Personalized travel planning Public transport information and marketing Travel awareness campaigns
	Work schedules and telecommuting	Teleworking Compressed workweeks Flexible work schedules

1.2.2. Structural measures

Structural measures, also referred to as hard measures, aim to change the external context in which choices are made (Steg and Vlek, 2009). Depending on how behavioral changes may be elicited, three types of hard measures can be distinguished (Steg and Tertoolen, 1999).

First, levels of car use can be reduced improving the relative attractiveness of alternative travel modes through physical changes such as the provision of new infrastructure for walking and cycling, the introduction of new public transport routes, or removing parking places. Recently, different natural experiment studies have found that interventions in the built environment may encourage a modal shift, supporting the association between the built environment and mode choice (*e.g.* Handy *et al.*, 2005; Saelens and Handy, 2008; Pinjari *et al.*, 2007; Ewing and Cervero, 2010). Examples of structural measures include improved infrastructure for public transport, walking and cycling and reduction of car parking places.

A second type of strategy to enforce car use is legal regulation measures. They can include prohibiting car traffic in city centers, decreasing speed limits, and introducing parking regulations. The assumption is that people will observe these regulations and in the longer term they will lead to changes in social norms (Gärling and Schuitema, 2007). However, such strategies require an adequate organization for supervision, monitoring and enforcement, with their success depending on majority public support, or at least compliance (Steg and Tertoolen, 1999).

Third, financial/economic measures can help to diminish car use making it relatively more expensive. Examples of economic policies include congestion or road pricing, taxation of fuels and cars, and reducing costs of public transport. The main assumption underlying this kind of measures is that people's behavior is rational, and they choose the alternative with the highest utility at the lowest costs.

1.2.3. Informative measures

However, hard measures do not always produce the desired results as they do not succeed in making car use less appealing or triggering behavior change. Traveling by car offers many personal benefits, both instrumental (flexibility, shorter travel times, comfort) and psychological (feeling of freedom, *etc.*) (Steg, 2005), that make it difficult to change a choice that has become habitual, even though other alternatives, objectively more advantageous in time, cost, and so on, do exist (Gärling and Fuji, 2009). Furthermore, people who behave habitually are less likely to be interested in information about available alternatives and tend to make decisions that involve few cognitive resources (Verplanken *et al.*, 1998). For all these reasons a structural measure may not be sufficient to influence car drivers to change their travel behavior: they need to be informed and educated to enhance their knowledge (for example of transport alternatives), heightening their awareness (for example of environmental impacts) and modifying attitudes, thereby strengthening the propensity to adopt non-motorized forms of travel (Steg *et al.*, 1999). On the basis of the studies of these process, soft measures, known as Voluntary Travel Behavior Change (VTBC) programs (Ampt, 2003), were developed. They aim to change individuals' perceptions, motivations, knowledge and norms (Steg and Vlek, 2009), so as to steer them toward a more pro-environmental behavior. Informational strategies aim to enhance car users' knowledge of behavioral alternatives and their pros and cons (Steg and Vlek, 2009). They may include social modeling (*e.g.* preeminent public figures using alternative travel modes) to strengthen social norms as well as individualized marketing (providing people with customized information about their travel options) (Gärling and Schuitema, 2007). Cairns *et al.* (2008) identified and defined 10 different soft measures, including:

- Workplace/school travel plans, in which a package of measures is introduced at an individual school/ workplace to encourage students/employees to travel more sustainably;
- Personalized travel planning (PTP), in which individuals are provided with personalized information, to make them aware of available sustainable means of transport and encourage their use;
- Public transport information and marketing (PTIM), which includes advertising campaigns and the provision of information in more accessible formats;
- Travel awareness campaigns, in which several media are involved and attempt to enhance general public awareness of the negative consequences of car use.

Numerous PTP programs have been implemented worldwide over the last 20 years adopting different methodological approaches. Most of them are inspired by two popular works: IndiMark

(TravelSmart) and Travel Blending. IndiMark (TravelSmart) (Brög *et al.*, 2009) is a social marketing approach that attempts to increase users' knowledge of transport system. The program targets single individuals, but it is not fully customized as participants are only provided with a package of general information. IndiMark has been implemented in Australia, and several European countries. Travel Blending (Brög *et al.*, 2009; Taylor and Ampt, 2003) aims to reduce the number of car trips through a "blend" of 1) different means of transport (car, public transport, bike, walk), and 2) of activities in space and time (Ampt, 1999). Since its approach provides a quantitative feedback tailored to each individual, the measure is implemented at a smaller scale than the TravelSmart approach. The program, consisting in two phases, has been implemented in Australia since 2002.

Other early programs include Travel Feedback Programs (Fuji and Taniguchi, 2006), which are the most common mobility management tool in Japan. Participants in the program receive feedback (*e.g.* CO_2 produced by car use) and useful information such as timetables and frequency of public transport service. TPF are programs conducted on a small scale by university researchers. The first TFP was implemented in 1999, followed by several to date.

Following these experiences, different types of soft measures have been implemented in other countries. Despite being a sparsely populated country that does little promote the use of public transport, in Sweden soft measures have been implemented adopting various techniques (Friman *et al.*, 2013). The programs have been conducted mainly in residential areas, at workplaces, or at schools and included travel plans, incentives, information and feedback provision. Casteddu Mobility Styles (Sanjust di Teulada, *et al.*, 2015) is a VTBC program implemented in Cagliari, Italy, between 2011 and 2012 for promoting a light rail service. The program comprised two phases. In the first phase data on current travel behavior were collected via a smartphone app, which were then used to create a personalized travel plan (PTP). In the second, after PTP delivery, travel and activity patterns were observed to detect any changes in travel behavior. Interestingly, Hsieh *et al.* (2017) conducted a randomized social experiment in Taipei city aimed at examining the effect of implementing two different types of personalized travel plan interventions, action plans and coping plans. The action plan intervention mainly followed the PTPs assisting commuters to create alternatives to car use; in contrast, the action-plus-coping plan intervention combined the action plan intervention with a strategy triggering the participants to reflect on the potential barriers to switching from car use to public transport, so that it was possible to reinforce the formulated action plans or make the plans adjustable. In Valencia, Spain, between 2010 and 2011, Arroyo *et al.* (2018) implemented a travel behavior change program based on three different actions following Cialdini's principles of persuasions of (2009). In the first phase all the participants received a report with the characteristic of their past travel behavior and information concerning the available sustainable travel alternatives. Then they were invited to attend a talk given by a cardiologist and a sports trainer on the relationship between health and physical activity. Finally, respondents were invited to watch a video session where people who had recently reduced car use were interviewed.

1.2.3.1. Demonstrated effects of soft measures

Analysis of all these experiences demonstrated that soft interventions have strengths and weaknesses. In the UK Cairns *et al.* (2008) describe several experiences of soft measure implementation, achieving an average reduction in car use (vehicle kilometres) of 4-5%. Workplace travel plans reduced car use by 10-30%, school travel plans by 8-15%. Personalized travel planning in households achieved a reduction in car driver trips of 7-15% in urban areas and 2-6% in rural and smaller urban areas. The implementation of the program Casteddu Mobility Styles (Sanjust di Teulada, *et al.*, 2015) showed that 30% of participants were observed to have changed their travel behavior and a general average decrease of 8% in the distance traveled and of 11% in the number of trips made by car was observed.

Two meta-analyses (a statistical technique that combines the results of multiple scientific studies in an effort to provide quantitative estimates of treatment effects) of previous research have also been conducted. Möser and Bamberg (2008) examined the data set of 141 studies evaluating three types of soft transport policy measures implemented in various countries. According to the authors, all the studies analyzed used a weak quasi-experimental evaluation design.[1] The soft measures evaluated were divided into three categories: work travel plans, travel planning/awareness campaign/PT marketing and school travel plans. Across all three soft policy measures, the results showed percent point increase of the no-car use proportion from 39% to 46%. In particular, for work travel plans the analysis indicated a mean increase of 12% in the proportion of employees not commuting by car and a 5% increase of the trip proportion not conducted by car for travel planning/awareness campaign/PT marketing. In the second meta-analysis, Tanigushi *et al.* (2007) investigated the results of Travel Feedback Programs (TFP) conducted in Japan. The programs were classified into three categories depending on the location where they were conducted: residential areas, workplaces or schools. The authors limited their analysis to those programs implemented in residential areas, because of the higher number of reports and larger sample sizes. The analysis results indicated a 7.3% reduction in car use, a 68.8% increase in public transport, and a 10.4% and 7.5% increase in intentions to limit car use and increase public transport use, respectively. Closer analysis of just the TFP interventions that had control groups revealed a

[1] Wall *et al.* (2011) in their article in response to the one by Möser and Bamberg (2008) criticized this assumption, stating that at least fifteen studies used control groups to evaluate interventions.

12.1% reduction in car use and a 38.6% increase in public transport as well. However, the total number of studies was small and most of them were based on small non-representative sample.

Hence, PTP can reduce car usage by around 10% and lead to significant increases in walking, cycling, and public transport use. Other experiences have shown that ‘‘soft’’ measures, such as PTP, generally work better when used in conjunction with relevant infrastructure improvements (new cycle paths, new bus routes, new pedestrian crossings, *etc.*) or with the introduction of ‘‘sticks’’ such as parking charges or reduced capacity for road traffic or parked vehicles (Bonsall, 2009). Cairns *et al.* (2008) report that plans which included parking management measures achieved an average reduction in car driver trips of 24%, compared with 10% for those that did not. A study by Transport for London shows that the implementation of smarter choice measures combined with a road pricing measure achieved a reduction in total traffic levels of 8–17%, compared with 2–4% had they been applied alone (Cairns *et al.*, 2008).

1.2.3.2. Procedures for the success of soft measures

Analysis of the numerous projects conducted over the past 20 years identified the key factors distinguishing a successful VTBC program (Meloni and Sanjust di Teulada, 2015):

- Target mobility context: if the context in which the measure is to be implemented is inappropriate, it may be ineffective.
- Target behavior and population: selecting clear and measurable behavioral goals facilitates behavior change (Davies, 2012). It is also essential to target a suitable segment of the population for the alternative to be promoted. This implies an in-depth knowledge of participants’ socio-economic and psychological characteristics.
- Removing barriers to behavior change. Barriers can be identified as:
    - External barriers, associated with the transport and social-economic context;
    - Internal barriers, associated with the psycho-motivational nature of the travel choice and reflecting the individuals’ point of view on the travel choice made (negative attitude toward a certain means of transport, poor awareness of automobile externalities, *etc.*);
    - Barriers caused by habitual behavior.
- Personalization: this is the most effective tool for breaking down internal barriers to a sustainable means of transport (Fujii and Taniguchi, 2006). The greater the level of information customization of a VTBC program, the greater its effectiveness will be (Meloni and Sanjust di Teulada, 2015). The proposed alternative must take into account individuals’ limits and needs (Stopher, 2005). Another important element is the contact between the VTBC program team and participants, to exchange opinions and offer advice (Friman *et al.*, 2019).

- Information: this is the core of a VTBC program. The information provided must be able to influence those psychological aspects crucial to mode choice. It should be useful, usable, used and readily acquired (Abrahamse *et al.*, 2005). Feedback serves to compare the disadvantages of actual behavior with the advantages of sustainable behavior, as individuals are unable to quantify them.
- Persuasion: information, by itself may be ineffective in evoking behavior change (Abrahamse *et al.*, 2005). Information needs to be integrated with communication and persuasion techniques, based on the knowledge of individuals' psychology (Meloni and Sanjust di Teulada, 2015).
- Use of one of three behavioral approaches or a combination thereof: 1) behavioral theories and models such as the Theory of Planned Behavior (Ajzen, 1991), the Norm Activation Model (Schwartz, 1977), the Value-Belief-Norm theory (Stern *et al.*, 1999), or the Theory of Interpersonal Behavior (Triandis, 1979); 2) behavior change theories such as Lewin's Theory of Change, Transtheoretical Model (Prochaska *et al.*, 1982), Self-Regulation Theories (Carver *et al.*, 1989), or the Model of Implementation Intention (Gollwitzer, 1993); 3) DEFRA social marketing (Defra, 2008) and persuasion techniques (Cialdini, 2001).
- Evaluating the effectiveness of interventions and monitoring: evaluation of a soft measure is important for researchers and policy-makers (Steg and Vlek, 2009). In general, these are assessed in terms of car use reduction (time traveled, distance, number of trips) and benefits for the single individual (kilocalories [kcal] burned]) or community (*e.g.*, carbon dioxide [CO_2] emissions reduction). Monitoring is another key factor for observing any changes in travel behavior and whether changes are long lasting.

1.2.3.3. Criticism of soft measures

Notwithstanding the above results and the well-known procedures for conducting a successful VTBC, some contrary opinions can also be found in the literature. Tertoolen *et al.* (1998) found that although information can affect attitudes, it does not affect travel behavior. Eriksson *et al.* (2006) found that information campaigns are considered by car drivers to be a relatively ineffective measure compared with improving public transport and increasing fuel taxes. One problem associated with information is the level of personalization, as it affects the effectiveness of the measure. Eriksson *et al.* (2008) found that less than 2% of respondents believed that improving timetable information would persuade them to use public transport. However, a high degree of personalization can be costly and difficult to achieve in large scale programs (Meloni and Sanjust di Teulada, 2015). Another problem is how to evaluate the effectiveness of a measure, particularly in the absence of a control group. VTBC programs are often conducted within broader campaigns launched to reduce car travel, with a focus on the overall effect (Tørnblad *et al.*, 2014). So, without a control group, it is difficult to assess the effectiveness of the implementation of a single measure and disentangle it from a combination of effects. A study by Fuji *et al.* (2009) shows that outcomes of 15 Travel Feedback Programs conducted in Japan differ depending on whether suitable control groups have been included in the study or not. Another problem that affects evaluation of VTBC programs is the variability in travel from day to day, from season to season, and in response to external stimuli (Stopher *et al.*, 2009). Multi-day self-report surveys, with the compilation of a 2-day activity diary among those who agreed to participate, may be the best option to overcome this problem (Stopher *et al.*, 2009). Another criticism is that the evaluation is highly dependent on self-reporting (Bonsall, 2009). A solution could be the use of passive GPS devices, which require almost no action on the part of respondents (Stopher *et al.*, 2009).

1.3. EXISTING DISCRETE CHOICE MODELS

1.3.1. Determinants of travel choice behavior

As mentioned in 1.2.3.2, a deep understanding of the determinants of mode choice is essential to design environmentally sustainable transport system in line with people's preferences (Bhat, 1998; Vredin Johansson *et al.*, 2005). Modal choice can be defined as the decision process that leads an individual to make a travel choice among different alternatives. This process can take place consciously or unconsciously and its analysis includes a wide range of factors from different disciplines (economy, sociology, geography and psychology) (De Witte *et al.*, 2013). In their literature review De Witte *et al.* (2013) distinguish four major determinants of modal choices:

- Socio-demographic indicators, which include age, gender, education, occupation, income, household composition and car availability.
- Spatial indicators, that characterize the spatial environment in which the trip takes place. Some examples are density, diversity, proximity to infrastructure and services, parking availability.
- Journey characteristics indicators such as travel motive, distance, travel time, travel cost, departure time, interchange.
- Socio-psychological indicators, concerning the personal such as experiences, lifestyle, habits and perceptions.

Table 2. Classification of modal choice determinants (adapted from De Witte *et al.*, 2013)

Socio-demographic indicators	Spatial indicators	Journey characteristics indicators	Socio psychological indicators
Age	Density	Travel motive	Experiences
Gender	Diversity	Distance	Lifestyle
Education	Proximity to infrastructure and services	Travel time	Habits
Occupation	Parking availability	Travel cost	Perceptions
Income		Departure time	
Household composition		Interchange	
Car availability			

Traditionally, travel choices have been studied and analyzed using the rationalist approach that assumes that individuals take decisions based on utility maximization (see 1.3.2.) This microeconomic approach treats individuals as an "optimizing black box" who select the alternative that minimizes their travel times and costs. The factors included in these works are mostly travel time and costs, individual and household characteristics. More recently, different studies have started to include in their analysis spatial indicators (*e.g.* travel tour motivation and complexity, residential location and neighborhood type), treating demand for travel as a derived demand, where people travel to pursue activities distributed in space and time (De Witte *et al.*, 2013).

However, this representation of consumers as individuals with predetermined wants and needs is inconsistent with the findings from studies of social sciences (Vij and Walker, 2016) that suggest that socio-psychological indicators play a key role in influencing people's mode choice. These studies have frequently shown that subjective constructs such as perceptions, lifestyle, habits can often have a greater influence on behavior than that exerted by objective observable variables (see for example Bamberg *et al.*, 2003; Anable, 2005; Beirão and Cabral, 2007)

This is especially the case when exploring the choice process that leads to drive a car. Steg (2005) distinguishes three different factors, of a psychological nature, influencing the use of the car: instrumental motives, symbolic motives and affective motives. *Instrumental motives* are related to personal benefits deriving from the use of the car (such as its speed, flexibility, and convenience). *Symbolic motives* refer to the fact that people can, through the possession and use of the car, express their preferences and show their social status: in fact, they can compare the value of their cars - and therefore their social status - with that of others (Steg, 2005). *Affective motives* refer to emotions (freedom, light-heartedness, vitality) that are transmitted by driving the car.

1.3.2. Utility-based choice theory

The most common theoretical construct for deriving discrete outcome models is the random utility theory (RUM) (McFadden, 1974), which postulates the utility-maximizing behavior by the decision maker.

Let $C = \{c_1, ..., c_j, ..., c_n\}$ be a finite set of mutually exclusive alternatives and Q a given homogenous population of decision makers. Each alternative $c_j \in C$ has associated a certain level of utility U_{qj} for each individual q. This utility is known to the decision maker but not by the modeler. The decision maker will choose alternative j if and only if $U_{qj} > U_{qi}\ \forall i \neq j$, with $i, j \in C$.

The modeler cannot observe the decision maker's utility, but only some attributes of the alternative and the tastes and sociodemographic attributes of the decision-maker. Therefore, the modeler assumes that:

$$U_{qj} = V_{qj} + \varepsilon_{qj} \tag{1.1}$$

Where V_{qj} is the observed part of the utility and ε_{qj} captures the factors that affect utility but are not included in V_{qj}. V_{qj} can be expressed as $f\ (\beta_q,\ x_{jq})$, where x_{jq} represents a vector of measured attributes of alternative j as faced by decision-maker q and β_n is a vector of parameters representing the tastes of decision-maker q, which is to be estimated from the data. Usually all parameters enter the utility function linearly, although, in many cases, the use of a non-linear formulation has clear advantages (see for example Mandel *et al.*, 1997).

Since the modeler does not know ε_{qj}, he treats these terms as random, meaning that the deterministic choice process is represented as a probabilistic phenomenon. The probability that decision maker q chooses alternative j is:

$$\begin{aligned} P_{qj} &= Prob\left(U_{qj} > U_{qi}\ \forall i \neq j\right) \\ &= Prob\left(V_{qj} + \varepsilon_{qj} > V_{qi} + \varepsilon_{qi}\ \forall i \neq j\right) \\ &= Prob\left(\varepsilon_{qi} - \varepsilon_{qj} < V_{qj} - V_{qi}\ \forall i \neq j\right) \end{aligned} \tag{1.2}$$

with the unobserved part of utility varying randomly across respondents. The random vector ε_q possesses a joint density distribution $f(\varepsilon_q)$, with zero mean and covariance matrix Ω. The cumulative probability can be written now as:

$$\begin{aligned} P_{qj} &= Prob\left(\varepsilon_{qi} - \varepsilon_{qj} < V_{qj} - V_{qi}\ \forall i \neq j\right) \\ &= \int_{\varepsilon} I\left(\varepsilon_{qi} - \varepsilon_{qj} < V_{qj} - V_{qi}\ \forall j \neq i\right) f\left(\varepsilon_q\right) d\varepsilon_q \end{aligned} \tag{1.3}$$

where $I(\cdot)$ is the indicator function, equaling 1 if the term in parentheses is true and 0 otherwise. Different discrete choice models may be obtained from different specifications of the density, though the integral takes a closed form only for certain distributions of ε_q.

Equation (1.2) shows that if a constant is added to the utility of all alternatives, or all utilities are multiplied by the same constant, the alternative with the highest utility does not vary, leading to the conclusion that the absolute level of utility is irrelevant and only differences in utility matter.

The use of the RUM paradigm has many advantages, so that forcing choice models to that paradigm, where possible, is often found to be beneficial, even though there it could entail a loss of explanatory power or clarity of theoretical exposition of the model (Hess *et al.*, 2018). The main benefit of the RUM approach is its link with microeconomic theory and its large apparatus of methodologies and tests of behavior. Positioning the modeling approach within such an accepted behavioral framework helps in achieving acceptance for the approach, as it is always possible to provide evidence of its strength and weaknesses (Hess *et al.*, 2018).

1.3.3. The multinomial logit model

The multinomial logit model is the simplest and one of the most used choice model forms (Domencich and McFadden, 1975). The logit model is obtained by assuming that each ε_{qi} is an independently, identically distributed (iid) extreme value (also called Gumbel) with variance $\pi^2/6$. With this assumption, the choice probability for alternative i and individual q is given by (Train, 2009):

$$P_{qi} = \frac{e^{V_{qi}}}{\sum_{c_j \in C} e^{V_{qj}}} \tag{1.4}$$

If utility is specified to be linear in parameters $V_{iq} = \boldsymbol{\beta}' \boldsymbol{x}_{qi}$, the logit probability becomes:

$$P_{qi} = \frac{e^{\beta' x_{qi}}}{\sum_{c_j \in C} e^{\beta' x_{qj}}} \tag{1.5}$$

As demonstrated by McFadden (1974), the log-likelihood with these choice probabilities is globally concave in parameters β.

The multinomial logit model satisfies the axiom of independence of irrelevant alternatives (IIA), meaning that the relative probability of choosing one alternative over another is independent of the attributes or even existence of other alternatives. Another effect of the IIA property is that the cross elasticities of logit probabilities are uniform, thus the percentage change in the probability of choosing alternative i, given a percentage change in attribute m of alternative j, is constant for all $i \neq j$ (Ben-Akiva and Lerman, 1985; Mokhtarian, 2016). Even though in different travel behavior contexts the IIA property is not a realistic assumption (see Ben-Akiva and Lerman, 1985 and Train, 2009), there are however also cases where the IIA is valid and its employment is a considerable advantage, such as when we want estimated model parameters on a subset of alternatives for each sampled decision maker or we are only interested in analyzing choices among a subset of alternatives (Train, 2009).

1.3.3.1. The role of the scale parameter

As stated in utility can be expressed as $U_{qj}^* = V_{qj} + \varepsilon_{qj}^*$, where the unobserved part has variance $\sigma^2 \times (\pi^2/6)$. Since the scale of utility is irrelevant and only differences in utilities matter, U_{qj} can be divided by σ without changing behavior, so that:

$$U_{qj} = \frac{V_{qj}}{\sigma} + \varepsilon_{qj} \tag{1.6}$$

Where $\varepsilon_{qj} = \varepsilon_{qj}^*/\sigma$. The choice probability can be expressed as

$$P_{qi} = \frac{e^{\frac{V_{qi}}{\sigma}}}{\sum_{c_j \in C} e^{\frac{V_{qj}}{\sigma}}} \tag{1.7}$$

If V_{qj} is linear in parameters with coefficients β^*, the choice probability is:

$$P_{qi} = \frac{e^{(\boldsymbol{\beta}^*/\sigma)' x_{qi}}}{\sum_{c_j \in C} e^{(\boldsymbol{\beta}^*/\sigma)' x_{qj}}} \tag{1.8}$$

with each coefficient scaled by $1/\sigma$. The parameter σ is called scale parameter, because it scales the utility to reflect the variance of the unobserved portion of utility (Train, 2009). Because σ cannot be identified, only the ratio β^*/σ can be estimated.

We assumed that the variance of unobserved factors is the same for all individuals in the sample. However, in some situations, the variance of the error terms can vary among different segments of the population. Swait and Louviere (1993) examine the role of the scale parameter in discrete choice models, investigating the several reasons that lead the variances to differ over observations.

1.3.4. The nested logit model

To overcome the IIA limitation in simple MNL models a class of models known as generalized extreme value models was developed by McFadden (1978). All the models are based on the use of the extreme-value distribution, which allows for various levels of correlation among the unobserved part of utility across alternatives.

The nested logit (NL) model is one of the more commonly used models in this class. The idea behind a nested logit is to divide the choice set into nests of alternatives, in such a way that for any two alternatives that are in the same nest, the ratio of probabilities is independent of the attributes or existence of all other alternatives and for any two alternatives in different nests, the ratio of probabilities can depend on the attributes of other alternatives in the two nests (this sets up an IIA violation in the nest).

Mathematically, McFadden (1981) has shown the NL disturbance assumption leads to the following model structure for observation *q* choosing outcome *i* in nest *j* is

$$P_q(i) = \frac{e^{\boldsymbol{\beta}' \boldsymbol{x}_{qi} + \phi_i L_{qi}}}{\sum_I e^{\boldsymbol{\beta}' \boldsymbol{x}_{qI} + \phi_I LS_{qI}}} \tag{1.9}$$

$$P_q(j|i) = \frac{e^{\beta'_{j|i} x_q}}{\sum_J e^{\beta'_{J|i} x_q}} \tag{1.10}$$

$$LS_{qi} = ln\left(\sum_{j} e^{\boldsymbol{\beta}'_{J|i}\boldsymbol{x}_q}\right) \quad (1.11)$$

where $P_q(i)$ is the unconditional probability of observation q having discrete outcome i, the x are vectors of observed characteristics, the β are vectors of estimable parameters, $P_q(j|i)$ is the probability of observation q having discrete outcome j conditioned on the outcome being in outcome category i, J is the conditional set of outcomes (conditioned on i), I is the unconditional set of outcome categories, LS_{iq} is the inclusive value (logsum), and ϕ_i is an estimable parameter. Note that this equation system implies that the unconditional probability of having outcome j is:

$$P_q(j) = P_q(i) \times P_q(j|i) \quad (1.12)$$

1.3.5. The heteroskedastic logit

Instead of capturing correlations among alternatives, the analyst may only want to allow different variances on the random components across alternatives. Bhat (1995) developed a random utility model with independent, but non-identical error terms distributed with a type I extreme value distribution, called *heteroskedastic logit*.

The random utility of alternative i, U_{qi}, is specified as $U_{qi} = V_{qi} + \varepsilon_i$, where V_{qi} is the systematic component of the utility of alternative i which is a function of observed attributes of alternative i and observed characteristics of the individual q, and ε_i is the random component of the utility function. It is assumed that the random components in the utilities of the different alternatives have a type I extreme value distribution and are independent, but non-identically distributed. It is also assumed that the random components have a location parameter equal to zero and a scale parameter equal to θ_i for the i_{th} alternative. Thus, the probability density function and the cumulative distribution function of the random error term for the i_{th} alternative are:

$$f(\varepsilon_i) = \frac{1}{\theta_i} e^{-\frac{\varepsilon_i}{\theta_i}} e^{-e^{-\frac{\varepsilon_i}{\theta_i}}} \ and\ F_i(z) = \int_{\varepsilon_i=-\infty}^{\varepsilon_i=z} f(\varepsilon_i) d\varepsilon_i = e^{-e^{-\frac{z}{\theta_i}}} \quad (1.13)$$

Thus, the choice probabilities for this heteroskedastic logit are:

$$P_{qi} = \int_{\varepsilon_i=-\infty}^{\varepsilon_i=+\infty} \prod_{j\in C, j\neq i} \Lambda\left[\frac{V_i - V_j + \varepsilon_i}{\theta_j}\right] \frac{1}{\theta_i} \lambda\left(\frac{\varepsilon_i}{\theta_i}\right) d\varepsilon_i \quad (1.14)$$

where $\lambda\,(\cdot)$ and $\Lambda(\cdot)$ are the probability density function and cumulative distribution function of the standard type I extreme value distribution, respectively,

$\lambda\,(t) = \mathrm{e}^{-t} e^{-\mathrm{e}^{-t}}\ and\ \Lambda(t) = e^{-e^{-t}}$.

The integral does not take a closed a form, but it can be approximated by simulation.

1.3.6. The multinomial probit model

An attractive solution to overcome the limits of multinomial logit model is to use the multinomial probit (MNP) framework (Hausman & Wise, 1976; Daganzo, 2014). The multinomial probit model assumes that the underlying utility functions follow a joint multivariate normal (MVN) distribution with zero mean and arbitrary covariance matrix. This means that the variances may be different, and the error terms may be correlated in any fashion, so that probit models can accommodate random taste variation, allow any pattern of substitution and handle panel data.

The only limitation of probit models is that the normal distribution assumption for model parameters might be inappropriate in some situations and can lead to issues in results interpretation. A prominent example concerns cost coefficient values. In a probit model with random taste variations, the normal distribution assumption for price coefficient would imply the presence of people with a wrongly positive price coefficient. The same applies to travel time coefficients, which, in a transportation context, are expected to be negative. Other than this limitation, the probit model is quite general.

The MNP choice probabilities are defined as:

$$P_{qi} = Prob\left(V_{qi} + \varepsilon_{qi} > V_{qj} + \varepsilon_{qj}\ \forall\, j \neq i\right)$$

$$= \int_{\varepsilon_q} I\left(V_{qi} + \varepsilon_{qi} > V_{qj} + \varepsilon_{qj}\ \forall\, j \neq i\right)\phi\left(\varepsilon_q\right)d\varepsilon_q \qquad (1.15)$$

Where *I(·)* is the indicator function and $\phi\left(\varepsilon_q\right)$ is the density of ε_q given by:

$$\phi\left(\varepsilon_q\right) = \frac{1}{(2\pi)^{\frac{J}{2}}|\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}\varepsilon_q'\Sigma^{-1}\varepsilon_q} \qquad (1.16)$$

The integral is over all values of ε_q and does not have a closed form, hence it must be evaluated numerically through simulation.

In the case of a binary probit model, it can be shown (Ortúzar and Willumsen, 2010) that the choice probability can be expressed as:

$$P_1 = \phi[(V_1 - V_2)/\sigma_\varepsilon] \qquad (1.17)$$

with σ_ε defined as

$$\sigma_\varepsilon^2 = \sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2$$

where ρ is the correlation coefficient between the utilities U_1 and U_2.

1.3.7. The mixed multinomial logit model

The mixed logit (ML) model, also called random parameter or error-component logit, is a generalization of the standard logit model and can approximate any random utility model to any degree of accuracy (McFadden and Train, 2000). The first applications of mixed logit date to the early 1980s for evaluating market demand for car attributes (Boyd and Mellman, 1980; Cardell and

Dunbar, 1980). The current form originated from the work of Train *et al*. (1987), Ben-Akiva *et al.* (1993), Bhat (1998), Brownstone and Train (1998) for cross-sectional data, Revelt and Train (1998), Bhat (2000) for panel data.

Mixed logit choice probabilities are calculated as the integrals of multinomial logit probabilities over a density distribution of parameters that can vary randomly across customers, such as

$$P_{qi} = \int L_{qi}(\beta) f(\beta)\, d\beta \tag{1.18}$$

Where $L_{qi}(\beta)$ is the logit probability evaluated at parameters β for individual q and alternative i

$$L_{qi} = \frac{e^{V_{qi}(\boldsymbol{\beta})}}{\sum_{c_j \in C} e^{V_{qj}(\boldsymbol{\beta})}} \tag{1.19}$$

and $f(\beta)$ is a density function. If $f(\beta)$ is degenerate at fixed parameters b ($(\beta) = 1$ for $\beta = b$ and $f(\beta) = 0$ otherwise), the choice probability becomes the simple multinomial logit formula.

The mixing distribution $f(\beta)$ can be discrete or continuous. In the first case the mixed logit becomes the *latent class model*, used in many fields such as psychology and marketing (see Greene and Hensher, 2003). In the second case, given a choice of distribution $f(\beta)$, described by a set of parameters θ, the mixed logit choice probability is given by

$$P_{qj} = \int L_{q_i}(\beta) f(\beta \mid \theta)\, d\beta \tag{1.20}$$

1.3.7.1. Model specification

The mixed logit model can be derived under different behavioral specifications, each approach being conceptually different, though mathematically equivalent.

The more classical derivation is based on random coefficients structure, in which the coefficients vary over decision makers in the population with density $f(\beta)$. The utility of person q from alternative j is specified as

$$U_{qj} = \boldsymbol{\beta}'_q \boldsymbol{x}_{qj} + \varepsilon_{qj} \tag{1.21}$$

where x_{qj} are observed variables, β_q is a vector of coefficient of these variables representing a person's tastes and ε_{qj} is a random term iid extreme value. The density function $f(\beta)$ is a function of parameters θ and can take different forms such as normal, lognormal, uniform or triangular.

For a given value of β_q, the conditional probability for choice i is standard logit, since the ε_{qj} are iid extreme value:

$$L_{qi}(\beta_q) = \frac{e^{\boldsymbol{\beta}'_q x_{qi}}}{\sum_j e^{\boldsymbol{\beta}'_q x_{qj}}} \tag{1.22}$$

However, it is not possible to know the true value of β_q and therefore to condition on β. So, the unconditional choice probability is given by the integral of the expected value of L_{qi} weighted by the density of β_q. The unconditional probability is

$$P_{qi} = \int \frac{e^{\boldsymbol{\beta}'_q x_{qi}}}{\sum_j e^{\boldsymbol{\beta}'_q x_{qj}}} f(\beta)\, d\beta \tag{1.23}$$

The second interpretation of the mixed multinomial logit model is that of the error-components specification. In this case the utility of an alternative *j* for a decision-maker *q* can be specified as:

$$U_{qj} = \alpha' x_{qj} + \mu'_q z_{qj} + \varepsilon_{qj} \tag{1.24}$$

Where x_{qj} is a vector of observed variables relating to the alternative j, z_{qj} is a vector of 0 and 1 terms which determine what error-components enter the utility of alternative j, α is a vector of fixed coefficients, μ is a vector of random terms with zero mean and a covariance matrix Σ, and ε_{qj} is iid extreme value. The unobserved portion of utility is $\eta_{qj} = \mu'_q z_{qj} + \varepsilon_{qj}$. With z_{qj} containing only zero entries $\forall j$, the model reduces to a standard multinomial logit. If z_{qj} is different from zero, correlation among alternatives is introduced: $Cov(\eta_{qi}, \eta_{qj}) = E(\mu'_q z_{qi} + \varepsilon_{qi})(\mu'_q z_{qj} + \varepsilon_{qj}) = z'_{qi}\, W z_{qj}$, where W is the covariance of μ_q.

The choice probability for alternative j is then obtained by integration over the distribution of μ_q, with

$$P_{qi} = \int \frac{e^{\alpha' i + \mu'_q z_{qi}}}{\sum_j e^{\alpha' x_{qj} + \mu'_q z_{qj}}} f(\gamma_q \mid 0, \Sigma)\, d\gamma_q \tag{1.25}$$

1.3.8. Models of ordered discrete data

Many transportation applications involve the assessment of influences on a choice amongst ordered discrete alternatives (Greene and Hensher, 2010). Examples include situations where respondents are asked to provide ratings, ordered opinions or categorical frequency. Although these response data are discrete, use of standard or nested multinomial discrete models is not a proper way to model data of an ordered nature. To address the problem of ordered discrete data, ordered probability models have been developed (McKelvey and Zavoina, 1975).

An ordered-response model postulates the presence of a latent continuous variable for each individual *q*, such that

$$y_q^* = \boldsymbol{\beta}' \boldsymbol{x}_q + \varepsilon_q \tag{1.26}$$

where $\boldsymbol{\beta}'$ is the vector of parameters to be estimated and $\boldsymbol{x}_q$ and ε_q are vectors of explanatory variables and random error term, respectively. Using this equation, observed ordinal data, *y*, are defined as

$$
\begin{aligned}
&y_q = 1 \ if \ y_q^* \leq \mu_0 \\
&y_q = 2 \ if \ \mu_0 < y_q^* \leq \mu_1 \\
&y_q = 3 \ if \ \mu_1 < y_q^* \leq \mu_2 \\
&y_q = \cdots \\
&y_q = I \ if \ y_q^* \geq \mu_{I-2}
\end{aligned}
\tag{1.27}
$$

where the μ are unknown parameters, known as "cut-points" or “threshold parameters”, to be estimated and ordered from the lowest to the highest. I is the highest integer ordered response. Once a distribution for ε is specified, the probability of I specific ordered responses for each observation q can be calculated exactly. If ε_q is assumed to be normally distributed across observations with mean = 0 and variance =1 an ordered probit is obtained. The probabilities are as follows:

$$
\begin{aligned}
&P(y_q = 1) = \Phi(-\boldsymbol{\beta}'\boldsymbol{x}_q) \\
&P(y_q = 2) = \Phi(\mu_1 - \boldsymbol{\beta}'\boldsymbol{x}_q) - \Phi(\boldsymbol{\beta}'\boldsymbol{x}_q) \\
&P(y_q = 3) = \Phi(\mu_2 - \boldsymbol{\beta}'\boldsymbol{x}_q) - \Phi(\mu_1 - \boldsymbol{\beta}'\boldsymbol{x}_q) \\
&\ldots \\
&P(y_q = I) = 1 - \Phi(\mu_{I-2} - \boldsymbol{\beta}'\boldsymbol{x}_q)
\end{aligned}
\tag{1.28}
$$

where $\Phi(\cdot)$ is the cumulative normal distribution and the threshold μ_0 is set equal to 0 (this implies that one need only estimate *I-2* thresholds).

The likelihood function over the population of Q observations is

$$
L(y \mid \beta, \mu) = \prod_{q=1}^{Q} \prod_{i=1}^{I} \left[\Phi(\mu_i - \boldsymbol{\beta}'\boldsymbol{x}_q) - \Phi(\mu_{i+1} - \boldsymbol{\beta}'\boldsymbol{x}_q)\right]^{\delta_{iq}} \tag{1.29}
$$

where δ_{iq} is a dummy variable taking the value of 1 if the observed discrete outcome for the observation q is i, zero otherwise. The corresponding log-likelihood function is:

$$
LL = \sum_{q=1}^{Q} \sum_{i=1}^{I} \delta_{iq} ln\left[\Phi(\mu_i - \boldsymbol{\beta}'\boldsymbol{x}_q) - \Phi(\mu_{i+1} - \boldsymbol{\beta}'\boldsymbol{x}_q)\right] \tag{1.30}
$$

If the assumption is made that ε_q is logistically distributed across observations with mean=0 and variance=1, an ordered logit model results. In such case, probabilities are expressed as:

$$
\begin{aligned}
&P(y = 1) = \frac{e^{[-\boldsymbol{\beta}'\boldsymbol{x}_q]}}{1 + e^{[-\boldsymbol{\beta}'\boldsymbol{x}_q]}} \\
&P(1 < y < I) = \frac{e^{[\boldsymbol{\beta}'\boldsymbol{x}_q - \mu_{y-1}]}}{1 + e^{[\boldsymbol{\beta}'\boldsymbol{x}_q - \mu_{y-1}]}} - \frac{e^{[\boldsymbol{\beta}'\boldsymbol{x}_q - \mu_y]}}{1 + e^{[\boldsymbol{\beta}'\boldsymbol{x}_q - \mu_y]}} \\
&P(y = I) = \frac{e^{[\boldsymbol{\beta}'\boldsymbol{x}_q - \mu_{I-1}]}}{1 + e^{[\boldsymbol{\beta}'\boldsymbol{x}_q - \mu_{I-1}]}}
\end{aligned}
\tag{1.31}
$$

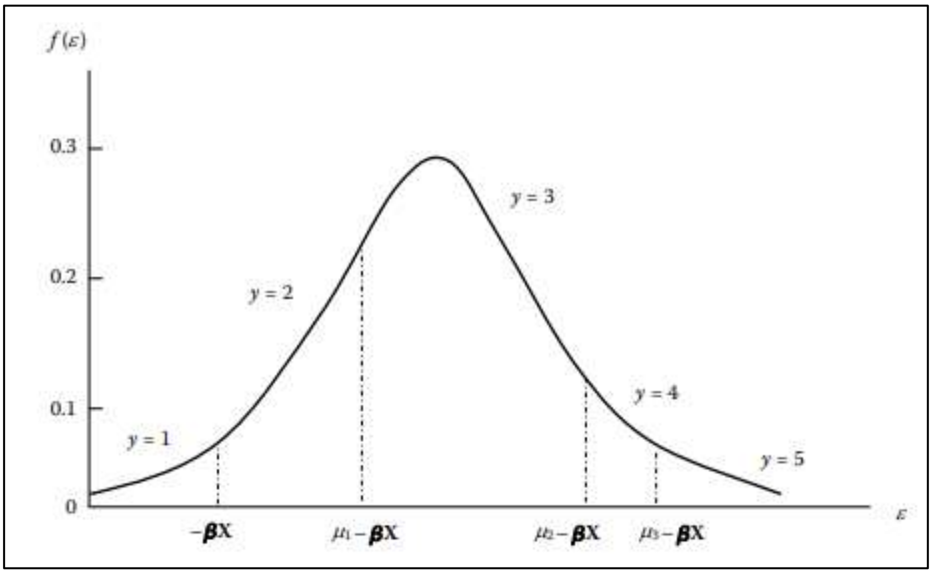


Figure 1. Illustration of an ordered probability model with $\mu_0 = 0$ (from Washington *et al.*, 2010)

The most natural way to interpret model parameter β is to determine how a marginal change in one of the explanatory variables x_k changes the distribution of the outcome variable, *i.e.* all the outcome probabilities. These marginal probability effects can be calculated as

$$\frac{\partial P(y_q = k)}{\partial x} = [\phi(\mu_{i-1} - \beta x_q) - \phi(\mu_i - \beta x_q)]\beta \tag{1.32}$$

where $\phi(\cdot)$ is the standard normal density. In case of discrete variables, it is more appropriate to consider the changes in probability before and after the change in the variable instead of the partial effects using:

$$\Delta P(y = k|x_q, \tilde{x}_q) = P(y_q = k|\tilde{x}_q) - P(y_q = k|x_q) \tag{1.33}$$

where all elements of $\tilde{x}_q$are equal to x_q except for the v-th element, which is equal to $\tilde{x}_{qv} = x_{qv} + \Delta x_{qv}$ for the discrete change Δx_{qv} in the variable x_v .

1.3.8.1. The generalized ordered probit/logit

A limitation of the ordered probit/logit is that it assumes that the thresholds μ are the same for every individual in the sample. This can lead to biased and inconsistent estimates of the effect of variables (Srinivasan, 2002; Eluru *et al.*, 2008). Following the formulation proposed by Terza (1985), the latent propensity can be expressed as:

$$y_q^* = \boldsymbol{\beta'}_q \boldsymbol{x}_q + \varepsilon_q, \quad y_q = k \; if \; \mu_{q,k-1} < y_q^* < \mu_{q,k} \tag{1.34}$$

The β vector and the μ thresholds are now subscripted by the index q to indicate that these parameters can vary across ordered alternatives of different individuals due to observed and unobserved factors.

The thresholds $\mu_{q,k}$, now a function of exogenous attributes of the decision maker with corresponding parameter vectors, are expressed as

$$\mu_{q,k} = \mu_{q,k-1} + \alpha_{q,k} + \gamma'_{q,k} z_{q,k} \quad (1.35)$$

To immediately guarantee the ordering of the thresholds ($-\infty < \mu_{q,1} < \mu_{q,2} < \cdots < \mu_{q,I} < \infty$) for each and every individual q, Eluru *et al.* (2008) proposed the following specific parametric form of the thresholds:

$$\mu_{q,k} = \mu_{q,k-1} + \exp(\alpha_{q,k} + \gamma'_{q,k} z_{q,k}) \quad (1.36)$$

where z_{qk} is a set of exogenous variables associated with the *k-th* threshold (excluding a constant), $\gamma'_{q,k}$ is a corresponding vector of coefficients, and $\alpha_{q,k}$ is a parameter associated with the outcome level $k = 1, 2, \ldots, I-2$.

1.3.8.2. The multivariate ordered probit

While applications of the ordered response models are quite widespread, many of them are confined to the analysis of a single outcome. However, respondents' answers to different questions could be related and the analyst might want to include this phenomenon in the analysis. To account for possible within-outcome correlation, a multivariate ordered-response system framework can be used (for a complete discussion of the problem the reader is referred to Srinivasan and Bhat 2005; Ferdous *et al.*, 2010; Greene and Hensher, 2010).

A multivariate ordered response model structure assumes an underlying set of multivariate continuous latent variables whose horizontal partitioning maps into the observed set of ordered outcomes. Such an ordered response system allows the use of a general covariance matrix for the underlying latent variables, which translates to a flexible correlation pattern between the observed ordered outcomes.

Let q be an index for individuals ($q = 1, 2, \ldots, Q$) and let i be the index for ordered variables (i = 1, 2, …, I), where I is the total number of dependent variables for each individual. Let the observed level for individual q and variable i be k_{qi}. As seen in the previous paragraph the latent propensity y* for each category can be written as:

$$y^*_{qj} = \beta'_j x_{qj} + \varepsilon_{qj}, \quad y_{qj} = k \quad if\ \mu_j^{k-1} < y^*_{qi} < \mu_j^k \quad (1.37)$$

The ε_{qj} terms are assumed to be independent and identical across individuals (for each and all i). For identification reasons, the variance of each ε_{qj} term is normalized to 1. However, the model allows correlation in the ε_{qi} terms across variables i for each individual q. If $\varepsilon_q = (\varepsilon_{q1}, \varepsilon_{q2}, \varepsilon_{q3}, \ldots, \varepsilon_{qI})'$, then ε_q is multivariate normally distributed (MVN) with a mean vector of zeros and a correlation matrix Σ as follows:

$$\varepsilon_q \sim N\left[\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_{12} & \rho_{13} & \cdots & \rho_{1I} \\ \rho_{21} & 1 & \rho_{23} & \cdots & \rho_{2I} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{I1} & \rho_{I2} & \rho_{I3} & \cdots & 1 \end{pmatrix}\right] \quad (1.38)$$

Or

$\varepsilon_q \sim N[0, \Sigma]$

The off-diagonal terms of Σ capture the error covariance across the underlying latent continuous variables of the different ordinal variables. In other words, the off-diagonal terms Σ capture the effects of common unobserved factors that influence the propensity of ordered response levels for each attitudinal variable. As a special case, if all the correlation parameters are zero, the model system collapses to a set of independent ordered response probit models. The parameter vector of the multivariate ordered probit model is

$$\delta = \left(\beta_1', \beta_2', \dots, \beta_I'; \boldsymbol{\mu}_1', \boldsymbol{\mu}_2', \dots, \boldsymbol{\mu}_J'; \Omega'\right)' \tag{1.39}$$

where $\mu_j = \left(\mu_j^1, \mu_j^2, \dots, \mu_j^{K_i - 1}\right)'$ for j = 1, 2, …, J and Ω' are the off-diagonal terms of Σ matrix. Let the actual observed ordered response level for individual q and dependent variable i be m_{qi}. Then, the likelihood function (L) for individual q may be written as:

$$L_q(\delta) = \Pr(y_{q1} = m_{q1}, y_{q2} = m_{q2}, \dots, y_{qI} = m_{qJ}) \tag{1.40}$$

1.3.9. The Integrated Choice Latent Variable Model

Microeconomic theory has tended to consider decision makers as rational self-interested actors engaged in a constant process of evaluating the costs and benefits associated with any choice, trying to maximize their personal well-being given market constraints (Vij and Walker, 2016). Traditional discrete choice models have focused their analysis on observable variables, such as product attributes and socioeconomic characteristics, and treated consumers as optimizing black boxes with predetermined wants and needs is at odds (Vij and Walker, 2016). However, as seen in 1.3.1, this approach is in contrast with the findings from studies in the social sciences which have shown that choice behavior can also be influenced by psychological factors such as affections, attitudes, norms, and preferences (Ben-Akiva *et al.*, 1999).

Only in the last two decades have models of disaggregate decision-making started to include latent constructs for capturing the impact of subjective factors on choice process. These kinds of models, called hybrid choice models (HCM) or integrated choice and latent variable (ICLV) models, proposed in the 1980's by McFadden (1986) and Train *et al.* (1987), only became popular in 2002 following the work of Ben-Akiva *et al.* (2002). As a result, an increasing number of researchers have begun to adopt these models in transportation and logistics contexts. Some examples include the study of travel mode choice (Paulssen *et al.*, 2014; Vij *et al.*, 2013; Kamargianni and Polydoropoulou, 2013; Abou-Zeid and Ben-Akiva, 2011), route choice (Bhat *et al.*, 2015; Prato *et al.*, 2012), departure time (Thorhauge *et al.*, 2016), fuel/vehicle type choice (Daziano and Bolduc, 2013; Glerum *et al.*, 2013), freight (Bergantino *et al.*, 2013), *etc.*

1.3.9.1. Methodology Framework

In the general formulation of the ICLV models, two components can be distinguished: the latent variable model and the discrete choice model (Figure 2).

The latent variable model is composed of a set of *structural equations*, which describe the latent variables (*e.g.* attitudes, perceptions) in terms of observable exogenous variables, and a group of *measurement equations* that link the latent variable to indicators. Because the latent constructs are not observable, the analyst obtains information about them from observed responses to questions of a survey: the indicators, which can be continuous, binary or ordered variables. By simultaneously integrating discrete choice and latent variable models, the latent variables can be seen as explanatory variables included in the utilities of choice alternatives.

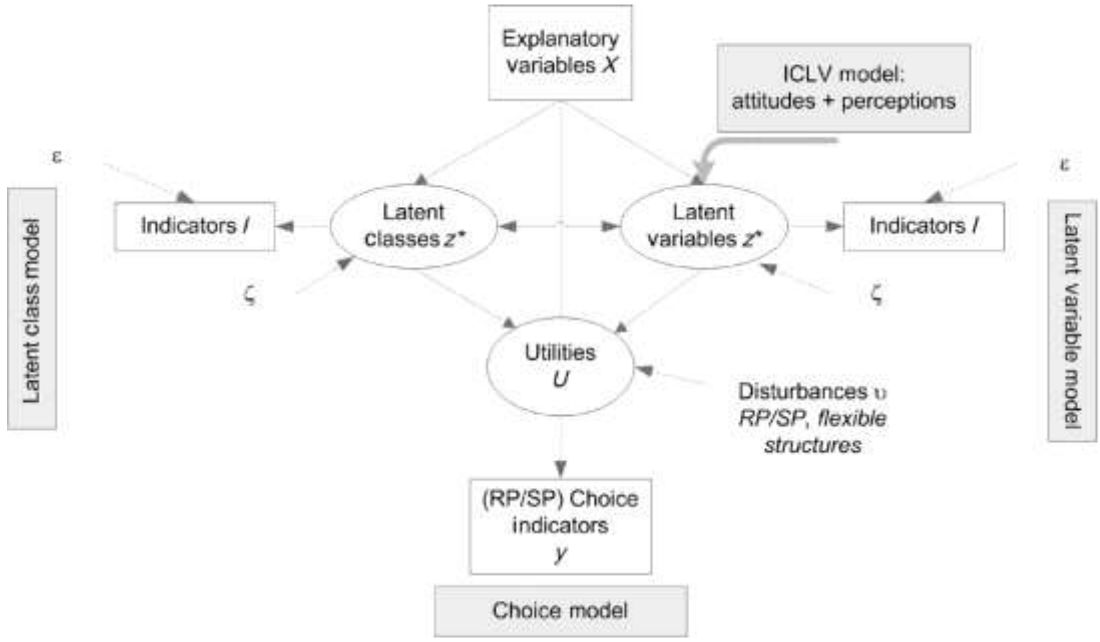


Figure 2 Framework for Integrated Choice Latent Variable model (from Ben-Akiva *et al.*, 2002 and Daziano and Bolduc, 2013)

1.3.9.1.1. Structural equations

For the latent variable model, the distribution of the latent variable given the explanatory variables is required. For example:

$$x_q^* = h(\boldsymbol{x}_q; \boldsymbol{\lambda}) + \omega_q \quad \text{and} \quad \omega_q \sim D(0, \Sigma_\omega) \tag{1.41}$$

where x_q is a vector of explanatory variables (observed or latent), λ is a vector of K parameters (to be estimated from data) and ω_q is the (random) error term. Note that the most common specification for the function h is linear:

$$h(x; \gamma) = \lambda_0 + \sum_{k=1}^{K} \lambda_k \boldsymbol{x}_k \tag{1.42}$$

For the discrete choice part of the model, the distribution of the utilities is

$$U_q = V(\boldsymbol{x}_q, \boldsymbol{x}_q^*; \boldsymbol{\beta}) + \varepsilon_q \text{ and } \varepsilon_q \sim D(0, \Sigma_\varepsilon) \quad (1.43)$$

Note that the systematic part of the random utility is a function of both observable and latent variables.

1.3.9.1.2. Measurement equations

The latent variable model requires the distribution of the indicators conditional on the values of the latent variables

$$I_q = m(x_q, x_q^*; \alpha) + \upsilon_q \text{ and } \upsilon_q \sim D(0, \Sigma_\upsilon) \quad (1.44)$$

where I_n is the reported value, x^* is the latent variable, x_q is a vector of explanatory variables, α a vector of parameters and υ_n is the random error.

The measurement equation of the discrete choice model is defined by a dummy variable that takes the value one if the alternative chosen has the highest utility, zero otherwise:

$$y_i = \begin{cases} 1, if\ U_i = max_j\{U_j\} \\ \quad 0\ \ otherwise \end{cases} \quad (1.45)$$

1.3.9.1.3. Likelihood function

If the latent variable were not present, the likelihood function would be

$$P(y|x; \beta; \Sigma_\varepsilon) \quad (1.46)$$

The choice model can take different forms, *e.g.*, logit, nested logit, probit, ordered probit, ordered logit, and can include the combination of different choice indicators such as stated and revealed preferences.

In a setting with the latent variables, if error components ε and ω are independent, the likelihood function is the integral of the choice model over the distribution of the latent constructs

$$P(y|\boldsymbol{x}; \boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}_\varepsilon, \boldsymbol{\Sigma}_\omega) = \int_{\boldsymbol{x}^*} P(y|\boldsymbol{x}, \boldsymbol{x}^*; \boldsymbol{\beta}, \boldsymbol{\Sigma}_\varepsilon)\, g(\boldsymbol{x}^*|\boldsymbol{x}; \boldsymbol{\lambda}, \boldsymbol{\Sigma}_\omega) d\boldsymbol{x}^* \quad (1.47)$$

which is an integral of dimension equal to the number of latent variables in $\boldsymbol{x}^*$ and g is the density function of the latent variable.

To characterize the unobserved latent variables and improve the accuracy of estimates of the structural parameters, indicators are introduced. Assuming the error components (*ω*, *ε*, *υ*) are independent, the joint probability of observing variables y and I is:

$$P(y, I|\boldsymbol{x}; \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\lambda}, \Sigma_\varepsilon, \Sigma_\upsilon, \Sigma_\omega) = \int_{\boldsymbol{x}^*} P(y|\boldsymbol{x}, \boldsymbol{x}^*; \boldsymbol{\beta}, \Sigma_\varepsilon) f(I|\boldsymbol{x}, \boldsymbol{x}^*; \boldsymbol{\alpha}, \Sigma_\upsilon)\, g(\boldsymbol{x}^*|\boldsymbol{x}; \boldsymbol{\lambda}, \Sigma_\omega) d\boldsymbol{x}^* \quad (1.48)$$

Note that the first term of the integrand corresponds to the choice model, the second term to the measurement equation from the latent variable model, and the third term to the structural equation from the latent variable model.

By jointly constructing the indicators in the latent variable model with the distribution function of the measurement relationship, the indicators do not only permit to identify the latent variables, but also provide efficiency in estimating the full model.

1.3.9.1.4. Distribution of the error terms

Different distributional assumptions about ε lead to different forms of the ICLV models. A common choice is to assume that each element of ε is i.i.d. Gumbel across alternatives and decision-makers with location zero and scale one, leading to a multinomial logit kernel for the discrete choice sub-model. If the vector ε is normally distributed with a mean of zero and covariance matrix Ω the discrete choice sub-model will be a multinomial probit (Bhat and Dubey, 2014; Kamargianni *et al.*, 2015).

Alternatively, the vector ε could be a mixture of normally distributed and Gumbel distributed vectors, resulting in the mixed logit kernel. Walker and Ben-Akiva (2002) have included the mixed logit model in the ICLV to include individual taste variations in terms of alternative attributes and the panel effect due to the combination of revealed preference and stated preference data. Similarly, other authors (see for example Daly *et al.*, 2012; Jensen *et al.*, 2013; Meloni *et al.*, 2013; Soto *et al.*, 2018) treated the panel effect using a mixed logit specification within the ICLV. With the mixed logit specification it is also possible to associate random coefficients with latent variables, whose effects over the choice process may vary significantly across individuals. For instance, Yáñez *et al.* (2010) estimated a mixed ICLV model with random parameter associated with cost and two latent variables.

1.3.9.1.5. Ordered choice model in the latent variable model

The indicators measuring the latent variables are normally measured with scales (*e.g.* Likert scale) leading to ordinary variables. Nevertheless, various authors (*e.g.* Habib and Zaman, 2012; Paulssen *et al.*, 2014; Scagnolari *et al.*, 2015) assume continuous indicators and linear relationships between the latent variables and their indicators, overlooking the ordinal nature of indicators. Daly *et al.* (2012) point out the inconsistency of this approach and suggest replacing the continuous specification by an ordered specification (ordered probit/logit). They conclude that this specification is qualitatively better as the use of an ordered model contributes to an improved explanation of choice behavior.

1.3.9.1.6. Estimation

To estimate the unknown parameters of the ICLV model maximum likelihood techniques are used. The model estimation process maximizes the logarithm of the sample likelihood function

$$max \sum_{q=1}^{Q} ln[P(y, I|\boldsymbol{x}; \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\lambda}, \Sigma_{\varepsilon}, \Sigma_{\upsilon}, \Sigma_{\omega})] \tag{1.49}$$

The likelihood function includes multi-dimensional integrals, with dimensionality equal to that of the integral of the underlying choice model plus the number of latent variables.

The proper way to estimate ICLV models is to use a simultaneous approach that jointly estimates the parameters of the structural equation and the parameters of the choice model, using both the indicators and the choice data (Bierlaire, 2018). The simulated maximum likelihood estimation method, like those developed for traditional mixed logit models, is used in the majority of studies (Bhat and Dubey, 2014). However, while these simulation techniques work quite well for the traditional mixed logit, their use in ICLV models has been problematic because of the prohibitive computational time of the estimation and convergence problems.

For these reasons, in some cases, sequential estimation is preferred. The sequential estimation method involves first estimating the latent variable model, using standard latent variable estimators. The second step is to include latent variables in the estimation process of the choice model. However, the two-step estimation method leads to consistent, but inefficient estimates. Bahamonde-Birke and Ortúzar (2014) introduced a method to correct the bias due to sequential estimation, though they only refer to the case of a multinomial logit model.

To address this issue, Bhat and Dubey (2014) proposed a different model formulation for the ICLV model, based on a multivariate probit (MNP) kernel and estimated using Bhat's maximum approximate composite marginal likelihood (MACML) inference approach (Bhat, 2011). With this approach, the dimensionality of integration in the composite marginal likelihood (CML) function that needs to be maximized is independent of the number of latent variables and the number of ordinal indicators variables, but only depends on the number of alternatives.

1.3.9.2. Criticism of ICLV models

Recent work in transportation research (Chorus and Kroesen, 2014) raised the question about the ability of ICLVs to derive policy implications that aim to change travel behavior, pointing out two potential issues. Firstly, latent attitudes and perceptions are partly endogenous with respect to travel behavior, and the analyst cannot be sure about the direction of causality in attitudes, perceptions, and choices. Moreover, latent variables and the observed choice could be influenced by the same underlying, unmeasured factors. Secondly, the data collected are typically cross-sectional, measured at a single moment in time. This means only between-person comparisons based on differences in latent variables can be evaluated, and not within-person comparisons and how latent variables may change over time as a result of variation in socio-economic (SE) characteristics or the implementation of travel demand management policies. However, Vij and Walker (2016) argue that even if forecasting by means of an ICLV model has no added value from an econometric point of view, they could still be a useful tool to explain mode choice and to suggest how public policies and TDM measures may be designed to promote sustainable mobility.

1.3.10. The Generalized Heterogeneous Data Model

The development and spread of new measurement tools and technologies has given rise to new ways of collecting and storing data. As a result, an increasing quantity of data (big data), often characterized by complex interdependent structures, is now available. These complex configurations often need to be treated with unconventional statistical methodologies, which sometimes require great computational effort. Multivariate data including mixtures of nominal outcomes, ordinal variables, count variables and continuous variables are an example of non-standard correlated data.

The easiest approach to handle these complex data structures is to carry out separate analysis and ignore the dependencies. However, this kind of approach is inefficient because it fails to account for any associations that exist between the mixed variables.

Another approach would be to transform discrete variables into continuous variables through some scoring scheme. Conversely, all the variables can be treated as discrete if the continuous ones are discretized with a grouping criteria treatment. However, one problem associated with these techniques is the subjectivity in the numerical scoring scheme adopted or the loss of information due to categorization of continuous variables (de Leon and Chough, 2013).

Recently, Bhat proposed a different way of treating mixed data, formulating the Generalized Heterogeneous Data Model (GHDM) (Bhat, 2015). The model is an evolution of the Integrated Choice and Latent Variable models developed by Bhat and Dubey (2014), and jointly handles mixed types of dependent variables by representing the covariance relationships among them through a reduced number of latent factors.

1.3.10.1. The GHDM formulation

The GHDM model consists of two components: the latent variable structural equation model (SEM), and the latent variable measurement equation model (MEM).

In the SEM component, latent variables relevant to the endogenous outcomes of the MEM system are hypothesized, based on theoretical psychological considerations and earlier qualitative/quantitative studies. These latent variables are expressed as linear functions of exogenous observed variables and stochastic error terms. A general covariance structure for the latent variables is adopted, therefore, no causal relationship between latent variables is allowed.

In the MEM component, the endogenous variables are described as functions of both latent variables and exogenous variables. The measurement equations have different characteristics depending on the type of dependent variable (continuous, ordinal, count, or nominal), however all have continuous underlying functions.

Let q be the index for individual $(q = 1,2, \dots, Q)$, which we will suppress in parts of the presentation below. Assume that all error terms in the GHDM model for an individual are independent of other individual error terms.

1.3.10.1.1. Structural Equation Model

Let z_l^* be the lth latent variable (l = 1, 2, …, L) for a specific person. Write z_l^* as a linear function of covariates:

$$z_l^* = \boldsymbol{\alpha}_l' \boldsymbol{w} + \eta_l, \tag{1.50}$$

where w is a $(\tilde{D} \times 1)$ vector of observed covariates (excluding a constant), α_l is a corresponding $(\tilde{D} \times 1)$ vector of coefficients, and η_l is a random error term assumed to be standard normally distributed for identification purposes (see Stapleton, 1978). Next, define the $(L \times \tilde{D})$ matrix $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_L)'$, and the $(L \times 1)$ vectors $z^* = (z_1^*, z_2^*, \dots, z_L^*)'$ and $\boldsymbol{\eta} = (\eta_1, \eta_2, \eta_3, \dots, \eta_L)'$. Let $\boldsymbol{\eta} \sim MVN_L[\mathbf{0}_L, \boldsymbol{\Gamma}]$, where $\mathbf{0}_L$ is an $(L \times 1)$ column vector of zeros, and $\boldsymbol{\Gamma}$ is an $(L \times L)$ correlation matrix. In matrix form, we may write Equation (1) as:

$$\boldsymbol{z}^* = \boldsymbol{\alpha w} + \boldsymbol{\eta}.$$

1.3.10.1.2. Measurement Equation Model Components

Consider N ordinal outcomes (indicator variables) for the individual, and let n be the index for the ordinal outcomes $(n = 1,2, \dots, N)$. Also, let J_n be the number of categories for the n^{th} ordinal outcome $(J_n \geq 2)$ and let the corresponding index be $j_n (j_n = 1,2, \dots, J_n)$. Let y_n^* be the latent underlying continuous variable whose horizontal partitioning leads to the observed outcome for the n^{th} ordinal variable. Assume that the individual under consideration chooses the a_n^{th} ordinal category. Then, in the usual ordered response formulation, for the individual, we may write:

$$\boldsymbol{y}_n^* = \boldsymbol{\gamma}_n + \boldsymbol{d}_n' \boldsymbol{z}^* + \boldsymbol{\varepsilon}_n, \text{ and } \psi_{n,a_n - 1} < y_n^* < \psi_{n,a_n} \tag{1.51}$$

where γ_n is a scalar constant, d_n is an $(L \times 1)$ vector of latent variable loadings on the n^{th} continuous outcome, the ψ terms represent thresholds, and ε_n is the standard normal random error for the n^{th} ordinal outcome. For each ordinal outcome, $\psi_{n,0} < \psi_{n,1} < \psi_{n,2} \dots < \psi_{n,J_n - 1} < \psi_{n,J_n}$; $\psi_{n,0} = -\infty$, $\psi_{n,1} = 0$, and $\psi_{n,J_n} = +\infty$. For later use, let $\psi_n = (\psi_{n,2}, \psi_{n,3} \dots, \psi_{n,J_n - 1})'$ and $\psi = (\psi_1', \psi_2', \dots, \psi_N')'$. Stack the N underlying continuous variables y_n^* into an $(N \times 1)$ vector y^*, and the N error terms ε_n into another $(N \times 1)$ vector ε. Define $\gamma = (\gamma_1, \gamma_2, \dots, \gamma_N)'$ $[(N \times 1)$ matrix$]$ and $d = (d_1, d_2, \dots, d_N)$ $[(N \times L)$ matrix$]$, and let $IDEN_N$ be the identity matrix of dimension N representing the correlation matrix of ε (so, $\varepsilon \sim MVN_N(0_N, IDEN_N)$; again, this is for identification purposes, given the presence of the unobserved z^* vector to generate covariance. Finally, stack the lower thresholds for the decision-maker $\psi_{n,a_n-1}(n = 1,2,\ldots,N)$ into an $(N \times 1)$ vector ψ_{low} and the upper thresholds $\psi_{n,a_n}(n = 1,2,\ldots,N)$ into another vector ψ_{up}. Then, in matrix form, the measurement equation for the ordinal outcomes (indicators) for the decision-maker may be written as:

$$\boldsymbol{y}^* = \boldsymbol{\gamma} + \boldsymbol{dz}^* + \varepsilon, \psi_{low} < y^* < \psi_{up} \tag{1.52}$$

Consider G nominal (unordered-response) variables $(g = 1,2,3,\ldots,G)$, with I_g being the number of alternatives corresponding to the g^{th} nominal variable ($I_g \geq 2$) and i_g being the corresponding index $\left(i_g = 1,2,3,\ldots,I_g\right)$. Let the individual under consideration choose the alternative m_g for the g^{th} nominal variable and assume the usual random utility structure for each alternative i_g:

$$U_{gi_g} = b'_{gi_g}x + \vartheta'_{gi_g}\left(\beta_{gi_g}z^*\right) + \varsigma_{gi_g}, \tag{1.53}$$

where x is a fixed $(A \times 1)$ vector of exogenous variables (including a constant), b_{gi_g} is an $(A \times 1)$ column vector of corresponding coefficients and $\varsigma_{gi_g}$is a normal error term. β_{gi_g} is a $\left(N_{gi_g} \times L\right)$ -matrix of variables interacting with latent variables to influence the utility of alternative i_g, and ϑ'_{gi_g} is an $\left(N_{gi_g} \times 1\right)$ -column vector of coefficients capturing the effects of latent variables and its interaction effects with other exogenous variables . Let $\varsigma_g = (\varsigma_{g1}, \varsigma_{g2},\ldots,\varsigma_{gI_g})'$ $(I_g \times 1$ vector), and $\varsigma_g \sim MVN_{I_g}\left(0, \Lambda_g\right)$.

Define $U_g = (U_{g1}, U_{g2},\ldots,U_{gI_g})'$ $(I_g \times 1$ vector), $b_g = (b_{g1}, b_{g2}, b_{g3},\ldots,b_{gI_g})'$ $(I_g \times A$ matrix), and $\beta_g = (\beta'_{g1}, \beta'_{g2},\ldots,\beta'_{gI_g})'$ $\left(\sum_{i_g=1}^{I_g} N_{gi_g} \times L\right)$ matrix. Also, define the $\left(I_g \times \sum_{i_g=1}^{I_g} N_{gi_g}\right)$matrix ϑ_g, which is initially filled with all zero values. Then, position the $\left(1 \times N_{g1}\right)$ row vector ϑ'_{g1} in the first row to occupy columns 1 to N_{g1} , position the $\left(1 \times N_{g2}\right)$ row vector ϑ'_{g2} in the second row to occupy columns N_{g1}+1 to $N_{g1} + N_{g2}$, and so on until the $\left(1 \times N_{gI_g}\right)$ row vector ϑ'_{gI_g} is appropriately positioned. Further, define $\varpi_g = (\vartheta_g\beta_g)(I_g \times L$ matrix), $\overleftrightarrow{G} = \sum_{g=1}^{G} I_g$, $\tilde{G} = \sum_{g=1}^{G}\left(I_g - 1\right)$, $U = \left(U'_1, U'_2,\ldots,U'_G\right)'$ $(\overleftrightarrow{G} \times 1$ vector), $\varsigma = (\varsigma_1, \varsigma_2,\ldots,\varsigma_G)'$ $(\overleftrightarrow{G} \times 1$ vector), $b = (b'_1, b'_2,\ldots,b'_G)'(\overleftrightarrow{G} \times A$ matrix), $\varpi = (\varpi'_1, \varpi'_2,\ldots,\varpi'_G)'(\overleftrightarrow{G} \times L$matrix), and $\vartheta_{vec} = \text{Vech}(\vartheta_1, \vartheta_2,\ldots,\vartheta_G)$ (that is, ϑ_{vec} is a column vector that includes all elements of the matrices $\vartheta_1, \vartheta_2,\ldots,\vartheta_G$). Then, in matrix form, we may write Equation (6) as:

$$U = bx + \varpi z^* + \varsigma,$$

where $\varsigma \sim MVN_{\overleftrightarrow{G}}\left(\mathbf{0}_{\overleftrightarrow{G}}, \Lambda\right)$, with Λ as follows:

$$\mathbf{\Lambda} = \begin{bmatrix} \mathbf{\Lambda}_1 & 0 & 0 & 0. \cdots 0 \\ 0 & \mathbf{\Lambda}_2 & 0 & 0 \cdots 0 \\ 0 & 0 & \mathbf{\Lambda}_3 & 0 \cdots 0 \\ \vdots & \vdots & \vdots & \vdots \ldots \vdots \\ 0 & 0 & 0 & 0 \ldots \mathbf{\Lambda}_G \end{bmatrix} (\vec{G} \times \vec{G} \text{ matrix }), \tag{1.54}$$

Let δ be the collection of parameters to be estimated: $\delta = [\text{Vech}(\alpha), \text{Vech}(\Gamma), \gamma, \text{Vech}(\text{d}), \psi, \vartheta_{vec}, \text{Vech}(\Lambda)]$, where the operator "$\text{Vech}(\cdot)$" vectorizes all the non-zero unique elements of the matrix/vector on which it operates. We will assume that the error vectors $\boldsymbol{\eta}$, ε, and ς are independent of each other. Additional details on the GHDM formulation, including sufficiency conditions for identification of model parameters, and the MACML estimation approach for the formulation may be found in Bhat (2015).

1.3.11. Goodness of fit and model selection

1.3.11.1. The use of the t-test for significance of coefficient estimates

The t-test is mainly used to test if an estimated parameter in the model is significantly different from some known constant, often zero. The t-test for nonlinear models is used in the same way as the t-test in nonlinear regression, except that it is valid only asymptotically and thus for large samples. The critical values at which we reject the null hypothesis and hence we accept that an attribute has a relevant effect are percentiles of a standardized normal distribution, which for two-tailed tests at the usually 10% and 5% level of significance are ±1.65 and ±1.96 respectively.

1.3.11.2. The likelihood ratio test

For more complex hypotheses than the ones regarding individual parameters, the likelihood ratio test can be used. To determine if a model is statistically significant, it is possible to test the null hypothesis that all the coefficients, except for the alternative specific constant, are zero. The test statistic is

$$-2(LL_{market\ share\ model} - LL_{estimated\ model}) \tag{1.55}$$

where $LL_{market\ share\ model}$ is the loglikelihood of a model with only constants (market share model) and $LL_{estimated\ model}$ is the loglikelihood of the estimated model. The statistic is asymptotically χ^2 distributed with $K - J + 1$ degrees of freedom, where K is the number of parameters and J is the number of alternatives in the choice set.

Another useful application of the likelihood ratio test is to test if an explanatory variable is generic or alternative-specific, with the former possessing the same weight across all alternatives. The likelihood ratio test statistic for the null hypothesis of generic attribute is

$$-2(LL_{restricted\ \ model} - LL_{estimated\ model}) \tag{1.56}$$

Where $LL_{restricted\ model}$ denote the loglikelihood of the model with a generic specification (restricted model). The statistic is asymptotically χ^2 distributed with number of degrees of freedom equal to the number of linear restrictions.

1.3.11.3. Goodness of fit measures

To evaluate how a model fits a set of observations and compare models with different specifications it is useful to calculate goodness of fit measures. A commonly used statistic is the likelihood ratio index (rho squared):

$$\rho^2 = 1 - \frac{LL_{estimated\ model}}{LL_{EL}} \tag{1.57}$$

Where LL_{EL} is the loglikelihood of the model estimated with all the parameters equal to 0 (equally likely model). ρ^2 is a function of the sample data, the choice set size and the number of parameters, hence it can be used only for comparing models when the previous are the same. For this reason, we also use the adjusted likelihood ratio index (rho-squared bar):

$$\overline{\rho}^2 = 1 - \frac{LL_{estimated\ model} - K}{LL_{EL}} \tag{1.58}$$

where K is the number of parameters. The introduction of a penalty term for the number of parameters permits to compare models with different number of attributes.

1.3.11.4. The Percentage Right or First Preference Recovery (FPR) Measure

The Percentage right or First Preference Recovery (FPR) proposed by Foerster (1979), is an aggregate measure that calculates the percentage of individuals who actually choose the alternative to which the model assigns the greatest probability. However, the FPR could be an ambiguous criterion of model performance in the sense that too high a value of FPR may be an indication of poor model performance, rather than *viceversa*. Because of this, for any model, we compare the FPR with the chance recovery (CR) and the expected value of FPR (ER). The chance recovery is given by:

$$CR = \sum_q \frac{\left(\frac{1}{N_q}\right)}{Q} \tag{1.59}$$

where Q is the number of individuals and N_q the number of available alternatives for each individual q. The expected value of FPR can be expressed as:

$$ER = \sum_q P_q \tag{1.60}$$

where P_q is the computed (maximum) probability associated with the best option for individual q. If the three measures are relatively close the model is reasonable but uninformative; if FPR and ER are similar, with a value larger than CR, the model is reasonable and informative; finally, if FPR is very much larger or very much smaller than expectations, we have to conclude that the model does not explain the variation in the data (inconsistent data) and should be rejected (see Gunn and Bates (1982) and Ortúzar and Willumsen (2011) for a complete discussion of the problem).

1.3.12. Aggregation and forecasting

Discrete choice models allow us to estimate probabilities at the level of individual decision makers. However, the analyst is often interested in the prediction of some aggregate measure, so that he can make forecasts and examine the average response to change in some factor. We report here three aggregate indicators: market share, elasticities and willingness to pay (WTP).

Market share. For a population of Q individuals, the market share of an alternative I according to the model is the weighted sum of the probabilities of each individual:

$$P_{iQ} = \frac{1}{Q}\sum_{q} w_q P_q(i) \tag{1.61}$$

Elasticities. Because choice probabilities are a function of observed variables, it is useful to know to what extent these probabilities vary in response to a change in one of the explanatory variables. The calculation of demand elasticities is a common way to address this problem.

The aggregate direct elasticities, defined as the percent change in the market share for alternative i following a change of one percent in the value of a variable x_i, is

$$E_{x_i}^{i} = \frac{\sum_{q=1}^{Q} w_q P_q(i) E_{x_{iq}}^{i}}{\sum_{q=1}^{Q} w_q P_q(i)} \tag{1.62}$$

where x_i is an attribute of alternative i, $w_q$is the sample weight of observation q, $P_q(i)$ is the probability that an individual q chooses alternative i and $E_{x_{iq}}^{i}$ is the disaggregate direct elasticity of the demand for observation q with respect to a variation in variable x_{iq}. The disaggregate elasticity is:

$$E_{x_{iq}}^{i} = \frac{\partial P_q(i)}{\partial x_{iq}} \frac{x_{iq}}{P_q(i)} \tag{1.63}$$

It is also possible compute the extent to which the probability of choosing an alternative *i* changes if there is a variation in an observed variable of alternative *j*. The aggregate cross elasticities are given by the following expression:

$$E_{x_j}^{i} = \frac{\sum_{q=1}^{Q} w_q P_q(i) E_{x_{jq}}^{i}}{\sum_{q=1}^{Q} w_q P_q(i)} \tag{1.64}$$

where $E^i_{x_{jq}}$ is the disaggregate cross elasticity for individual q with respect a change in variable x_{jq}. The disaggregate cross elasticity is

$$E^i_{x_{jq}} = \frac{\partial P_q(i)}{\partial x_{jq}} \frac{x_{jq}}{P_q(i)} \tag{1.65}$$

To summarize, an aggregate cross elasticity represents the percentage change of probability in the market share in response to a change of 1% in the value of the attribute x_{jq}.

Willingness to pay. Willingness to pay measures the amount of money at or below which a consumer will definitely pay for one unit of change in the attribute x_i of choice alternative *i*. It is given by:

$$WTP_q(i) = \frac{\partial U_q(i)/\partial x_i}{\partial U_q(i)/\partial cost_i} \tag{1.66}$$

The aggregate WTP over a sample of Q individuals is:

$$WTP(i) = \sum_{q=1}^{Q} w_n WTP_q(i) \tag{1.67}$$

PART 2 - ASSESSMENT AND FORECASTING OF BIKE USE IN AN URBAN CONTEXT

2.1. INTRODUCTION

A key element of the multi-dimensional toolbox of sustainable transportation strategies aimed at easing traffic congestion and reducing environmental impact is the encouragement of the use of active transportation travel modes such as walking, bicycling and bicycle-sharing. In particular, there has been a surge of interest in bicycling (Pucher and Buehler, 2017) as a physically active transportation option able to bring immediate multiple benefits at both an individual as well as a community level. These physical activity benefits include a reduction in the incidence of obesity and the concomitant mental and physical health benefits, improved cardiovascular fitness and decreasing heart disease, diabetes, high blood pressure, and several forms of cancer side effects, reduced community parking space that can then be utilized for green use, no mobile source emissions, affordable accessibility to activity locations, and flexibility in departure time compared to public transport.

Understanding why in many countries where cycling share is low, individuals perceive the bicycle as a form of recreation more than as a means of transport, is of special interest. This is so even though many localities and communities have invested in infrastructure and policies to promote bicycling, including separate bicycling lanes, continuous bicycle facilities, safe and secure bicycle parking facilities, car parking prohibitions on roadways (or sections of roadways close to intersections), reduced speed limit laws, and even shower/locker rooms at work places. In Italy, as reported by Legambiente in its 1st Report on Bike Economy and Urban Cycling in Italy, published in 2017, despite the fact that between 2008 and 2015 the cycling infrastructure in major cities has increased by 50%, in the same period the percentage of Italians using the bicycle as a mode of transportation has remained unchanged. In fact, it was 3.6% in 2008 and remained at 3.6% in 2015.

Hence, it is not surprising that many studies have been aimed at better understanding what deters individuals from bicycling to work (and, conversely, what encourages it). And this body of research has been rapidly growing only within the past few years. However, there are some research gaps in investigating biking behavior and more in-depth research will benefit transport planners and public health officials in improving their understanding of urban biking behavior. One is the consideration of the choice of biking for different purposes under a comprehensive analytical framework, which would examine behavioral linkages between them. The other is accommodating psychological variables within a modeling framework and understanding to what extent they affect the choice to use the bike. In fact, the traditional random utility modeling approach failed in some cases to explain the decision process leading people to cycle, making it necessary to incorporate those psychological factors attached to bicycle use and derived from its special characteristics.

This work provides knowledge and tools for the proper incorporation of the bicycle as an option in travel demand models, with different methodologies developed to contribute to this line of research. To achieve this general object, several steps were taken. In particular:

- We investigate the impact of habitual behavior in the choice of biking to work, estimating a mode choice model where biking is one of four alternative modes
- We study specifically whether psycho-attitudinal factors vary among people with different cycling experience, for any purpose, and to quantify the determinants influencing cycling frequency
- We explore the relationships among psycho-attitudinal factors related to bicycling, urbanization level and socio-demographics for bicycle commuting and cycling for other purposes
- We examine how facilitators to cycling are perceived by different segments of individuals, in view of assessing how to best promote cycling in an urban area.

For this purpose, we designed a survey in order to explore all those attributes affecting bicycle choice, including the less investigated ones (socio-economic characteristics, physical characteristics, psycho-attitudinal factors or/and trip characteristics).

The survey was conducted among a group of employees from different local authorities but within similar territorial contexts (fairly homogeneous in socio-economic terms) who commute daily to work to explore why, under similar conditions, individuals choose different travel modes.

The chapter is organized as follows: Section 2.2 provides a literature review of the different determinants influencing the choice to cycle for transport and of the different methodological frameworks adopted. Section 2.3 describes the survey designed for data collection and provides a descriptive analysis for highlighting differences, if any, in socio-economic characteristics, physical characteristics, psycho-attitudinal factors or/and trip characteristics among different segments of individuals. The remaining sections describe the different model results.

2.2. DETERMINANTS ASSOCIATED WITH BICYCLE USE

Several studies have investigated the key variables associated with bicycle use, especially those related to the bicycle mode choice problem (Heinen *et al.*, 2010; Muñoz *et al.*, 2016). In their review Muñoz *et al.* (2016) identify two categories of explanatory variables, depending on how they are measured: objective, related to individual socio-economic characteristics, trip characteristics, built environment, natural environment and cycling facilities, and subjective, associated with the personal sphere of each individual.

2.2.1. Socio-economics characteristics

Socio-economics characteristics have always been an important category of determinants in travel behavior research. Regarding gender, most conclude that men cycle more than women, but this tendency could be related to cycling culture (Heinen *et al.*, 2010; Fernández-Heredia *et al.*, 2014). Some studies (*e.g.* Garrard *et al.*, 2008; Emond *et al.* 2009; Akar *et al.* 2013) report that in countries with low rates of cycling males are more likely to use bike than females. Both Garrard *et al.* (2008) and Bhat *et al.* (2015), the former in the Australian, the latter in the US context, suggest that women tend to be more safety-conscious than men and are more likely to perceive the negative consequences of sharing roads with motorized vehicles. By contrast, in countries with high cycling rates, such as the Netherlands and Denmark, cycling is more evenly spread over the two genders (Heinen *et al.*, 2013).

Instead, the relation between cycling and age seems to be unclear. Some state that cycling levels decline with age (Dill & Voros 2007; Pucher *et al.* 1999; Heinen *et al.*, 2013; Muñoz *et al.*, 2016) but others (Bhat *et al.*, 2015; Wardman *et al.*, 2007; De Geus, 2007) declare that age is not a significant determinant.

Also, the relation between cycling and level of income is ambiguous. If some (Stinson and Bhat, 2004; Dill and Voros, 2007) observe a positive relationship between income and use of the bike, Parkin *et al.* (2008) affirm that in England there is a link between lower incomes and lower bicycle share. This latter finding is in line with studies conducted in other contexts (Schwanen and Mokhtarian, 2005; Guo *et al.*, 2007).

The level of education is another factor that can be related to bike usage, though previous research has reported mixed findings. For example, Bhat *et al.* (2017) find that highly educated individuals have greater propensity to cycle. They suggest that a high education level is associated with a high level of environmental awareness and consequently, with a greater tendency to use environmental-friendly modes of transportation. Conversely, Cervero *et al.* (2009) show that the likelihood to cycle for non-recreational activities decreases with education level.

In general, car availability is widely reported as being negatively related to bicycle choice (Wardman *et al.*, 2007; Heinen *et al.*, 2013; Fernández-Heredia *et al.*, 2016; Bhat *et al.*, 2017). Interestingly, Ton *et al.* (2019) do not find any relationship between the availability of cars and bicycle use in the Netherlands.

Household size and composition are known to be related to mode choice as well. Some studies reported that the number of children (Bhat *et al.*, 2017) or the number of household members (Fernández-Heredia *et al.*, 2016), are negatively associated with bike commuting.

2.2.2. Trip characteristics

Trip characteristics, such as travel time and cost, are widely recognized as key aspects in travel behavior. Travel time has been found to have a huge influence on travel mode choice (Börjesson and Eliasson, 2012), also on bicycle use. Previous studies have shown that cyclists prefer short travel times (Stinson & Bhat 2005; Fu & Farber, 2017). Hunt and Abraham (2007) found that biking is competitive with all other motorized means of transport under certain distances. Akar and Clifton (2009) discovered that people with flexible departure time were more likely to cycle. Only a few studies have examined the association between cost and cycling (Handy *et al.*, 2014). Cycling is almost free, and for Bergström and Magnusson (2003) this is one of the reasons why commuters choose to cycle. Other research shows that also the cost of other means of transport, *e.g.* parking tolls and fees, influence the choice to cycle (Rodríguez, and Joo, 2004; Buehler, 2012).

Beside travel time, distance is one of the most common factors taken into consideration (Heinen *et al.*, 2010). In general, it has been found that the greater the distance the lower the bike share in mode choice for commuting (Parkin *et al.*, 2008; Heinen *et al.*, 2013). Non-cyclists often indicate long distances as a deterrent to commute by bike (Dickinson *et al.*, 2003; Stinson e Bhat, 2004). Moreover, there could be a maximum admissible travel distance that differs among individuals and gender (Heinen *et al.*, 2010). Regarding gender, some studies suggest that women tend to cover shorter distances by bike than men (Howard and Burns, 2001; Garrard *et al.*, 2008)

2.2.3. Natural environment

Natural environment is another important explanatory factor in bike choice for utilitarian purposes. In contrast to motorized transport, cycle choice is strongly determined by topography, weather and climate (Heinen *et al.*, 2010). Hilliness and the presence of steep slopes have a negative effect on bike use (Cervero & Duncan 2003; Parkin *et al.*, 2008; Cervero *et al.*, 2009; Winters *et al.*, 2010; Goetzke & Rave, 2011; Buehler & Pucher, 2012), and maximum gradient seems more important than average gradient (Menghini, *et al.*, 2010). A well-known example of this is bike share in the City of York (UK) and the City of Bradford (UK) (Heinen *et al.*, 2010). York, with slopes of more than 3% on only 5% of its surface area, has a cycling share of 13.1%. Bradford, which is characterized by steep slopes throughout its surface area, only has a cycling share of 0.8%. Moreover, Motoaki & Daziano (2014) found that hilly topography (slope inclination) has a relation with the physical condition of the cyclist, *i.e.* the fitter the cyclist, the less bothersome a steeper route.

The weather (temperature, rain, wind, snow) can influence bicycling as well. Previous research shows that cycling is more common in summer (Stinson & Bhat, 2005; Guo *et al.*, 2007; Sener *et al.*, 2009; Böcker *et al.*, 2013). Rain and snow are considered as some of the most unfavourable weather conditions that stop people from cycling (Nankervis, 1999; Sears *et al.*, 2012; Motoaki & Daziano, 2015). A less widely considered variable is temperature, that has been found to have a non-linear effect, with cold and very hot weather negatively associated with bike use (Ton *et al.*, 2019).

Another seasonality aspect is linked to the number of hours of daylight (Muñoz *et al.*, 2016). Stinson and Bhat (2004) and Gatersleben and Appleton (2007) showed that darkness has a negative effect on cycling. In particular, women are more concerned about this aspect (Bergström and Magnusson, 2003; Cervero and Duncan, 2003).

2.2.4. Built environment

A large body of literature has investigated the relationship between the built environment and bike travel behavior. Different terms have been used when referring to the built environment. *Urban design* refers to the design of buildings and infrastructure within an urban area and can include both their arrangement and their appearance in connection with their function (Handy *et al.*, 2002). *Land use* is the way in which human activities are distributed across space, including the location and density of different activities, that can be classified in residential, commercial, office, industrial, and other activities. The *transportation system*, in the specific case of bike use, includes the physical infrastructure of roads, bike paths, bike trails, bridges, and so on, as well as the presence of those complementary infrastructures such as showers and dressing rooms. Here we consider the *built environment* consisting of urban design, land use, and the transportation system.

2.2.4.1. Urban design and land use

Looking in more detail at the urban design and land use problem, it has been observed that higher levels of urban density positively influence the decision to ride a bike (Baltes, 1996; Cervero and Duncan, 2003; Pucher and Bueheler, 2006; Zharan *et al.*, 2008; Parkin *et al.*, 2008; Fraser and Lock, 2011; Wang *et al.*, 2016; Braun *et al.*, 2016). This can be explained by the fact that, compared to low-density areas, denser urban areas are characterized by shorter distances between origins and destinations. Interestingly, Witlox and Tindermans (2004) found that depending on the place of residence there are differences with regard to the chosen mode, with residents in urban areas more likely to use the bike than in the suburbs. Dill and Voros (2007) observed that individuals who live in neighborhoods closer to downtown were more likely to make utilitarian bike trips. Stinson and Bhat (2004) found a higher bicycle commuting propensity for individuals residing in urban areas compared with those residing in suburban and rural areas.

Beside density, another key factor is the land-use mix, which depends on the level of diversity of land-use types (commercial industrial, residential and so on) across a neighborhood. Areas with a traditional layout, street-level shops and residences above make travel distances smaller and so easier to cycle from home to shops or places of work (Cervero and Duncan, 2003;

Pucher and Buehler, 2006; Heinen *et al*., 2010; Habib *et al.*, 2014; Braun *et al.*, 2016; Muñoz *et al.*, 2016; Winters *et al.*, 2017; Ton *et al.*, 2019). Another important element regards the aesthetics, as it has been shown that the presence, among others, of parks, street plants and garbage bins are positively associated with cycling (Sacks, 1994; Heath *et al.*, 2006; Fraser and Lock, 2011; Habib *et al.*, 2014; Wang *et al.*, 2016; Ton *et al.*, 2019).

Despite the large number of studies on cycling and neighborhood characteristics analysis, far less attention has been given to the residential self-selection problem. Many of the studies listed above considered the characteristics of residential location as an exogenous variable in the decision to cycle, ignoring the possibility of the residential neighborhood choice process of households (Pinjari *et al.*, 2008). In fact, bike travel may not only be influenced by residential location, but individuals could choose their home because they intend to cycle, preferring to live in areas that allow them to do so easily. Although this issue has been largely investigated in studies of walking as well as travel behavior more generally (see Bhat and Guo (2007) and Cao *et al.* (2009) for a complete overview of the problem), little research exists on the role of residential preference specifically influencing bicycle use. Pinjari *et al.* (2008) present a joint model of residential neighborhood type choice and bike ownership, showing that ignoring self-selection may lead to an underestimation of the impact of neighborhood attributes on bicycle ownership. Pinjari *et al.* (2011) use an integrated simultaneous multi-dimensional choice model to capture the jointness of residential location, auto ownership, bicycle ownership, and commute tour mode choices. They found that some socio-demographic variables influence both bike ownership and residential location choices, indicating, for the authors, the presence of a residential self-selection effect. However, one limitation of these works is that their sub models are only a function of socioeconomic and trip characteristics variables and do not include in their analysis any psychological factors. Another limitation concerns the use of bike ownership level as a dependent variable, assuming that the ownership of a bicycle automatically leads to its use. Recently, Ettema and Nieuwenhuis (2017) explored whether and to what extent built-environment factors, travel attitudes and reasons for location choice affect the use of different travel modes within two years after relocation. They show that active travel attitudes positively influence cycling frequency and the variable cycling accessibility, being a reason for location choice, has a stronger impact, in terms of magnitude, than locational factors.

2.2.4.2. Infrastructure

Several studies have revealed the existence of a relationship between bike infrastructure and cycling levels, as it may encourage cycling by augmenting awareness and raising visibility, increasing perceived safety and decreasing the number of conflict points with motorized vehicles and pedestrians (Braun *et al*., 2016). Bike infrastructure can come in different forms:

- in mixed traffic, where cyclists share the full roadway with other traffic without any longitudinal separation;
- on-road bicycle lanes, usually designated with a white stripe, a bicycle icon on the pavement and signage;
- cycle tracks, similar to bike lanes but physically more separated from motorized vehicles, for example with a curb, vehicle parking, or other barriers;
- off-street paths, completely separated from motor vehicle traffic and usually designed to accommodate non-motor vehicle traffic.

Results from aggregate cross-sectional studies in Europe and in US indicate that cities with more extensive networks of bicycle lanes and paths have higher shares of bicycle commuting (Dill and Carr, 2003; Parkin *et al.* 2008; Buehler and Pucher, 2012; Santos *et al.*, 2013). Using the data of 43 cities in US, Dill and Carr (2003) show that that for typical U.S. cities with a population of more than 250,000, each additional mile of bike lanes per square mile is associated with a one percent increase in the share of workers commuting by bicycle. Buehler and Pucher (2012) use a dataset on the length of bike lanes and paths in 2008 collected directly from 90 of the 100 largest U.S. cities to show a similar positive association between bike commute rates and the presence of both off-street paths and on-street lanes. In the attempt to identify factors that influence modal split for journeys to work in 112 medium-size cities in Europe, Santos et *al.* (2013) found that bicycle share is positively associated with length of the bicycle network.

Disaggregate revealed-preference studies indicate that a bike network could increase the likelihood to bicycle for utilitarian purposes (Hunt and Abraham, 2007; Dill and Voros, 2007; Bhat *et al.*, 2017) as well. Some works report in general a predilection for bicycle paths to both bicycle lanes and roads without bicycle facilities (Braun *et al.*, 2016). However, in some cases the findings of these studies are contradictory. For instance, Krizek and Johnson (2006) found that living within 400 m of an on-street bicycle lane but not of off-street facilities, was associated with a greater likelihood to use the bike in the Twin Cities of Minneapolis and St. Paul, Minnesota. On the other hand, Moudon *et al.* (2005) report that close proximity to bike paths in Seattle, Washington, increases the likelihood to cycle by 20%, but they found no effect for bike lanes, while Winters *et al.* (2010) found that neither type of infrastructure was correlated with cycling in Vancouver, British Columbia.

Also stated-preference studies investigated the preference for different infrastructure types. Aultman-Hall *et al.* (1997) use GIS software to investigate bicycle commuter routes in Guelph, Canada. Their results indicate that cyclists do not divert too much from the shortest path and found little use of off-road trails, with a preference for in-traffic facilities. Abraham *et al.* (2002) investigate cyclist preferences in the context of route choice in Calgary, Canada, finding that cyclists prefer off-street cycling facilities and low-traffic residential streets. Stinson and Bhat (2003) found that respondents preferred bicycling on minor residential streets to major arterials, likely because of the low traffic volumes on residential streets. They also detected a positive effect linked to the presence of some types of bicycle facilities and a preference for routes designed for bicycle use, with bicycle lanes being the most preferred facility type, followed by separate paths. Wardman *et al.* (2007) used RP/SP data to understand the likelihood of cycle commuting, finding that the best forecasting scenario to increase the number of bike commuters over the base situation is when all travel time would be spent on a completely segregated cycleway.

This whole discussion is further complicated by the influence of personal characteristics. In fact, preferences for different types of infrastructure may differ across socio-economic groups and across experienced and non-experienced cyclists (Heinen *et al.*, 2010). Some studies provided evidence that women have stronger preferences for segregation from motorized vehicles than men (Gardner, 1998; Berggren *et al.*, 2012; Caulfield *et al.*, 2012; Akar *et al.*, 2013; dell'Olio *et al.*, 2014; Aldred *et al.*, 2017). Regarding age, some works reported that older people prefer cycle infrastructure separated from motor traffic, others detected no differences (Aldred *et al.*, 2017). Stinson and Bhat (2005) found that, for both experienced and inexperienced users, separation from motorized traffic through separate bicycle paths or clearly designated bicycle lanes are among the most important attributes in route choice decisions, with all these variables considered more important by inexperienced users. Taylor and Mahmassani (1996) observed that for experienced cyclists bike lanes are not considered to be more attractive than wide curb lanes. However, it is worth highlighting that many of these studies generally measure the preferences of existing cyclists rather than the ability of facilities to entice new cyclists (Handy *et al.*, 2014).

The question of bicycle infrastructure is very much related to safety. Two types of safety can be identified: objective and subjective safety. Objective safety is 'real' safety for cyclists, measured in terms of the number of bicycle-related incidents per million inhabitants. Subjective safety refers to how individuals perceive safety and is mostly measured in terms of the stated safety experience of users or other respondents. Most cyclists are reluctant to cycle on a particular infrastructure or across an intersection if they perceive it as dangerous (Lawson *et al.*, 2013). Therefore, understanding subjective measures of safety plays an important role in cycling promotion.

The mere provision of bicycle facilities may not be sufficient to encourage use of the bike (Fernández-Heredia *et al.*, 2014). It has been shown that the provision of a well-designed network, with direct routes and a small number of stops, clearly contribute to the attractiveness of the bicycle as a transport mode (Rietveld and Daniel 2004). In fact, often cities have bike lanes that go nowhere, end in unsafe conditions, or pass through dangerous intersections. Different studies indicate that continuity and connectivity are key ingredients to attract users and make cycling more viable and comfortable for the everyday commuter. A work by Caulfield *et al.* (2012) revealed that direct routes with short journey times were found to be the most important positive variable for existing cyclists and non-cyclists in determining route choice. Similarly, Bhat *et al.* (2015) found that routes with a continuous bicycle facility (the whole route has a bicycle lane or wide outside lane) are preferred to those with a discontinuous facility.

2.2.4.3. End-of-trip facilities and transit integration

Fewer studies have considered the role of end-of-trip facilities such as bicycle parking and showers, although their presence can positively contribute to cycling.

Bicycle parking can protect bicycles from theft, damage, and weather and their presence may facilitate cycling (Heinen and Bueheler, 2019). The vast majority of studies show a positive relationship between parking at workplaces and bicycle mode choice. For example, Wardman *et al.* (2007) found an improvement in cycle market share, introducing bike parking at work, particularly the provision of indoor parking. Hunt and Abraham (2007) established that the provision of bike parking at destination has a positive effect equivalent to a reduction of 26.5 minutes cycling in mixed traffic. They also report that for young cyclists, particularly for under-16s, secure parking is more important than for other age groups. Only a few papers have focused on bicycle parking at the residential location and only one (Nkurunziza *et al.*, 2012) shows a positive relationship between the presence of parking at home and the likelihood to cycle. Bike parking is also a key aspect in integrating bicycling with transit. In particular in Europe and in Japan a large amount of bike parking has been provided at both suburban rail and metro stations (Martens, 2007; Harden, 2008; Pucher and Buehler, 2009; Pucher *et al.*, 2010), while their presence is less common at bus stops, due to the lack of bike racks on buses (Pucher et *al.*, 2010).

Besides the above, another determinant in the choice to use the bike is the presence of complementary infrastructure such as showers, dressing rooms and lockers on site (Abraham *et al.*, 2002; Wardman *et al.*, 2007; Sener *et al.*, 2009; Heinen *et al.*, 2013; Hamre and Buehler, 2014). Interestingly, Abraham *et al.* (2002) found that cyclists and non-cyclists value the provision of showers to the same extent. However, the research findings in this area are ambiguous. Taylor and Mahmassani (1996) found that showers at workplace were a disincentive to bike and ride for males and was not relevant for females. In the same vein, Stinson and Bhat (2004) found that the presence of showers and clothing lockers were not relevant variables in modeling the propensity to commute by bike.

2.2.5. Bicycle access

People are not able to cycle if they do not have access to a bicycle, and several studies indicate that the possession of a bicycle is the strongest predictor of bicycling in transportation (Wardman *et al.*, 2007; Heinen *et al.*, 2013; Fernández-Heredia *et al.*, 2016).

Interestingly, some studies consider bike ownership as the dependent variable in their modeling framework. Yamamoto (2009) studied the effects of the built environment on motorcycle and bicycle ownership as well as car ownership in Japan and Malaysia. Handy *et al.* (2010) estimate a nested logit model to jointly examine bike ownership and level of cycling experience jointly in six small US cities. Habib *et al.* (2014) estimate a joint econometric model of bike usage, choice of purposes of biking and bike ownership level, in Toronto. Maness and Cirillo (2016) present a latent class discrete choice model to explore the effect of social influence on bike ownership in US. Recently, Huang *et al.* (2017) applied a joint mixed multinomial logit-ordered model to explore the impacts of metro transit on the ownership of four mobility instruments, including the bicycle, in China.

The importance of bicycle access in the choice to cycle is also demonstrated by the number of programs, facilitating bike ownership or enabling temporary use of a bicycle, implemented throughout the world (Pucher *et al.*, 2010). In particular, bike sharing programs, which have been around since 1960s, have recently begun to receive large scale acceptance as a feasible transportation alternative (Fishman, 2016; Nikitas, 2018; Barbour *et al.*, 2019). In Barcelona the proportion of trips by bicycle increased from 0.75% to 1.76% (Demaio, 2009) and in Paris from 1.0% to 2.5% (Nadal, 2007) thanks to bike sharing systems. Fishman *et al.* (2016) using the data from bike share programs in Melbourne, Brisbane, Washington D.C., London, and Minnesota found a significant reduction in motor vehicle use due to the presence of this kind of system. Li *et al.* (2018), who studied bike sharing in East Asia, reported benefits such as reductions in greenhouse gas emissions and fuel consumption and increased public transport use.

2.2.6. Subjective factors

2.2.6.1. Perceptions of environmental and cycling facilities

Recent research has started to focus on how individuals can be influenced in their choice to cycle by their perceptions related to variables that can be measured objectively, such as hilliness, weather, traffic risks, distance and bicycle facilities. In fact, perceptions are subjective representations of objective factors and depend on many personal characteristics and experiences. Hence, their effect may play a much larger role than the objective environment. In an analysis of the relationship between the objectively measured and perceived built environments, Ma *et al.* (2014) show that the latter had a direct and significant effect on bicycling behavior, while the direct effect of the objective environment on bicycling behavior became insignificant when controlling for perception.

A key aspect of perceptions is the concept of risk. Fear of riding in motorized traffic could be felt as an obstacle to cycle (Rietveld and Daniel, 2004; Garrard *et al.*, 2008; Fernández-Heredia *et al.*, 2014). However, perception of risk is a subjective matter that is not always correlated with actual risk (Fernández-Heredia *et al.*, 2014).

2.2.6.2. Psychological indicators

The other subgroup of subjective variables includes those factors that are not explicitly identified by the users but are suspected to have an influence on the choice to cycle, such as personal attitudes, self-efficacy and social norms (Heinen *et al.*, 2010; Fernández-Heredia *et al.*, 2014; Muñoz *et al.*, 2016). Almost all studies using this psychometric approach conduct their analysis following the Theory of Planned Behavior (Ajzen, 1985), which assumes that attitudes, subjective norms and perceptions can affect actual behavior (Maldonado-Hinarejos *et al.* (2014), Lois *et al.* (2015), Fernández-Heredia *et al.* (2016), and Muñoz *et al.*, 2016). Others use the Transtheoretical Model of Behavior Change (Prochaska *et al.*, 1992) for their analysis (Gatersleben and Appleton, 2007; Thigpen *et al.*, 2015), focusing on how people's attitudes and perceptions vary for each stage of change.

Attitudes refer to the general evaluation individuals have about places, objects, or activities. It has been found that a positive attitude toward cycling increases the likelihood of using the bike (Dill and Voros, 2007; Sener *et al.*, 2009; Muñoz *et al.*, 2013; Fernández-Heredia *et al.*, 2014; Muñoz *et al.*, 2016; Gao *et al.*, 2019). Heinen *et al.* (2011) found that an increase in distance corresponds with a decrease in the average value of attitudes toward the various characteristics of bicycle travel.

Earlier research also confirmed the importance of social norms. Dill and Voros (2007) found that people living in households with other adults who cycled regularly, have co-workers who cycle to work, or people who see adults cycling on their street frequently are more likely to be regular cyclists themselves. Similarly, Muñoz *et al.* (2016) show that the influence of friends, family and co-workers/classmates also fosters the decision to cycle for commuting purposes.

Further, habit can shape bike behavior over attitudes. Stinson and Bhat (2004) stated that cycling more in leisure time could increase the frequency of bicycle use for commuting. In a study of cycling in six small U.S. cities, Xing *et al.* (2010) found that while over one-quarter of cyclists only cycled for leisure and sport, only 10% cycled only for transportation and the majority cycled for a mix of recreational and transportation purpose, suggesting a relationship between the two. Among a sample of Korean cyclists, Park *et al.* (2011) found that 57% of commuter-cyclists began as leisure-cyclists. A broader perspective has been adopted by Kroesen and Handy (2014) who argue that the effect of bicycle commuting on non-work cycling is greater than *vice versa.* However, this study uses a sample from a country, the Netherlands, with a high level of cycling.

2.2.7. Methodological approach

From a methodological point of view, previous research has used a variety of statistical and econometric models including descriptive analysis, discrete choice models and structural equation models to understand either bicycle mode choice or its frequency.

2.2.7.1. Aggregate level methods

Aggregate-level methods predict demand by mode for an area served by a facility, based on the characteristics of the area served, such as population and land use mix. Aggregate methods include measures of potential demand, comparison studies, aggregate behavior studies or a mixture of these (Maldonado *et al.*, 2014). In particular:

- Facility Demand Potential: these methods predict the maximum potential demand for bicycles by characterizing facilities and using population characteristics, proximity to activity centers, and trip distribution. Often, they are used to prioritize projects based on potential usage.
- Comparison Studies: These studies compare bicycle levels before and after a change in the travel context or an improvement in bike facilities. Counts on existing facilities also have been used to forecast demand for proposed facilities with similar characteristics.
- Aggregate Behavior Studies: these studies estimate models to predict mode split or other travel behavior characteristics at a zonal level, such as for residents of census sections or metropolitan areas. An example of methods is regression analysis.

2.2.7.2. Discrete choice models

Several studies have developed discrete choice models to investigate the relation between explanatory variables and the likelihood to cycle (see Table 3). Regarding data collection, some works used revealed preference (RP) data, others stated preference (SP) data. Most studies use specifications based on logit formulations (Muñoz *et al.*, 2016). Some works investigate the choice between cycling or not cycling, while others have focused on quantifying the factors influencing the choice between cycling and other modes. Finally, some research works analyze cycling frequency through the construction of ordered data models.

Table 3. Discrete choice models used for modeling bicycle choice

Authors	Year	Data collection	Methodology	Location
Taylor and Mahmassani	1996	SP	Nested logit	Texas
Stinson and Bhat	2004	RP	Ordered logit	Austin, TX
Ryley	2006	RP/SP	Multinomial logit	West Edinburgh, Scotland
Hunt and Abraham	2007	SP	Binary logit	Edmont, Canada
Wardman *et al.*	2007	RP/SP	Hierarchical logit	Different towns in UK
Parkin *et al.*	2008	RP	Binary logit model	UK
Akar and Clifton	2009	SP	Multinomial logit	University of Maryland
Sener *et al.*	2009	RP	Bivariate ordered probit	Austin, TX
Börjesson and Eliasson	2010	SP	Binary logit	Stockholm
Handy *et al.*	2010	RP	Nested Logit	USA
Yi *et al.*	2011	RP/SP	Hierarchical logit	Sidney
Akar *et al.*	2013	RP	Multinomial logit	Ohio State University
Heinen *et al.*	2013	RP	Binary Logit	Delft and Zwolle, Netherlands
Hamre and Buehler	2014	RP	Multinomial Logit	Washington DC
Braun *et al.*	2016	RP	Binary Logit	Barcelona
Bhat *et al.*	2017	RP	Spatial generalized ordered probit	USA
Ton *et al.*	2019	RP	Multinomial Logit	Netherlands
Gao *et al.*	2019	RP	Tobit	Netherlands

2.2.7.3. The incorporation of latent variables in bicycle mode choice studies

As explained in 2.2.6, a particularly appealing characteristic of the more recent commute bicycling-oriented research efforts is the explicit recognition that, in addition to objective factors (such as individual and household socio-demographics, variables characterizing the bicycling facilities environment, and trip characteristics), subjective factors linked to individuals' attitudes, perceptions, beliefs and social norms also play an important role in commute bicycling decisions. This recognition may be attributed to the general finding of the presence of substantial unobserved individual heterogeneity in the evaluation of bicycle infrastructure attributes, as well as the observation that individuals with similar socio-economic and bicycling environment characteristics tended to manifest quite different commute bicycling behavior. Some earlier studies in the past decade, such as Sener *et al.* (2009), considered such unobserved factors, but only implicitly, by allowing random distributions to capture sensitivity variations across individuals to route attributes (that is, taste heterogeneity). But this random distribution approach treats unobserved psychological preliminaries of choice (*i.e.*, attitudes and preferences) as being contained in a “black box” to be integrated out. On the other hand, with the development of ICLV models it is possible to unpack the unobserved heterogeneity and have deeper understanding of bicycling choice by considering “soft” (latent) psychometric measures of individual attitudes and perceptions.

The use of ICLV models is well established in this field. Kamargianni and Polydoropoulou (2013) estimate an ICLV model to study the impact of the latent variable willingness among teenagers to walk or cycle in mode choice. The same survey was also used by the study Kamargianni *et al.* (2014) to test Bhat and Dubey (2014) ICLV probit kernel model. Habib *et al.* (2014) present an integrated econometric model of bike ownership and choice of biking for utilitarian and/or recreational purposes using simultaneous estimators and three latent variables, namely the impact of comfort, safety awareness and perceptions of bikeability (in terms of quality of bike facilities). Maldonado-Hinarejos *et al.* (2014), using data from a stated preference experiment, develop an ICLV model using sequential estimation. They incorporate attitudes toward cycling, perceptions of the image associated with cycling and the stress arising from safety concerns, identified using a principal component analysis in the utility to commute by bike. Motoaki and Daziano (2015) estimate a hybrid choice latent class model to investigate the effects of weather (temperature, rain, and snow), cycling time, slope, cycle facilities (bike lanes), and traffic on cycling decisions. Fernández-Heredia *et al.* (2016) estimate a hybrid model to study the intention to bicycle combining a structural equations model that captures intentions and a choice model. Sottile *et al.* (2019) use a hybrid choice model to estimate the effect of people's perception on the propensity to bike, finding that, besides individual characteristics, latent aspects related to the perception of the context and of the bicycle as a means of transport, strongly affect the propensity to cycle.

2.2.7.4. Inference

There are a few studies that use forecasting techniques for bike share prevision. Wardman *et al.* (2007) evaluate the likelihood to commute by bike under a variety of scenarios, including modest financial incentives, cycle facilities for around half the journey to work and good parking and shower facilities at work in the British context.

Goetzke and Rave (2011) found that in a case study in Germany in a scenario with the introduction of new cycling infrastructure, bike share increases only for shopping and errand trips. Moreover, commuting trips by bicycle seemed to be largely independent of any policy variables.

Hamre and Buehler (2014) found that in the region of Washington DC in a scenario simply introducing bike/walk facilities (showers/lockers and/or bike parking) walk share for commuting increases from 1.4% to 2.1% and cycle share from 0.5% to 1.0%.

Maldonado-Hineros *et al.* (2014) use an Integrated Choice Latent Variable model, in the London (UK) context, to evaluate choice behavior in three different policy scenarios: (a) a scenario in which the image of cycling increases by one scale point for each related attitudinal and perceptual indicator (b) a scenario in which cycle parking facilities are improved for all travelers assumed to have access to locked cycle parking compounds (c) the combination of scenarios (a) and (b). The results show that under both the image enhancement and parking facility improvement scenarios the cycling mode share increases, and that in combination the increase in market share is mildly super-additive. Nevertheless, cycling is predicted to draw demand principally from public transport and walking rather than from the car.

2.3. DATA COLLECTION

The data used in this study originates from an on-line survey implemented by the Regional Government of Sardinia (RAS) and the Research Centre for Mobility Models (CRiMM) at the University of Cagliari, Italy, in the metropolitan areas of Cagliari and Sassari (Italy). Potential candidates were contacted both via mailing lists provided by the universities of Cagliari and Sassari and the Regional Government of Sardinia (around 9,600 invitation mails were sent) and through a promotional campaign conducted via traditional communication channels and social media.

A total of 4,691 individuals responded to the survey. However, following careful screening - excluding incompleteness or missing crucial information - the final sample size includes 2,128 individuals (corresponding to 45.4% of respondents).

The on-line questionnaire, built on a SaaS (*Software as a service*) platform called WUFOO (www.wufoo.com), comprised three sections:

1. *Bicycles and cycle paths*

The first section aimed to identify for what purpose people choose to cycle, as well as their impressions and opinions about cycling and cycle paths from cyclists and non-cyclists.

- Cyclists section: to be completed by those who stated they cycled, to identify what type of bike they used, for what purpose, distance traveled for each purpose, type of route cycled;
- Bicycles and cycle paths: to be completed by cyclists and non-cyclists. Respondents are asked to express their agreement/disagreement (on a Likert scale from 1 to 5) with regard to a series of statements concerning the bicycle and its use, existing cycle paths and factors that are likely to encourage people to cycle or to cycle more;
- Reasons for not cycling: to be completed by those who stated they did not cycle. Non cyclists are asked to indicate the importance (on a Likert scale from 1 to 5) of a series of factors that influence the decision not to cycle.

2. *Description of home-work trip*

The second section aimed to identify the form of transport used for the home-work trip, especially to obtain a detailed description for the car-as-driver, public transport and bicycle modes.

- Modes used: workplace address, means of transport used to reach workplace, departure time, means of transport not usually available for the home-work trip; depending on the trip mode reported (car-as-driver, public transport and bicycle) the respondent is directed to a specific section for describing the trip

- Car as driver: time taken to walk from home to where car is parked, in-vehicle travel time, time taken looking for a parking place at destination, type of parking and if paid parking, cost; any stops along the way;
- Public transport (alone or combined with other means of transport): time taken to walk to the bus stop/station, waiting time, in-vehicle travel time (and on board second mode if travelling by a combination of modes), time taken on alighting from bus/train to walk to final destination, type of ticket used;
- Cycling: duration of trip, type of parking, alternative means of travel in bad weather, mode used prior to taking up cycling.

3. *Socio-economic information*

Lastly, the third section contains questions of a socio-economic nature:

- Socio-economic information: address, age, gender, occupation, education, income, marital status, number and age of children living at home, number of household members, number of cars and bicycles in the household, possession of driving license;
- Information about height and weight: for the purpose of calculating body mass index and evaluating the correlation with active transportation (bicycle);
- Possession of a smartphone and type;
- E-mail address for participating in the prize draw.

2.3.1. Socio-economic characteristics

The sample is distributed throughout the island of Sardinia (70.1% in the metropolitan city of Cagliari and in the province of South Sardinia, 25.6% in the metropolitan area of Sassari and the remaining 4.3% divided between the other Sardinian provinces).

Regarding individual characteristics (Table 4), the sample is practically equally divided between females and males with a slight preponderance of the latter. In terms of age distribution, the sample was composed as follows, 18-30 (3.8%), 31-40 (16.6%), 41-60 (72.8%), over 60 (6.8%).

The level of education of the sample is quite high. This is not surprising as a large number of respondents who completed the questionnaire worked at Cagliari and Sassari universities, where it is more feasible to find individuals with a degree or a postgraduate diploma. In particular, 36.3% had a high school diploma, 34.7% had a bachelor's or master's degree and 23.0% pursued a PhD program.

The majority of respondents (93.4%) are employees, while 5.9% are students/graduate students. The remaining quota (0.8%) stated they were unemployed or retired. As for personal monthly income, 6.3% stated they earned less than € 1,000 a month, 66.2% € 1.000-2.000, 13.6% € 2000-4000, 13.8% > € 4.000. Analysis of marital status shows that the majority of individuals are married/living with partner in a household with an average of 3 members. Almost all have a driving license and own a car. The average number of cars per household is 2, slightly less for bicycles, though 79.6% reported owning at least one.

Another interesting aspect analyzed, that can be significant when considering active mobility *vs* non active mobility but often overlooked in most studies, was the body mass index (BMI) that was calculated based on the data on people's height and weight. The BMI is a biometric datum, based on the ratio of a person's weight to height squared and is used as an indicator of ideal weight. On average the sample had normal weight with an average BMI of 23.62 kg/m^2.

Table 4. Data description

Variables	N.	[%]	AVG.
Total sample	2,128		
Gender (male)	1029	48.4%	
Age		-	48.02
Age 18-30	82	3.9%	
Age 31-40	341	16.0%	
Age 41-60	1559	73.3%	
Age > 60	146	6.9%	
Level of education			
Low (High school and lower)	901	42.3%	
Medium (Graduate)	738	34.7%	
High (Higher than Master's degree)	489	23.0%	
Employment status			
Student	125	5.9%	
Worker	1987	93.4%	
Unemployed or retired	16	0.8%	
Marital status: married	1550	72.8%	
With children	1159	54.5%	
# of members in the household		-	2.88
Driving license	2098	98.6%	
Personal car available	1930	90.7%	
# of cars per household	-	-	1.72
# of bikes per household	-	-	1.54
Personal income per month			
Income 0-1000 €	140	6.6%	-
Income 1000-2000 €	1382	64.9%	-
Income 2000-3000 €	205	9.6%	-
Income >3000 €	301	14.1%	-

2.3.2. Transport analysis

As was to be expected from the above results on car ownership, practically the entire sample (94.0%) had access to a car for the home-work commute, 82.4% could travel to the workplace by public transport, 51.9% by bicycle and lastly 43.0% on foot (Table 5). In spite of the fact that almost half the sample could active commute to work either on foot or by bike or both, the private car dominates the modal split, in keeping with the Sardinian population as a whole. In fact, 67.5% of the respondents choose to commute by car (67.5%), followed by walking (22.6%), transit (9.9%) and cycling (7.9%).

Table 5. Transport analysis

	Observed mode share [%]	**Availability [%]**
Private car	67.5%	94.0%
Public transport	9.9%	82.4%
Walking	14.7%	43.0%
Cycling	7.9%	51.9%

2.3.3. Analysis of bicycle use

Analysis of the questionnaires revealed that around 50% of the sample (1063 individuals) stated they used the bike for some purpose or another (

Figure 3). Further analysis revealed the following share by frequency:

- I never use the bike (50.0% of the sample);
- I use the bike 1-10 times per year (14.6% of the sample);
- I use the bike 1-5 times per month (14.2% of the sample);
- I use the bike more than once a week (14.7% of the sample);
- I use the bike everyday (6.5% of the sample).

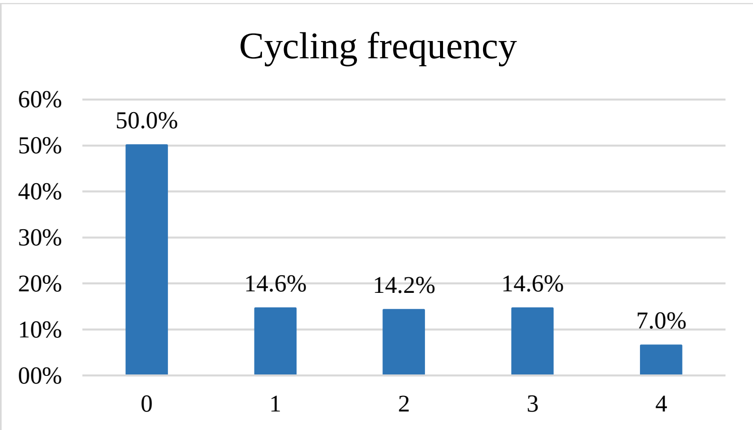


Figure 3. Results on the share of people using a bike by frequency (frequency scale 0=never; 1=1-10 times per year; 2=1-5 times per month; 3=more than once a week; 4=daily)

Table 6 shows the results of the share of people using a bicycle by frequency and by purpose. As can be clearly seen, all individuals who choose to cycle use the bike for leisure and sport. Instead, 55.8% of the sample reported they never used the bike for shopping or for commuting (63.6%). This finding is in line with other countries where cycling share is low, and individuals perceive the bicycle as a form of recreation more than as a means of transport.

Table 6. Analysis of cycling frequency by purpose for those who decide to cycle.

Cycling frequency	**Choice = Use the bike (1063 individuals)**		
	Leisure and sport	*Shopping*	*Work*
Never	0 (0.0%)	593 (55.8%)	676 (63.6%)
1-10 times per year	366 (34.4%)	236 (22.2%)	127 (11.9%)
1-5 times per month	342 (32.1%)	146 (13.7%)	65 (6.1%)
More than once a week	297 (27.9%)	211 (10.8%)	98 (9.2%)
Daily	58 (5.4%)	0 (0.0%)	97 (9.1%)

2.3.3.1. Comparisons: socio-economic characteristics for different levels of cycling

Table 7 gives the socio-demographic variables of respondents for different levels of cycling. No significant differences were detected between bike users and non-users. The most interesting difference concerns bicycle ownership per household. Indeed, bike users tend to own more bikes in their household than individuals who do not choose to cycle.

Table 7. Socio-economic characteristics for different levels of cycling

	Frequency of bike use									
	Never		*1-10 times per year*		*1-5 times per month*		*More than once a week*		*Every day*	
	N.	*%*	*N.*	*%*	*N.*	*%*	*N.*	*%*	*N.*	*%*
Questionnaires completed	1065	50.1%	310	14.6%	303	14.2%	311	14.6%	139	6.5%
Gender (male)	423	39.7%	128	41.3%	159	52.5%	212	68.2%	107	77.0%
Age (average)	49.07	-	46.13	-	46.49	-	47.46	-	48.82	-
18-30	24	2.3%	20	6.5%	14	4.6%	19	6.1%	5	3.6%
31-40	164	15.4%	59	19.0%	53	17.5%	44	14.1%	21	15.1%
41-60	786	73.8%	215	69.4%	223	73.6%	232	74.6%	103	74.1%
>60	91	8.5%	16	5.2%	13	4.3%	16	5.1%	10	7.2%
Level of education										
Low (High school and lower)	451	42.3%	107	34.5%	94	31.0%	172	55.3%	77	55.4%
Medium (Graduate)	381	35.8%	110	35.5%	125	41.3%	85	27.3%	37	26.6%
High (Higher than Master's degree)	233	21.9%	93	30.0%	84	27.7%	54	17.4%	25	18.0%
Employment status										
Student	46	4.3%	30	9.7%	21	6.9%	22	7.1%	6	4.3%
Employed	1011	94.9%	277	89.4%	281	92.7%	287	92.3%	131	94.2%
Unemployed or retired	8	0.8%	3	1.0%	1	0.3%	2	0.6%	2	1.4%
With Children	581	54.6%	170	54.8%	163	53.8%	171	55.0%	74	53.2%
Married	757	71.1%	230	74.2%	224	73.9%	236	75.9%	103	74.1%
# members of household (average)	2.83	-	2.99	-	2.88	-	2.96	-	2.81	-
Driving License	1049	98.5%	308	99.4%	300	99.0%	305	98.1%	136	97.8%
Own car	978	91.8%	276	89.0%	274	90.4%	275	88.4%	127	91.4%
# of cars (average)	1.71	-	1.82	-	1.72	-	1.73	-	1.52	-
# of bikes (average)	0.97	-	1.98	-	2.07	-	2.19	-	2.32	-
Monthly personal income										
Income 0-1000 €	52	4.9%	37	11.9%	16	5.3%	22	7.1%	13	9.4%
Income 1000-2000 €	706	66.3%	183	59.0%	205	67.7%	206	66.2%	82	59.0%
Income 2000-3000 €	148	13.9%	46	14.8%	41	13.5%	50	16.1%	20	14.4%
Income >3000 €	159	14.9%	44	14.2%	41	13.5%	33	10.6%	24	17.3%

2.3.4. Analysis of built environment characteristics

A crucial aspect in the choice to use the bike concerns the built environment that in many cases may represent a barrier to the choice to use the bike. Table 8 provides a summary of participants' built environment characteristics. The micro-environments of the home address area (presence of bike lanes and percentage of green areas) were assessed within the GIS environment using a buffer of 400 m radius. Using the digital land use maps downloaded from the Sardinian Government website, it was possible to calculate the characteristics of the residence location (urban or suburban).

It is apparent from Table 8 that no significant differences were found among cycling frequency categories. The majority of individuals live in urban areas and fewer than half have access to a bike lane within 400m of home.

Table 8. Built environment characteristics

Variables	Cycling frequency				
	Never	1-10 times per year	1-5 times per month	More than once a week	Everyday
% of individuals living in urban areas	79.1%	72.2%	74.3%	79.6%	83.0%
% of individuals who have access to a bike lane within 400m of home	48.7%	48.7%	45.1%	47.1%	59.2%
Average % of green areas within 400m of home per individual	5.2%	5.3%	5.1%	4.9%	5.1%

2.3.5. Analysis of psycho-attitudinal characteristics

Psycho-attitudinal characteristics were measured by means of the questions of the 5-point Likert scale (1 = Totally disagree to 5 = Totally agree). Specifically, the questionnaire contained three questions for measuring:

- The perception of the bicycle as a means of transport at the personal (in terms of travel time, travel cost, comfort, health, *etc.*) and societal level (environment, quality of life, standard of living).
- The perception of bikeability in terms of usefulness and safety.
- Facilitators of cycling with reference to the context characteristics (*e.g.* existence of bike lanes in urban area, presence of racks and secure parking, integration with the public transport service).

Overall most respondents recognized the advantages associated with cycling in terms of cost, travel times, benefits for health and the environment, as can be seen from the average values close to 5 for items A1, A4, A6, A7, A9 and A11.

Table 9. Analysis of the responses regarding perception of the bicycle as a means of transport at the personal

FIRST QUESTION "The perception of the bicycle as a means of transport at the personal and societal level"	1=Totally disagree		2=Disagree		3=Neutral		4=Agree		5=Totally agree		AVG
	N.	%	N.	%	N.	%	N.	%	N.	%	
A1. It is a rapid means of transport (it avoids queues and traffic)	60	2.8%	138	6.5%	402	18.9%	411	19.3%	1117	52.5%	4.12
A2. Cycling in traffic is not dangerous	1184	55.6%	497	23.4%	295	13.9%	91	4.3%	61	2.9%	1.75
A3. It is not likely to be stolen and there are no adequate parking areas	863	40.6%	463	21.8%	403	18.9%	263	12.4%	136	6.4%	2.22
A4. It is not expensive	36	1.7%	30	1.4%	99	4.7%	225	10.6%	1738	81.7%	4.69
A5. It does not imply exposure to bad weather and air pollution	605	28.4%	518	24.3%	601	28.2%	279	13.1%	125	5.9%	2.44
A6. It avoids wasting time looking for parking	68	3.2%	78	3.7%	226	10.6%	359	16.9%	1397	65.6%	4.38
A7. It is healthy	25	1.2%	18	0.8%	121	5.7%	274	12.9%	1690	79.4%	4.69
A8. It is easy to carry heavy items	1102	51.8%	476	22.4%	340	16.0%	127	6.0%	83	3.9%	1.88
A9. It allows one to appreciate historic centers and increases accessibility to city services	51	2.4%	83	3.9%	238	11.2%	430	20.2%	1326	62.3%	4.36
A10. No need for cycling gear	279	13.1%	484	22.7%	713	33.5%	355	16.7%	297	14.0%	2.96
A11. It contributes to reducing polluting emissions	23	1.1%	18	0.8%	36	1.7%	127	6.0%	1924	90.4%	4.84
A12. It does not hamper daily activity patterns	244	11.5%	453	21.3%	683	32.1%	373	17.5%	375	17.6%	3.09

The second question concerned the perception of the context in terms of bikeability and issues with existing cycle paths. Most individuals agree that existing bike lanes are not useful and not safe. This result is certainly due to the lack of a protected and connected system of bike lanes in the reference context. Moreover, the majority of respondents prefer to use the existing bike lanes than ride in traffic, even if they consider them not secure.

Table 10. Analysis of the responses regarding the perception of bikeability in terms of usefulness and safety

SECOND QUESTION "The perception of bikeability in terms of usefulness and safety"	1=Totally disagree		2=Disagree		3=Neutral		4=Agree		5=Totally agree		AVG
	N.	%	N.	%	N.	%	N.	%	N.	%	
B1. Existing bike lanes are not useful for travelling	302	14.2%	244	11.5%	529	24.9%	387	18.2%	666	31.3%	3.41
B2. Existing bike lanes and crossings are safe, comfortable and well-marked	791	37.2%	608	28.6%	510	24.0%	146	6.9%	73	3.4%	2.11
B3. It is better to ride in traffic than use the existing bike paths	1136	53.4%	363	17.1%	370	17.4%	128	6.0%	131	6.2%	1.95
B4. Car drivers do not respect dedicated bike lanes and often invade them	103	4.8%	165	7.8%	373	17.5%	481	22.6%	1006	47.3%	4.00

Examination of the facilitating factors revealed that the creation of a network of dedicated bike lanes and the possibility of safe bike parking were the most important factors in encouraging individuals to cycle or cycle more. Another decisive element appears to be the possibility of combining cycling with public transport, suggesting that people could use the bike as a means of transport for intermodal trips.

Table 11. Analysis of the responses regarding the facilitators of cycling with reference to the context characteristics

THIRD QUESTION: "The perceived importance of context characteristics"	1=Totally disagree		2=Disagree		3=Neutral		4=Agree		5=Totally agree		AVG
	N.	%	N.	%	N.	%	N.	%	N.	%	
C1. An extensive network of dedicated bike lanes in urban area	63	3.0%	75	3.5%	206	9.7%	402	18.9%	1382	64.9%	4.39
C2. The presence of racks and secure parking for bicycles	63	3.0%	104	4.9%	250	11.7%	504	23.7%	1207	56.7%	4.26
C3. A greater extension of the LTZ or pedestrian zones	168	7.9%	192	9.0%	446	21.0%	471	22.1%	851	40.0%	3.77
C4. A bike-sharing station close to home or at public transport stops	180	8.5%	239	11.2%	403	18.9%	467	21.9%	839	39.4%	3.73
C5. If other people use it	473	22.2%	373	17.5%	485	22.8%	371	17.4%	426	20.0%	2.95
C6. Dedicated services at work / study (parking, showers, lockers for equipment, etc.)	130	6.1%	156	7.3%	365	17.2%	520	24.4%	957	45.0%	3.95
C7. An integrated ticket for bike-sharing and public transport services	153	7.2%	220	10.3%	353	16.6%	485	22.8%	917	43.1%	3.84
C8. Combination with public transport services	126	5.9%	176	8.3%	328	15.4%	508	23.9%	990	46.5%	3.97
C9. Increase of car parking fee	910	42.8%	380	17.9%	428	20.1%	178	8.4%	232	10.9%	2.27

2.3.5.1. Comparisons: psycho-attitudinal characteristics for different levels of cycling

Table 12 shows the results of the statistical test analysis for individuals with different levels of cycling experience with regard to the attitudinal factors. There are items that are not statistically significantly different among categories. For example, no differences were detected for the item regarding the benefits of bicycling in terms of cost, accessibility and reduced level of pollution. Interestingly, there were no significant differences between the groups for those aspects regarding the implementation of certain measures, such as a greater extension of limited traffic zones and presence of end-of-trip facilities. This latter result suggests that all the individuals consider the existence of this type of facilities important, regardless of their level of cycling experience.

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

Table 12. Psycho-attitudinal characteristics for different levels of cycling (0 = never, 1=1-10 times per year, 2=1-5 times per month, 3= more than once a week, 4 = everyday). * Significant at 90% confidence. ** Significant at 95% confidence.

	Avg 0	Avg 1	Avg 2	Avg 3	Avg 4	**Z-stat 0-1**	**Z-stat 1-2**	**Z-stat 2-3**	**Z-stat 3-4**
A1. It is a rapid means of transport	3.99	3.90	4.15	4.46	4.83	1.12	-2.79**	-3.72**	-5.31**
A2. Cycling in traffic is not dangerous	1.72	1.56	1.75	1.86	2.22	2.47**	-2.38**	-1.32	-2.98**
A3. It is not likely to be stolen and there are no adequate parking areas	2.32	2.25	2.12	2.03	2.11	0.80	1.35	0.86	-0.60
A4. It is not expensive	4.66	4.65	4.72	4.73	4.84	0.16	-1.01	-0.26	-1.78*
A5. It does not imply exposure to bad weather and air pollution	2.24	2.49	2.65	2.69	2.78	-3.31**	-1.79*	-0.41	-0.64
A6. It avoids wasting time looking for parking	4.30	4.24	4.46	4.55	4.78	0.79	-2.61**	-1.24	-2.70**
A7. It is healthy	4.57	4.67	4.85	4.85	4.86	-1.93*	-3.53**	-0.04	-0.16
A8. It is easy to carry heavy items	1.70	1.83	2.01	2.18	2.41	-1.91*	-1.99**	-1.81*	-1.91*
A9. It allows one to appreciate historic centers and increases accessibility to city services	4.31	4.25	4.38	4.50	4.63	0.96	-1.54	-1.59	-1.50
A10. No need for cycling gear	2.83	2.83	2.95	3.18	3.70	0.10	-1.35	-2.36**	-3.86**
A11. It contributes to reducing polluting emissions	4.83	4.79	4.86	4.86	4.93	0.78	-1.31	0.04	-1.78*
A12. It does not hamper daily activity patterns	2.89	3.08	3.08	3.39	3.91	-2.35**	-0.05	-3.18**	-4.24**
B1. Existing bike lanes are not useful for travelling	3.57	3.42	3.22	3.14	3.16	1.73*	1.77*	0.73	-0.10
B2. Existing bike lanes and crossings are safe, comfortable and well-marked	2.02	2.16	2.21	2.22	2.16	-1.98**	-0.55	-0.07	0.47
B3. It is better to ride in traffic than use the existing bike paths	1.97	1.86	1.81	1.98	2.17	1.29	0.56	-1.76*	-1.34
B4. Car drivers do not respect dedicated bike lanes and often invade them	3.88	3.93	4.13	4.19	4.30	-0.63	-2.08**	-0.78	-0.95
C1. An extensive network of dedicated bike lanes in urban area	4.23	4.44	4.61	4.61	4.61	-3.19**	-2.51**	0.06	0.01
C2. The presence of racks and secure parking for bicycles	4.15	4.23	4.42	4.43	4.46	-1.11	-2.49**	-0.16	-0.26
C3. A greater extension of the LTZ or pedestrian zones	3.67	3.78	3.81	3.94	4.12	-1.28	-0.28	-1.36	-1.50
C4. A bike-sharing station close to home or at public transport stops	3.87	3.72	3.53	3.57	3.42	1.71*	1.71*	-0.30	1.01
C5. If other people use it	2.99	2.83	2.75	3.08	3.13	1.79*	0.69	-2.81**	-0.29
C6. Dedicated services at work / study (parking, showers, lockers for equipment, etc.)	3.86	3.93	4.04	4.17	4.01	-0.91	-1.25	-1.32	1.19
C7. An integrated ticket for bike-sharing and public transport services	3.92	3.57	3.81	3.90	3.78	4.21**	-2.22**	-0.85	0.90
C8. Combination with public transport services	3.98	3.72	3.97	4.15	4.04	3.19**	-2.50**	-1.76*	0.88
C9. Increase of car parking fee	2.24	2.20	2.16	2.40	2.53	0.50	0.36	-2.17**	-0.87

2.3.6. Comparison: who chooses to cycle to work *vs* those who do not

Only 1,105 individuals of the 2,128 examined, *i.e.* 52%, reported having access to a bicycle for the home-work trip. For these 1,105 commuters, we examined the stated availability, checking, especially for those who did not choose that alternative, the feasibility of bike commuting in terms of distance traveled, route (urban or non-urban), access to at least one bicycle in the household. In no cases was cycling not a feasible option, the distance to be traveled by bike being at the most 25 km. 64% of the routes are on urban roads while the suburban routes are on minor roads and thus suitable for cycling, and all had access to at least one bicycle in the household. Thus, the availability stated in the questionnaire was considered correct.

Out of the subsample of 1,105 individuals who could commute by bicycle, 15.2% (168 individuals)[2], actually chose to do so, while 85% (937 individuals) travel to the workplace by other means of transport.

2.3.6.1. Socio-economic characteristics

No significant differences were observed between bike commuters and commuters who choose other forms of transportation in terms of socio-economic characteristics, as the sample was fairly homogeneous (public sector employees). The most interesting difference concerns car and bicycle ownership per household. In fact, bike commuters tend to own more bicycles and fewer cars in their household than individuals who commute by other means of transport (Table 13).

[2] The percentage of individuals who commute by bike is different than that shown in paragraph 2.3.2 since only habitual bicycle commuters were considered here, namely those who had a ratio $\frac{bicycle\ commute\ frequency}{commute\ frequency} \geq 0.8$.

Table 13. Comparison: who chooses to cycle to work *vs* those who do not. Socio-economic characteristics

	Total		*Choice = commute by bicycle*		*Choice ≠ commute by bicycle*	
	N.	*%*	*N.*	*%*	*N.*	*%*
Sample	1,105		168	15.2%	937	84.8%
Gender (male)	602	54.5%	130	77.4%	472	50.4%
Age (average)	47.92		47.36		48.02	
18-30	47	4.3%	8	4.8%	39	4.2%
31-40	165	14.9%	31	18.5%	134	14.3%
41-60	822	74.4%	121	72.0%	701	74.8%
>60	71	6.4%	8	4.8%	63	6.7%
Level of education						
Low (High school and lower)	470	42.5%	78	46.4%	392	41.8%
Medium (Graduate)	386	34.9%	53	31.5%	333	35.5%
High (Higher than Master's degree)	249	22.5%	37	22.0%	212	22.6%
Employment status						
Student	65	5.9%	7	4.2%	58	6.2%
Employed	1033	93.5%	158	94.0%	875	93.4%
Unemployed or retired	7	0.6%	3	1.8%	4	0.4%
With Children	622	56.3%	90	53.6%	532	56.8%
Married	821	74.3%	128	76.2%	693	74.0%
# members of household (average)	2.92		2.95		2.91	
Driving License	1085	98.2%	163	97.0%	922	98.4%
Own car	997	90.2%	144	85.7%	853	91.0%
# of cars (average)	1.74		1.59		1.77	
# of bikes (average)	2.01		2.32		1.95	
Monthly personal income						
Income 0-1000 €	78	7.1%	14	8.3%	64	6.8%
Income 1000-2000 €	702	63.5%	104	61.9%	598	63.8%
Income 2000-3000 €	168	15.2%	27	16.1%	141	15.0%
Income >3000 €	157	14.2%	23	13.7%	134	14.3%

2.3.6.2. *Analysis of bicycle use for non-commuting purposes*

With respect to the frequency of bicycle use for non-commuting trips (Table 14), it appears that the vast majority of respondents use the bicycle for leisure and sport (74.9%). Interestingly, a not negligible percentage of individuals (26.5%) reported to bicycle for leisure and sport more than once a week. It is found that 56.0% never used the bicycle for shopping. Only about 15.9% reported regular usage (one day per week or more) of the bicycle for shopping.

Table 14. Analysis of bicycle use

Level of frequency	*Leisure and sport*		*Shopping*	
	N.	***%***	***N.***	***%***
Never	277	25.1%	619	56.0%
1-10 times per year	262	23.7%	190	17.2%
1-5 times per month	273	24.7%	120	10.9%
More than once a week	243	22.0%	176	15.9%
Daily	50	4.5%	0	0.0%

2.3.6.3. *Psycho-attitudinal characteristics*

For each item, Table 15, Table 16 and Table 17 provide the mean calculated for the bike commuters (168 individuals), and separately the mean for the non-active commuters who had access to a bicycle (937 individuals) and the difference and t-stat between the means (mean bike commuters – mean commuters using other forms of transport).

The statements that warrant attention are those with the significant difference between the means obtained for the two groups. Compared to bike commuters, commuters who choose not to travel by bike, in order of importance: 1) are more likely to agree that cyclists need proper cycling gear and 2) that going by bike hampers daily activity patterns; are more likely to disagree 3) that it is a rapid means of transport 4) that it is not dangerous to cycle in traffic, 5) that it saves time looking for a parking place and 6) agree more that cyclists are exposed to bad weather and air pollution. All respondents recognized the advantages associated with cycling in terms of cost, travel times, benefits for health and the environment, as can be seen from the average values approaching 5 for items 1, 4, 6, 7, 8 and 11.

These differences, as reported in the literature, are also associated with the experience/inexperience of the two groups, who thus perceive cycling differently.

Table 15. Differences in perception of the bicycle as a means of transport

FIRST QUESTION: "The perception of the bicycle as a means of transport at the personal and societal level"	**Total**		**Choice = commute by bicycle**		**Choice ≠ commute by bicycle**		**Diff**	**Z-stat**
	Mean	*St dev*	*Mean*	*St dev*	*Mean*	*St dev*		
A1. It is a rapid means of transport (it avoids queues and traffic)	4.22	1.18	4.74	0.46	4.12	1.27	0.62	9.75
A2. Cycling in traffic is not dangerous	1.84	1.24	2.23	1.47	1.78	1.17	0.45	4.50
A3. It is not likely to be stolen and there are no adequate parking areas	2.16	1.65	2.05	1.52	2.18	1.67	-0.14	-1.32
A4. It is not expensive	4.73	0.55	4.75	0.48	4.71	0.58	0.04	0.63
A5. It does not imply exposure to bad weather and air pollution	2.53	1.46	2.82	1.43	2.47	1.45	0.35	3.49
A6. It avoids wasting time looking for parking	4.43	1.04	4.79	0.42	4.36	1.13	0.42	6.96
A7. It is healthy	4.76	0.41	4.83	0.30	4.74	0.45	0.09	1.91
A8. It is easy to carry heavy items	2.01	1.37	2.23	1.24	1.97	1.38	0.26	2.72
A9. It allows one to appreciate historic centers and increases accessibility to city services	4.36	1.09	4.60	0.78	4.31	1.15	0.29	3.73
A10. No Need for cycling gear	3.08	1.53	3.80	1.51	2.94	1.41	0.86	8.41
A11. It contributes to reducing polluting emissions	4.85	0.34	4.93	0.09	4.83	0.39	0.10	3.28
A12. It does not hamper daily activity patterns	3.22	1.55	3.90	1.33	3.09	1.49	0.81	8.35

The second question concerned the perception of the context at hand in terms of bikeability. Generally speaking, compared to the first question, the differences between the two groups are less marked, and they both agree that existing bike lanes are not useful and not safe. This result is certainly due to the lack of bike lanes in the reference context, which respondents are well aware of. Commuters who traveled using other means of transport, despite agreeing with the fact that car drivers do not respect dedicated bike lanes, ranked this item lower than bike commuters. They are also less likely to agree that it is better to ride in traffic than in existing bike lanes.

Table 16. Differences in perception of the bikeability

SECOND QUESTION: "The perception of bikeability in terms of usefulness and safety"	**Total**		**Choice = commute by bicycle**		**Choice ≠ commute by bicycle**		**Diff**	**Z-stat**
	Mean	*St dev*	*Mean*	*St dev*	*Mean*	*St dev*		
B1. Existing bike lanes are not useful for travelling	3.31	2.00	3.18	2.24	3.34	1.96	-0.16	-1.32
B2. Existing bike lanes and crossings are safe, comfortable and well-marked	2.14	1.21	2.20	1.40	2.13	1.19	0.07	0.70
B3. It is better to ride in traffic than use the existing bike paths	2.01	1.61	2.18	1.77	1.97	1.58	0.21	1.91
B4. Car drivers do not respect dedicated bike lanes and often invade them	4.09	1.34	4.30	1.00	4.04	1.40	0.26	3.04

The third question aimed to understand those factors that could encourage the use/greater use of the bicycle. Here, compared to cyclists, non-cyclists would have a greater incentive to cycle if 1) restricted traffic zones were extended, 2) other people cycled, hence social norms come into play and 3) there was a bike-sharing station near home or at bus stops/train stations.

Table 17. Differences in the perceived importance of context characteristics

THIRD QUESTION: "The perceived importance of context characteristics"	**Total**		**Choice = commute by bicycle**		**Choice ≠ commute by bicycle**		**Diff**	**Z-stat**
	Mean	*St dev*	*Mean*	*St dev*	*Mean*	*St dev*		
C1. An extensive network of dedicated bike lanes in urban area	4.43	0.98	4.58	0.84	4.40	1.01	0.18	2.32
C2. The presence of racks and secure parking for bicycles	4.28	1.08	4.38	1.09	4.25	1.08	0.13	1.45
C3. A greater extension of the LTZ or pedestrian zones	3.84	1.54	4.14	1.32	3.78	1.57	0.36	3.65
C4. A bike-sharing station close to home or at public transport stops	3.54	1.88	3.30	2.02	3.59	1.84	-0.29	-2.49
C5. If other people use it	2.92	2.15	3.22	2.35	2.87	2.07	0.35	2.75
C6. Dedicated services at work / study (parking, showers, lockers for equipment, etc.)	3.99	1.46	3.89	1.77	4.01	1.40	-0.12	-1.10
C7. An integrated ticket for bike-sharing and public transport services	3.75	1.77	3.67	1.84	3.77	1.75	-0.10	-0.90
C8. Combination with public transport services	3.93	1.56	3.95	1.46	3.93	1.58	0.02	0.24
C9. Increase of car parking fee	2.29	1.94	2.42	2.10	2.26	1.90	0.16	1.31

2.3.6.4. Trip characteristics

Another aspect we examined in detail concerns the trip characteristics that in certain cases may be a deterrent to cycling (Table 18). Note that for the 937 respondents who choose not to cycle to work the private car predominates in modal split (72%) and 20% of drivers make one stop on the way.

The distance and times by bike were simulated for the non-bike commuters for each O/D pair by means of an algorithm that uses Google Maps APIs, assuming the same route traveled by car, given the almost total lack of bike lanes. One important factor to consider is that the average distance traveled by bike commuters is around 4 km compared to the 6.7 km that non-bike commuters would have to travel. Clearly, bike commuters also spend less time traveling: 18 minutes compared to 33 minutes for those who do not commute by bike.

No differences were observed for departure time, as 78% of both groups left home during the morning peak hours between 7:30 and 9:30 am. 83% of trips by bike commuters are on urban roads within the same municipality of residence, while this percentage decreases to around 64% for non-bike commuters.

Table 18. Differences in trip characteristics

	Distance bike [km]	Time bike [min]	HP	Urban
Choice = commute by bicycle	3,9	18	77,4%	83,3%
Choice ≠ commute by bicycle	6,7	33	78,9%	63,6%

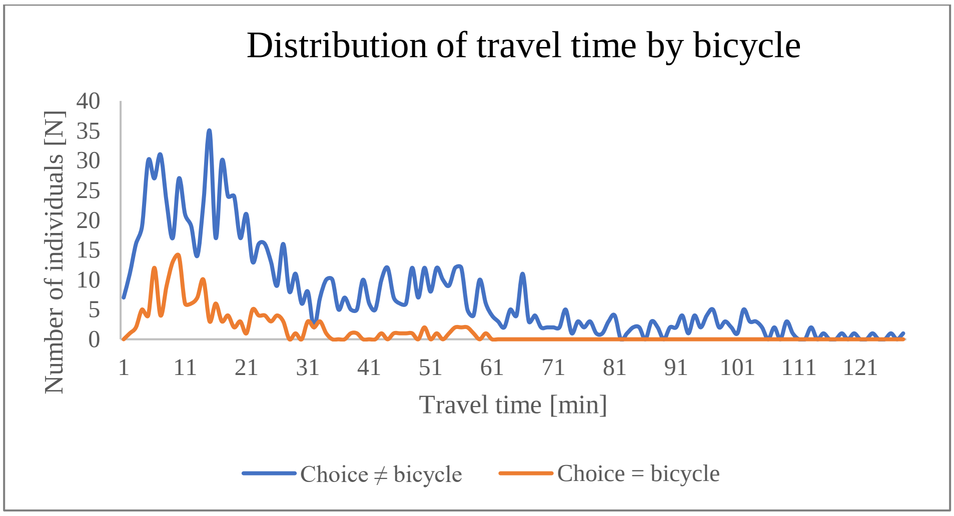


Figure 4. Distribution of travel time by bicycle

One important aspect for bike-commuting is certainly topography. Most of the trips examined here are made within the Cagliari municipality which lies mostly on hilly terrain. For this reason, we calculated the slopes along the routes travelled, also because this feature is often overlooked in studies on this topic.

Using the 10 m Digital Elevation Model (DEM) downloaded from the website of the Sardinian regional government, we extracted the routes with Google Maps and calculated slopes using GIS software. The results show similar average slopes for the two groups of commuters. Average down- and uphill slope was calculated to be around 2%. Figure 5 shows the percentage of routes identified for each class of slope.

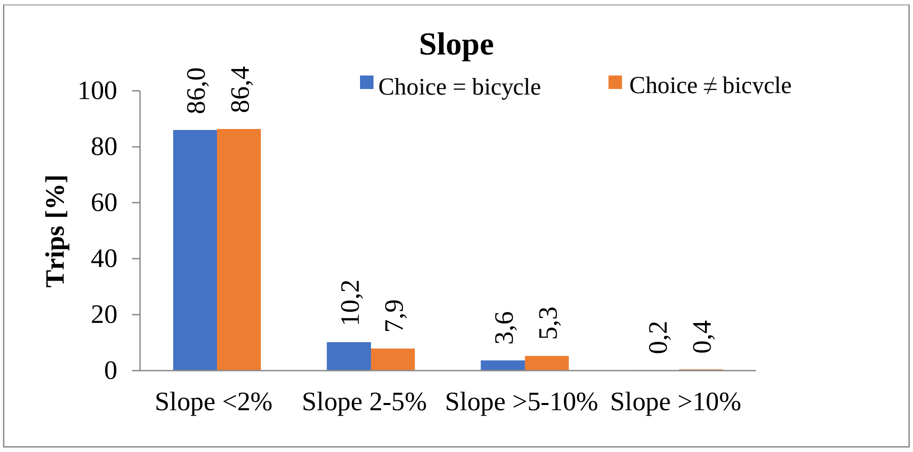


Figure 5. Percentage of trips for each slope class

For non-cyclists there is a slight prevalence over cyclists of routes that fall into the slope classes between 5 and 10% and >10%, while for both groups the majority of routes (around 86%) have slopes of less than 2%.

However, mean slope is calculated as the mean slope of each stretch of the route. Consequently, the existence of just one 10% uphill slope along the route, which could deter commuters from cycling to work, also considering the particularly hot summer temperatures in Sardinia, if the remainder of the route is downhill, mean slope along the entire route will be low. In light of the above we preferred to calculate the mean of maximum and minimum slopes, that can represent the real barrier (note that we only considered the home-work trip, but on the return journey an uphill slope becomes a downhill slope and *viceversa*). In fact, these values are much higher, much more so for non-cyclists (Table 19).

Table 19. Average value of slopes

	Uphill slope [%]	Downhill slope [%]	Max Slope [%]	Min Slope [%]
Choice = commute by bicycle	2.1	-2.2	4.6	-5.1
Choice ≠ commute by bicycle	2.4	-2.5	5.6	-6.3

2.4. THE ROLE OF INERTIA IN THE CHOICE OF COMMUTING TO WORK BY BIKE

2.4.1. Introduction

As seen in 2.2, several studies have investigated the explanatory variables, objective and subjective, that affect the propensity to cycle to work. All this research has emphasized the deliberate nature of individual decisions, assuming that individuals choose to cycle to work only reasoning on the good and bad consequences of their behavior. However, individuals may not go through such a conscious process when behavior is performed repeatedly and has become habitual (Aarts *et al.*,1997).

In the psychological literature habit is usually measured as the number of times the same trip is made using the same mode, but adopting such a methodology the transport choice would not be related to level of service characteristics. In fact, until behavior becomes habitual, individuals still look at the characteristics of the alternatives and adopt compensatory rules as well (Aarts *et al.*, 1998). Then it is crucial to account for the effect of both possible habitual behavior, leading to inertia, and trade-off among objective characteristics. Moreover, while frequency of usage (*i.e.* a repeated behavior) is probably the best indicator of habit, this is only an indicator of the tendency to repeat the same course of action.

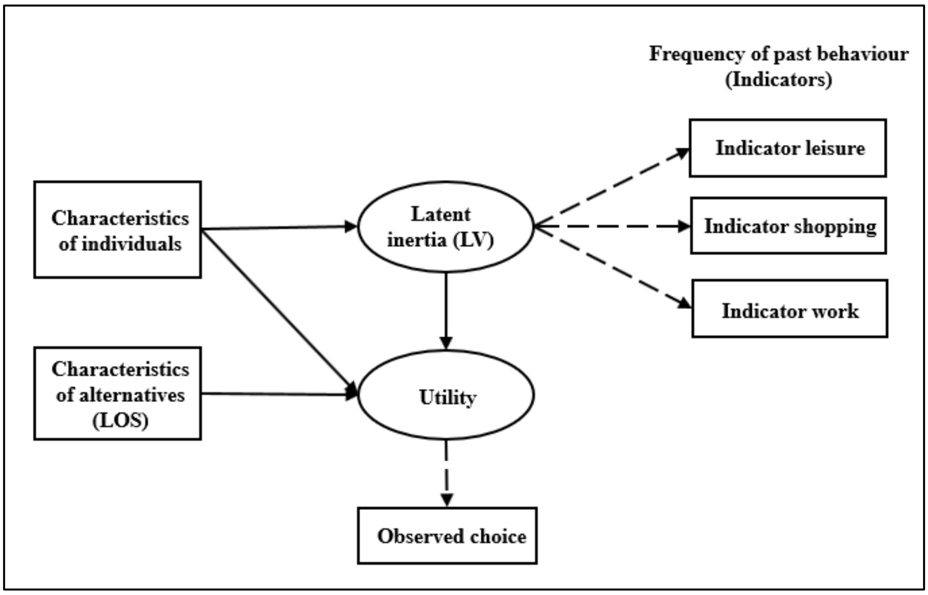
The aim of this section is to study specifically the impact of habitual cycling behavior on the choice to commute by bike. To do so, we estimate a mode choice model where cycling is one of four alternatives modes and we use a rich specification for all the modes to understand the relation between habit and compensatory evaluation of objective characteristics of the alternative modes. Following Cherchi *et al.* (2013) habitual behavior is measured as a latent variable, where three indicators reveal past behavior related to the use of the bike for different purposes. The discrete choice part of the model includes disaggregate level of service attributes for all modes, systematic heterogeneity, non-linear effects as well as other land use attributes such as topography.

The data used in this study originates from the survey “Bike I like you” conducted in 2014-2016 (see 2.2). The sample was composed of 2128 individuals who could choose at least two of four modes of travel (to ensure that their choice was actually a discrete choice and not mandatory) and the commute trip was made at least once a week.

2.4.2. Model specification

The methodology used to analyze the effect of habitual behavior is an Integrated Choice Latent Variable model, where it is assumed that the habitual behavior is latent and is revealed by the current frequency of bicycle usage. A graphical framework is shown in Figure 6.

Figure 6. Methodology framework



Let U_{jq} be the utility that each individual q associates to the alternative j. Under the assumption that latent habitual behavior affects the current choice, the utility can be written as:

$$U_{jq} = ASC_j + \boldsymbol{\beta}_{j,LOS}\boldsymbol{LOS}_{jq} + \boldsymbol{\beta}_{q,SE}\boldsymbol{SE}_q + \beta_{j,LI}LI_q + \varepsilon_{jq} \tag{2.1}$$

where $\boldsymbol{SE}_q$ is a vector of individual background characteristics; $\boldsymbol{LOS}_{jq}$ is the vector of vehicle attributes; LI_q is the latent variable that measures the habitual behavior of each individual q; $\boldsymbol{\beta}_{j,LOS}$, $\boldsymbol{\beta}_{q,SE}$ and $\beta_{j,LI}$ are the sets of coefficients associated with the attributes; and ASC_j are the typical alternative specific constants.

The structural equation for the *LI* has the following structure:

$$LI_q = \boldsymbol{\lambda}_{SE}\boldsymbol{SE}_q + \omega_q \tag{2.2}$$

where $\boldsymbol{SE}_q$ is a vector of individual background characteristics that can be different from the vector included in the discrete choice model; $\boldsymbol{\lambda}_{SE}$ is a vector of coefficients associated with these characteristics; and $\omega_{q.}$ is a normal distributed error term with zero mean and standard deviation σ_ω.

The measurement equation of the indicator is specified as:

$$I_r = \delta_r + \zeta_{r,LI}LI_q + \upsilon_{rq} \tag{2.3}$$

where δ_r is a constant of the r_{th} indicator, $\zeta_{r,LI}$ is the estimated effect of the LI_q on the r_{th} indicator, and υ_{rq} is a random disturbance term with zero mean and standard deviation σ_υ.

Three indicators were used in our study to measure habitual behavior: (1) the frequency of cycling for leisure and sport, (2) whether or not an individual uses the bike for shopping, and (3) whether or not an individual uses the bike for work.

The frequency of cycling for leisure and sport is measured using a four-point numerical scale, so the measurement equation of the indicator (I_{lq}) is expressed as an ordered logit model:

$$
\begin{aligned}
&(I_{lq} = 1) = \frac{1}{1 + e^{[\delta_l + \zeta_{l,LI}\boldsymbol{LI}_q - \eta_1]}} \\
&P(1 < I_{lq} < 4) = \frac{1}{1 + e^{[\delta_l + \zeta_{l,LI}\boldsymbol{LI}_q - \eta_I]}} - \frac{1}{1 + e^{[\delta_l + \zeta_{l,LI}\boldsymbol{LI}_q - \eta_{I-1}]}} \\
&P(I_{lq} = 4) = 1 - \frac{1}{1 + e^{[\delta_l + \zeta_{l,LI}\boldsymbol{LI}_q - \eta_3]}}
\end{aligned}
\tag{2.4}
$$

where η_i are thresholds defined respectively as: 1-10 trips per year, 1-5 trips per month, more than once a week, every day. For convenience we also defined $\eta_1 = 0$; $\eta_2 = \eta_1 + \delta_1$; $\eta_3 = \eta_2 + \delta_2$.

The other two indicators (I_{sq} and I_{wq}) are measured as dummy variables, so their measurement equations are binary logit models:

$$P(I_{sq} = 1) = \frac{1}{1 + e^{[\delta_s + \zeta_{s,LI}\boldsymbol{LI}_q]}} \tag{2.5}$$

$$P(I_{wq} = 1) = \frac{1}{1 + e^{[\delta_w + \zeta_{w,LI}\boldsymbol{LI}_q]}} \tag{2.6}$$

Because we assumed that ε_{iq} is i.i.d. Gumbel across alternatives, the probability that decision-maker q chooses alternative i is given by:

$$P_{iq}(\omega_q) = \frac{exp(ASC_i + \boldsymbol{\beta}_{i,LOS}\boldsymbol{LOS}_{iq} + \boldsymbol{\beta}_{q,SE}\boldsymbol{SE}_q + \beta_{j,LI}LI_q)}{\sum_{j \in D_q} exp(ASC_j + \boldsymbol{\beta}_{j,LOS}\boldsymbol{LOS}_{jq} + \boldsymbol{\beta}_{q,SE}\boldsymbol{SE}_q + \beta_{j,LI}LI_q)} \tag{2.7}$$

The joint probability for an individual q making the choice i over the distribution ω_q is:

$$P_{iq} = \int_{\omega} P_{iq}(\omega_q) f_{LI}(\omega_q) \prod_{r} f_{Ir}^{w}\left(I_{rq} \middle| LI_q(\omega_q)\right) d\omega \tag{2.8}$$

where $f_{LI}(\omega_q)$ and $f_{Ir}^{w}\left(I_{rq} \middle| LI_q(\omega_q)\right)$ are the distribution of the latent variable and the indicators, respectively.

The log-likelihood function is given by the logarithm of the product of the unconditional probability:

$$LL = \sum_{q \in Q} ln(P_{iq}) \tag{2.9}$$

All the models were estimated using PythonBiogeme (Bierlaire, 2016).

2.4.3. Results of model estimation

In this section we will discuss the results of the estimation of the ICLV model accounting for inertia. The choice set in the discrete choice model consists of four different modes of transport: 1) private car, 2) public transport, 3) walking, 4) cycling. The time and cost of travel for each mode were determined for each commuter, based on the location of the person's home and work. We simulated, for each individual, the values of the attributes of the non-chosen available alternatives using an algorithm that utilizes Google Maps APIs. This service determines the travel times taking into account the historical traffic conditions at specific times of the day. The algorithm was set so that it considered the route of shortest duration for trip departure time by car, bicycle and on foot, the option with the least number of interchanges for trips using public transport.

The following attributes were tested for each of the four alternatives:

- Motorised mobility
  - Car as a driver
    - Walking time from home to where car is parked;
    - In-vehicle travel time;
    - Time taken looking for a parking space;
    - Walking time from car park to final destination
    - Type of parking (dummy: 1 if paid, 0 otherwise)
    - Stops on the way (dummy: 1 if yes, 0 otherwise)
    - Cost (0.22 €/km that takes into account amortization, fuel, tyres, maintenance and repairs)
  - Public transport:
    - Walking time from home to bus stop/station;
    - In-vehicle travel time (if any transfers are involved this is given by the sum of travel time on each bus/train taken);
    - Walking time to transfer between different public transport lines or mode if any transfers are involved
    - Time taken on alighting from bus/train to walk to final destination
    - Waiting time (if any transfers are involved this is given by the sum of waiting times for each bus/train taken)
      - Equal to half the frequency for frequent services
      - Equal to a randomly assigned value between 5 and 15 minutes for scheduled services
    - Number of transfers
    - Cost (depending on type of ticket and trip frequency)
    - Departure time (dummy: 1 if trip is made between morning peak hours 7:30 to 9:30 am, 0 otherwise)
- Non-motorized mobility (active mobility)
  - Walking
    - Walking time
  - Cycling
    - Travel time

- Parking (dummy: 1 if protected parking at the workplace, 0 otherwise)
- Slope of route:
    - Mean slope of uphill stretches (continuous variable)
    - Maximum uphill slope along the route (continuous variable)
- Trip frequency by bike for work purpose

We tested several ICLV models and the estimation results of the best model are shown in Table 20. In arriving at the final specification, we also tested a number of interaction terms between mode specific attributes and latent inertia, but none of them were statistically significant.

The negative signs of travel times, travel costs, walking time to/from the bus station, walking time to/from the car park and time taken looking for a parking place are consistent with microeconomic theory. It is interesting to note the non-linear effect of waiting time, shown by the significant positive coefficient for the quadratic estimation of this attribute, indicating a threshold above which additional increases in waiting time have a declining effect on the likelihood of mode choice. The negative effect of the peak hour coefficient in the public transport utility function suggests that public transport passengers are more likely to travel during off-peak times to avoid crowds. In fact, as shown in other studies (*e.g.* Tirachini *et al.*, 2013), when the occupancy of buses or trains approaches capacity there might be an increase in both waiting and in-vehicle times, in addition to the discomfort of sharing a limited space with several people.

In terms of network characteristics, not surprisingly, the existence of bike lanes within 400m of home positively affects the utility of the bicycle mode, suggesting that investments in bicycle infrastructure could have a positive impact on the choice to cycle. On the other hand, hilly terrain has a negative impact on the choice to cycle: among the different specifications we tested, the mean slope of uphill stretches was the most significant.

A range of socio-economic variables were found to have a statistically significant influence on mode choice. Males are less likely to travel by car, while, by contrast, individuals with children are more likely to do so. Moreover, as expected, the number of cars per household positively affects the utility to commute to work by car, while the number of bikes positively influences the cycle to work choice.

The specification of the discrete part of the model included the coefficient frequency of going to work by bicycle, which measures the number of times per week an individual commutes by bike. The positive sign of the coefficient shows that as bicycle usage increases, the likelihood to commute to work by bike does so too.

The latent variable inertia, when included in the utility of the bike alternative, is both highly significant and positive. This result indicates that, like other modes of transport, commuting by bike is habit forming and guided by retrieved mental representation of past travel behavior.

Analysis of the estimated parameters of the structural equation shows that the latent variable inertia is positively affected by gender (the propensity to use the bike is greater among men) and the number of bikes in the household. Because the number of bicycles has a positive impact on habitual behavior, it may indicate the existence of a normative social influence.

At the same time, four factors were found to negatively affect bike habits. First, bike travel time from home to workplace negatively impacts the latent variable. Not surprisingly, the increasing number of cars in the family diminishes the probability of an individual's bike habit behavior, a finding that reflects the convenient aspects of owning a private vehicle. The results also show that the presence of children in the household negatively impacts bike habit. A feasible reason for this is that individuals with children have less flexible schedules than people without children, meaning that they do not perceive cycling as the best option to travel between multiple commitments. Lastly, a high level of education (degree or PhD) has a negative influence on bike habit.

Age and income effects were also considered, but they did not turn out to be statistically significant. This is probably because the majority of the sample is composed of public employees, resulting in a limited age and income range.

Table 20. Model results

	DCM alone		HCM	
Attributes	Values	Robust t-test	Values	Robust t-test
Constant Public Transport	0.92	2.05	0.91	2.04
Constant Walking	1.38	4.28	1.37	4.25
Constant Cycling	-3.48	-4.51	-4.25	-4.98
Car as a driver attributes				
Travel time	-0.04	-1.56	-0.04	-1.58
Parking time	-0.01	-0.71	-0.01	-0.67
Walking time	-0.02	-0.91	-0.02	-0.95
Travel cost	-0.46	-4.62	-0.46	-4.61
# of cars in the household	0.40	3.97	0.39	3.89
Male	-0.30	-2.44	-0.28	-2.25
Children	0.37	2.99	0.37	2.97
Public transport attributes				
Travel time	-0.04	-3.26	-0.04	-3.27
Walking time	-0.05	-3.31	-0.05	-3.31
Waiting time	-0.12	-3.31	-0.12	-3.31
Waiting time ^2	0.002	2.28	0.002	2.29
# of transfers	-0.23	-0.94	-0.23	-0.96
Travel cost	-0.29	-2.83	-0.29	-2.84
Peak hour	-0.54	-2.94	-0.54	-2.97
Walking attributes				
Travel time	-0.09	-7.02	-0.09	-7.02
Cycling attributes				
Travel time	-0.08	-4.55	-0.08	-4.63
AVG Slope Max	-0.30	-2.35	-0.30	-2.35
Presence of bike lanes within 400m of home	0.55	1.55	0.57	1.54
# of bikes in the household	0.66	3.37	0.64	3.22
Frequency of bike usage for going to work	1.53	9.18	1.37	7.73
Latent inertia	n/a	n/a	0.81	2.32
Latent inertia structural equation				
Intercept	n/a	n/a	0.82	5.26
Sigma	n/a	n/a	0.82	13.95
Male	n/a	n/a	0.62	6.71
Children	n/a	n/a	-0.33	-3.62
Graduate	n/a	n/a	-0.23	-2.55
# of cars in the household	n/a	n/a	-0.25	-4.07
# of bikes in the household	n/a	n/a	0.25	4.30
Bike travel time to workplace	n/a	n/a	-0.007	-4.51
Indicator: shopping				
Latent inertia	n/a	n/a	2.42	5.4
Intercept	n/a	n/a	-1.23	-4.1
Indicator: work				
Latent inertia	n/a	n/a	3.14	3.72
Intercept	n/a	n/a	-2.98	-3.75
Statistics				
Number of individuals	2128		2128	
LL(max)	-1023.927		-3074.595	
ρ^2 adj	0.744		0.517	

2.4.3.1. Validation

To validate the ICLV model, we randomly split the sample into two parts: the first part contains 80% of the observations of the original data set (estimation set) and the second part contains the remaining 20% of the observations (validation set). We then proceed as follows:

- We first re-estimate the parameters of the model on 80% of the data
- We apply the estimated model to the remaining 20%
- We test the prediction capabilities on the estimation and validation set

Table 21 gives the results of the validation process. Confidence bounds (5% and 95%) were generated by simulation, based on the values of the standard errors of the parameters. We observe that the choice probabilities are quite similar for the two datasets, indicating the robustness of model performance in terms of prediction.

Table 21. Validation of ICLV model

Variables	**Alternative car**		**Alternative Public Transport**		**Alternative walking**		**Alternative cycling**	
	20%	80%	20%	80%	20%	80%	20%	80%
Average choice probability	71.10%	72.39%	12.36%	11.34%	35.16%	35.12%	15.02%	14.44%
Average 5% confidence bound	64.47%	65.87%	8.70%	7.98%	28.70%	28.55%	12.10%	11.40%
Average 95% confidence bound	76.62%	77.79%	17.13%	15.95%	41.89%	41.80%	18.89%	18.44%

2.4.3.2. Demand elasticity

Table 22 shows the elasticity of the demand for bike commuter respect to travel time by bike and the probability of choosing the bike, computed using both models, the DCM and ICLV. Since the latent variable does not modify the marginal utilities, the value of elasticities is similar in both models. Nevertheless, the latent inertia influences the overall utilities and hence the probabilities. The DCM slightly overestimates the probability of choice and it is possible to see this effect also for those categories that are relevant in explaining the latent variable.

Table 22. Elasticity and probability

	Elasticity of the demand for bike commuter with respect to travel time by bike		Probability of choosing to commute by bicycle	
	DCM alone	*ICLV*	*DCM alone*	*ICLV*
Sample average	**-2.2162**	-2.3501	**15.2%**	14.51%
Male	-2.2069	-2.3318	21.2%	20.82%
Graduate	-2.0261	-2.1548	13.9%	12.89%
Children	-2.4070	-2.5533	14.1%	13.45%

2.4.3.3. Forecasting

Based on the results obtained, which provide the scientific evidence to be expected, two different project scenarios are formulated:

- Project 1: creation of a bicycle lane within 400 m of home with no connections with a metropolitan bike network.
- Project 2: creation of a dense network of cycle lanes evenly distributed throughout the areas concerned to enable destinations to be reached safely, using the least cost route (minimum distance) to avoid mixing with vehicle traffic. This would result in a reduction in travel time, in the first place due to the shorter distances traveled, but also due to the fact that being segregated from vehicle traffic, speed increases and less time is wasted at junctions. By way of example, at a conservative estimate this would produce a 20% reduction in travel time by bicycle, clearly to be verified at the design stage.

Both the projects result in an increase of bicycle use percentage compared to the current situation, though, as can be seen, project 2 produced a more marked effect (Table 23).

Table 23. Forecasting scenarios for different categories of individuals

	Current situation	Project 1	Project 2
	Probability of commuting by bicycle		
Total	14.51%	15.30%	16.60%
Male	20.82%	21.81%	23.46%
Female	6.96%	7.51%	8.39%
High level of education (bachelor's degree or higher)	12.89%	13.57%	14.82%
Low level of education	16.54%	17.45%	18.82%
Presence of children in the household	13.45%	14.16%	15.28%
No children in the household	15.88%	16.76%	18.31%

2.5. THE INFLUENCE OF PSYCHO-ATTITUDINAL FACTORS IN CYCLING USAGE AND FREQUENCY

2.5.1. Introduction

Although the vast literature on biking has emphasized the importance of psychosocial factors in understanding what contributes to shaping individual preferences (see 2.2.6), little work has been conducted on the effect of attitudes and perceptions on cycling frequency. In fact, people with different frequency levels are likely to have different attitudes towards cycling.

Heinen *et al.* (2011) studied differences in attitudes between cyclists and non-cyclists, and between full-time and part-time cyclists and analyzed the influence of attitudinal factors on bike commuting over different distances using multiple binary logit models. Fernández-Heredia *et al.* (2014) used a structural equation modeling approach to investigate the difference between the perceptions of users with cycling experience and non-habitual cyclists. Recently, Namgung and Jun (2019) examined attitudes among bicycle users with different experience levels and how these attitudes influence bicycle use in Ohio State University’s campus, employing binary logit models.

However, some works (Heinen *et al.*, 2011; Namgung and Jun 2019) used a two-stage sequential approach without integration for the estimation of their models, that can potentially lead to measurement errors and result in inconsistent estimates. Others (Fernández-Heredia *et al.*, 2014) have considered the frequency variable as a continuous variable, though they measured it in ordinal discrete categories, which is inappropriate from an econometric point of view (Bhat *et al.*, 2017).

A much more systematic approach would be advantageous from both a policy point of view and for representing the decision-making process. In fact, the implementation of effective strategies for the promotion of bike use, such as information campaigns that focus on those factors that could increase cycling, can benefit from an improved understanding of this phenomenon and help to avoid wasting limited resources, as well as failures that would reduce public support (Handy *et al.*, 2014).

In light of these considerations, the object of our research is to study specifically whether psycho-attitudinal factors vary among people with different cycling experience, for any purpose, and to quantify the determinants influencing cycling frequency. The second key contribution is the implementation of an Integrated Choice Latent Variable model with a generalized ordered probit choice kernel. By generalizing the ordered response model, the thresholds themselves are a function of both objectives and psycho-attitudinal variables.

The data used in this study were obtained from the survey “Bici Mi Piaci”. The sample of interest is composed of 2128 individuals. The current research was built around the cycling frequency questions asked in the survey. As seen in 2.2, we asked to individuals to identify their cycling frequency in five levels:

- I never cycle
- 1-10 times per year
- 1-5 times in the past 30 days
- 1-5 days per week
- Everyday

The trips considered for the individuals here are not confined simply to bicycle commuting, so the model developed examines bicycle frequency in general.

2.5.2. Factor analysis

Prior to the modeling phase, a factor analysis (Bollen, 1989; Spearman, 1904) was performed to identify the latent factors underpinning the set of our attitudinal statements. Kaiser-Meyer-Olkin (KMO) measure is used for sample adequacy.

Only two of the three latent constructs included in the survey were found to be suitable for factor analysis: perception of the bicycle as a means of transport (KMO = 0.765) and perception of the context (KMO = 0.806), whereas perception of bikeability (in terms of usefulness and safety) was below the reliability threshold (KMO = 0.572).

Principal axis factoring (PAF) with orthogonal "Varimax" rotation generated two factors for the perception of the bicycle as a means of transport and one factor for the perception of the context. Table 24 shows the results of FA, reporting the loadings of the survey items (table rows) on each of the three identified factors (table columns). Most of the Cronbach's alpha values are above 0.7, except for LV2 that is just acceptable since it is around the "criterion-in-use" of 0.6.

Table 24. Factor scores of the psycho-attitudinal factors towards the bike mode (values below 0.4 are not reported)

Factor	Variables	Loading	Cronbach's alpha
LV1	A1. It is a rapid means of transport (avoids queues and traffic)	0.582	0.690
	A4. It is not expensive	0.540	
	A6. It avoids wasting time looking for parking	0.594	
	A7. It is healthy	0.744	
	A9. It allows one to appreciate historic centers and increases accessibility to city services	0.699	
	A11. It contributes to reducing polluting emissions	0.621	
LV2	A5. It involves exposure to bad weather and air pollution	0.554	0.597
	A8. It is difficult to carry heavy items	0.629	
	A10. Need for cycling gear	0.629	
	A12. It limits daily activity patterns	0.693	
LV3	C1. An extensive network of dedicated bike lanes in urban area	0.907	0.778
	C2. The presence of racks and secure parking for bicycles	0.861	
	C3. A greater extension of the LTZ or pedestrian zones	0.721	

The factors obtained can be interpreted as follow:

- Six items can be considered capturing the *Perception of bike benefits* (LV1) which expresses the agreement related to generally recognized positive features of bikes (inexpensive, healthy, no need to look for parking, non polluting, better appreciation of historic city centers).
- Four items can be considered capturing the *Perception of bikeability* (LV2), which expresses the agreement related to generally recognized negative features of bikes (exposure to bad weather, carrying heavy items, limitations in daily activity patterns, fatigue)
- Three items can be considered capturing the *Perceived Importance of Bike Infrastructures* (LV3), *i.e.* the appeal related to the improvement of bike lanes, racks, parking, traffic calming and pedestrian zones.

2.5.3. Methodology framework

To perform our analysis, we employ an Integrated Latent Variable Choice Model (ICLV) approach with an ordered probit choice kernel (Figure 7).

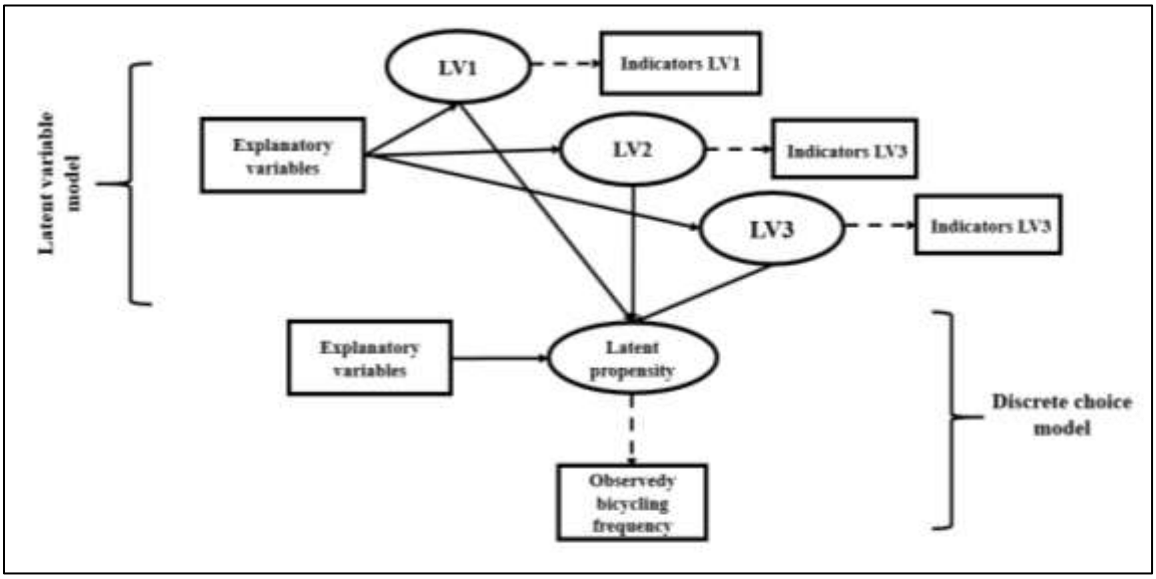


Figure 7. Methodology framework

In our case, the latent propensity underlying the ordered response observation, which is the cycling frequency reported for each individual, has been specified as a function of observed and latent variables:

$$y_q^* = \boldsymbol{\beta x}_q + \boldsymbol{\beta}^* \boldsymbol{LV}_q + \varepsilon_q \qquad (2.10)$$

where $\boldsymbol{x}_q$ is the vector of explanatory variables, $\boldsymbol{LV}_q$ is the vector of individual specific latent variables, $\boldsymbol{\beta}$ and $\boldsymbol{\beta}^*$ are the vectors of unknown parameters to be estimated and ε_q is the error term capturing the effects of unobserved factors on the latent propensity. In the usual ordered-response fashion, the latent propensity y_q^* is linked to the observed level y_q through the thresholds $\boldsymbol{\mu_{q,k}}$ $(\boldsymbol{\mu_{q,0}} = -\infty \textbf{ and } \boldsymbol{\mu_{q,K}} = \infty;\ \boldsymbol{\mu_{q,1}} < \boldsymbol{\mu_{q,2}} < \cdots < \boldsymbol{\mu_{q,K-1}}\ \forall \boldsymbol{q})$. To allow heterogeneity (across observations) in the thresholds, they are parametrized as a function of both objectives and latent variables, as in Williams (2006):

$$\mu_{q,k} = \mu_{q,k-1} + \alpha_k + \boldsymbol{\vartheta}_k' \boldsymbol{x}_q + \boldsymbol{\gamma}_k' \boldsymbol{LV}_q \tag{2.11}$$

Where α_k is a scalar, $\boldsymbol{\vartheta}_k'$ and $\boldsymbol{\gamma}_k'$ are vectors of coefficients associated with level k=1,2, …, K-1. For identifications reasons, we impose the normalization $\mu_{q,k} = const$ for all q. Following the framework of ICLV models, the structural equation for each latent variable is specified as a linear function of respondent's individual and household characteristics:

$$\boldsymbol{LV_q} = \boldsymbol{\kappa} + \boldsymbol{\lambda x_q} + \boldsymbol{\omega_q} \tag{2.12}$$

where $\boldsymbol{\kappa}$ is the intercept, $\boldsymbol{x_q}$ is a vector of individual background characteristics that can be different from the vector included in the discrete choice model; λ is a vector of coefficients associated with these characteristics; and $\boldsymbol{\omega_q}$ is a normal distributed error term with zero mean and standard deviation σ_ω. As in the typical ICLV theory, the measurement equation of the indicators is specified as

$$I_{rq} = \delta_r + \zeta_r LV_q + \upsilon_{rq} \tag{2.13}$$

where δ_r is a constant of the r_{th} indicator, ζ_r is the estimated effect of the LV on the r_{th} indicator, and υ_{rq} is a random disturbance with a mean of zero and a standard deviation of σ_I ($\delta_r$and ζ_r are normalized respectively to zero and 1 for one of the indicators of each latent variable for identification purposes). The distribution of the indicators I_{rq} is then defined by allocating the variables $\tilde{I}_{rq}$ to intervals given by the thresholds $\rho_k^{(0)}, \dots, \rho_k^{(S)}$:

$$\begin{aligned} I_{rq} &= 1 \ if \ \rho_r^{(0)} < \tilde{I}_{rq} \le \rho_r^{(1)} \\ I_{rq} &= 2 \ if \ \rho_r^{(1)} < \tilde{I}_{rq} \le \rho_r^{(2)} \\ I_{rq} &= \cdots \\ I_{rq} &= 5 \ if \ \rho_r^{(S-1)} < \tilde{I}_{rq} \le \rho_r^{(S-1)} \end{aligned} \tag{2.14}$$

Assuming the error components υ_{rq} are i.i.d. and follow a Gumbel distribution for all n and k, we obtain an ordered logit model. The probability for a certain response s to the r-th indicator given the latent variables and the parameters ζ_r is thus given by:

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

$$P(1 < I_{rq} < I) = \frac{e^{\left[\delta_r + \zeta_r LV_q - \rho_{I_{rq-1}}\right]}}{1 + e^{\left[\delta_r + \zeta_r LV_q - \rho_{I_{rq-1}}\right]}} - \frac{e^{\left[\delta_r + \zeta_r LV_q - \rho_{I_{rq}}\right]}}{1 + e^{e^{\left[\delta_r + \zeta_r LV_q - \rho_{I_{rq}}\right]}}} \quad (2.15)$$

$$P(I_{rq} = I) = \frac{e^{\left[\delta_r + \zeta_r LV_q - \rho_{I-1}\right]}}{1 + e^{\left[\delta_r + \zeta_r LV_q - \rho_{I-1}\right]}}$$

where I = 5. Because we assumed that ε_q is normal distributed the unconditional probability that decision-maker q belongs to category k is given by:

$$Prob(y = k) = \Phi(\mu_k - \boldsymbol{\beta x}_q - \boldsymbol{\beta}^* \boldsymbol{LV}_q) - \Phi(\mu_{k-1} - \boldsymbol{\beta x}_q - \boldsymbol{\beta}^* \boldsymbol{LV}_q) \quad (2.16)$$

Hence, the joint likelihood function for individual *q* may be written as follows:

$$P(y, I|\boldsymbol{x}; \boldsymbol{\zeta}, \boldsymbol{\beta}, \boldsymbol{\beta}^*, \boldsymbol{\lambda}, \Sigma_\varepsilon, \Sigma_\upsilon, \Sigma_\omega) = \int_{\boldsymbol{x}^*} P(y|\boldsymbol{x}, LV; \boldsymbol{\beta}, \boldsymbol{\beta}^*, \mu_k, \Sigma_\varepsilon) f(I|\boldsymbol{x}, \boldsymbol{x}^*; \boldsymbol{\zeta}, \Sigma_\upsilon)\, g(\boldsymbol{x}^*|\boldsymbol{x}; \boldsymbol{\lambda}, \Sigma_\omega) d\boldsymbol{x}^* \quad (2.17)$$

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

2.5.4. Model results

Table 25 presents results of the structural equation component of the model.

The *Perception of bike benefits* construct is affected by both the level of education and number of bikes in the household. It is clear that people with a lower level of education are more inclined to recognize the positive aspects of cycling. Owning several bikes has a significant influence on this latent construct as well.

The second construct is the *Perception of bikeability*. Compared to the female respondents, it seems that males are more comfortable with cycling, indicating a better perception of bikeability. This can be partly explained by the fact that in Italy women tend to make more trips for household activities (ISTAT, 2014), which often require the transport of goods or passengers (e.g. children) or trip chaining. The positive sign associated with the dummy variable *18 to 30 years old* indicates that younger people feel less strongly about the limitations of the bike mode. In line with other studies (Heinen *et al.*, 2010; Hamre and Buehler, 2014), we also found that the presence of children at home means that people are less inclined to cycle as they need to accompany youngsters usually by other means of transport. The possession of a personal car and the number of cars in the household negatively influence the latent construct. This effect could reflect the fact that car-addicted individuals, since they are used to the car's comfort, tend to overestimate the disadvantages of cycling. On the other hand, the number of bicycles in the household has a positive effect on the latent variable. Furthermore, people with a high level of education (bachelor's degree or higher) are likely to notice the negative aspects of cycling.

The structural model related to the *Perceived importance of bike infrastructure* indicates that women tend to consider the presence of bike infrastructure more important than men. This gender-based effect could reflect the fact that women are more concerned about traffic and safety, a finding consistent with the literature (Akar *et al.*, 2013; Bhat *et al.*, 2015; Manton *et al.*, 2016). Furthermore, people with a lower education (high school degree or lower) consider the presence of some complementary facilitators encouraging cycling more important than those with a higher education. Interestingly, the presence of children in the household negatively influences the latent construct, indicating the presence of barriers of other kinds for this segment of individuals.

Table 25. Determinants of latent constructs. "--"in a cell indicates that the variable in the corresponding row does not have a significant impact on the utility of the alternative in the corresponding column.

Explanatory Variables	**LV1 - Attitudes on bike performances & benefits**		**LV2 -Attitudes on bike limitations**		**LV3 - Perceived importance of bike infrastructure**	
	Coeff	R T-stat	Coeff	R T-stat	Coeff	R T-stat
18 to 30 years old	--	--	0.119	1.07	--	--
Gender (base: female)	--	--	0.090	1.93	-0.169	-2.48
Bachelor's degree or higher	-0.389	-3.28	-0.114	-2.40	-0.130	-1.74
# of bikes in the household	0.276	5.02	0.151	6.39	0.123	3.28
# of cars in the household	--	--	-0.078	-2.36	--	--
Presence of children	--	--	-0.153	-3.01	-0.124	-1.73
Constant	5.930	32.06	1.950	24.65	2.650	27.15
Variance	1.770	10.80	0.704	23.41	1.230	20.11

Table 26 presents the results of the measurement model. Several indicators were considered in the latent variable measurement model, which linked the latent variables of the responses to the qualitative attitudinal questions in the survey. The α parameters that indicate the associations between the responses to the scale items and the psychometric scale, all have the expected signs. For example, a more positive *Perception of bike benefits* will lead to respondents being more in agreement with the statements about the bike as a healthy and environmentally friendly mode of transport.

Table 26. Impact of Latent Variables on Non-nominal Dependent Variables. “n/a” not applicable.

Latent Variable	Indicators	Const	R T-stat	Coef.	R T-stat
Perception of bike benefits	A1. It is a rapid means of transport	-0.642	-1.46	0.803	8.79
	A4. It is not expensive	0.184	0.48	0.744	9.23
	A6. It avoids wasting time looking for parking	-0.747	-1.84	0.802	9.26
	A7. It is healthy	-1.560	-2.70	1.310	9.90
	A9. It allows one to appreciate historic centers and increases accessibility to city services	-0.606	-1.58	0.842	10.48
	A11. It contributes to reducing polluting emissions	0.000	n/a	1.000	n/a
Perception of bikeability	A5. It does not involve exposure to bad weather and air pollution	-1.050	-5.65	1.520	10.91
	A8. It is easy to carry heavy items	-2.740	-10.96	1.840	11.00
	A10. No need for cycling gear	0.000	n/a	1.000	n/a
	A12. It does not limit daily activity patterns	0.104	0.70	1.670	14.01
Perceived importance of bike infrastructure	C1. Presence of an extensive network of dedicated bike lanes	-0.878	-2.31	3.330	10.79
	C2. Presence of racks and secure parking for bicycles	-1.060	-2.70	3.560	12.12
	C3. Greater extension of the LTZ or pedestrian zones	0.000	n/a	1.00	n/a

The estimation results of the discrete part of the model are shown in Table 27. For reasons of identification, there is no constant in the latent bicycle propensity and the first threshold is considered fixed, with no socio-demographic attributes or latent variables in its specification.

Table 27 presents three columns. The first column corresponds to the estimate of the parameters characterizing the latent bicycle propensity. The second column corresponds to the estimates of constant and the parameters for the μ_2 threshold (threshold between occasional cycling and infrequent cycling). The third and the fourth columns correspond to the constants and the parameters linked to the third threshold, delimiting the infrequent and frequent cycling categories, and to the fourth threshold, delimiting the frequent and everyday cycling categories.

Some socio-economic variables were found to have a significant effect on the propensity to bike. In agreement with several other previous studies, males are more likely to cycle. In addition to the effect on cycling propensity, the gender variable also impacts the thresholds in the framework of the model. A negative coefficient in the vector μ_k for a variable has the effect of shifting the corresponding threshold to the left. In this case, the pattern of threshold effects indicates that men, compared to women are more likely to be infrequent cyclists than would be predicted by the standard-response model. Regarding age, older individuals have a lower propensity to cycle, which has been found in many other studies too. Interestingly, the Body Mass Index (a biometric datum used as an indicator of ideal weight) negatively affects the propensity to bike. The causal relationship underlying this correlation could go in either or both directions:

healthier people are more likely to use the bike or people who bicycle are likely to be healthier due to the benefits of physical activity gained from bicycling.

The household demographics that play an important role in bicycling frequency decisions are: number of cars, number of bicycles, presence of children. As expected, the number of bikes per family positively affects the utility of using the bike. By contrast, individuals with children are less likely to do so. This makes sense because cars are still the major means of transport for most individuals with children, who are less likely to bicycle. Also, the number of cars in the household has a significant negative impact on bicycle use. The effect on the threshold indicates that, compared to the standard ordered probit, the generalized ordered probit predicts a higher probability of individuals with a large number of cars not belonging in the everyday cyclists category.

All the attitudinal factors have positive effects on the propensity to bicycle. The positive influence of the latent variable *Perceived importance of bike infrastructure* emphasizes the importance of providing bicycle facilities such as bike paths or bike lanes, rental bikes, and bike storage. People who are aware of bicycling benefits, such as protecting the environment as well as keeping fit (latent variable *Perception of bike benefits*), were more likely to be cyclists. This finding suggests that if benefits of bicycling are better understood, more people are likely to cycle more frequently. *Perception of bikeability* is also positively associated with the propensity to cycle. This can be explained as follows: as the level of experience rises, people feel that bicycling is easier and more accessible and consequently they use the bicycle with greater frequency. Interestingly, the effects on the thresholds show that the generalized ordered probit predicts a lower probability with a higher value of the LV2 and LV3 of not being cyclists.

Table 27. Estimation results of the hybrid generalized ordered probit. "--"in a cell indicates that the variable in the corresponding row does not have a significant impact on the utility of the alternative in the corresponding column. "n/a" not applicable.

Variables	**Latent propensity to cycle**		**Threshold between occasional cycling and infrequent cycling**		**Threshold between infrequent bicycling and frequent cycling**		**Threshold between frequent bicycling and everyday bicycling**	
	Estimate	*R t-stat*	*Estimate*	*R t-stat*	*Estimate*	*R t-stat*	*Estimate*	*R t-stat*
Threshold constants	n/a	n/a	0.972	6.67	0.819	6.30	0.738	5.42
Age	-0.185	-6.58	--	--	--	--	--	--
Gender (male = 1; female =0)	0.591	9.15	-0.216	-3.81	--	--	--	--
Bachelor's degree or higher	-0.231	-3.90	0.107	1.87	--	--	--	--
Body Mass Index	-0.067	-8.66	--	--	--	--	--	--
# of bikes in the household	0.804	25.00	--	--	--	--	--	--
# of cars in the household	-0.066	-1.39	--	--	--	--	0.148	1.84
# of household members	-0.299	-10.56	--	--	--	--	--	--
Residential location (urban)	0.089	1.45	--	--	--	--	--	--
Presence of bike paths within 400m of home	0.090	1.66	--	--	--	--	--	--
LV1 – Perception of benefits	0.061	2.96	--	--	--	--	--	--
LV2 – Perception of bikeability	0.265	4.64	-0.119	-2.05	-0.127	-1.71	--	--
LV3 - Perceived importance of bike infrastructure	0.055	1.81	-0.065	-2.20	--	--	--	--

2.5.5. Elasticities effect

The coefficients in Table 27 do not provide a sense of the magnitude and direction of effects of each variable on each bicycling frequency category. But we can compute aggregate-level “pseudo-elasticity effects” of exogenous variables, that can be calculated as:

$$\Delta P\big(y = k | x_q, \tilde{x}_q\big) = P\big(y_q = k | \tilde{x}_q\big) - P\big(y_q = k | x_q\big) \tag{2.18}$$

where all elements of $\tilde{x}_q$are equal to x_q except for the v_{th} element, which is equal to $\tilde{x}_{qv} = x_{qv} + \Delta x_{qv}$ for the discrete change Δx_{qv} in the variable x_v . Hence, for dummy variables, we first predict the probabilities of each bicycling frequency level for each individual, assigning the base value of “1” for all individuals. All other exogenous variables take their values from the original data. Then we compute the percentage change in the expected number of individuals at each bicycling frequency level. It is important to note here that, since the latent variables influencing each level of cycling frequency are a function of exogenous variables, a change in the exogenous variables leads to a change in the value of the latent variables as well.

Table 28 provides the pseudo-elasticity effects. The numbers in the table may be interpreted as the percentage change in the probability of each bicycling frequency level due to a change in the exogenous variable. For example, the first entry in the table indicates that the probability of a man not being a cyclist is 16.1% lower than the probability of a woman not being a cyclist, everything else being equal. For the bike lanes variable, the probability of an individual not being a cyclist reduces by 2.4% if bicycle lanes are provided in the individual's residential neighbourhood. The directions of the elasticity effects of the model are consistent with the discussions in the previous section.

Table 28. Pseudo-elasticities effect

Variables	**I never cycle**	**1-10 times per year**	**1-5 times in the past 30 days**	**1-5 days per week**	**Everyday**
Gender (male)	-16.1%	-11.1%	15.9%	28.2%	48.9%
Education (bachelor's degree or higher level of education)	9.3%	9.5%	-8.5%	-17.9%	-32.8%
Residence location choice (urban)	-3.8%	1.8%	2.9%	4.6%	8.0%
Bike lanes within 400m of home	-2.4%	1.1%	1.8%	2.9%	5.2%

2.6. SUBJECTIVE AND OBJECTIVE FACTORS INFLUENCING THE CHOICE TO CYCLE FOR DIFFERENT PURPOSES

2.6.1. Introduction

One aspect often overlooked in transportation research is how sociodemographic, territorial and psycho-attitudinal factors influence cycling for different purposes. Much of the research has focused on cycling for all purposes, mixing utilitarian and recreational trips, but the determinants (both objective and subjective) triggering the choice to travel by bike may be different, depending on the reason people cycle. The other question concerns the existence of a relationship between cycling for leisure and the choice to cycle for utilitarian purposes. Some studies found that cycling for leisure may increase the likelihood to cycle for commuting or shopping, but others argue that the direction of causality may not be very clear.

To better understand the interplay between the use of the bike for different purposes, we conduct a multivariate analysis. In particular, a hybrid multidimensional choice model, capable of jointly accounting for mixed types of dependent variables, was developed and estimated. Three behavioral choice variables were considered: the choice to use the bike for commuting, the choice to use the bike for shopping and errands and the frequency of bicycling for leisure and sport. Further, we included in our analysis some psychological latent constructs that allowed us to identify differences in cycling proclivity within a context with non-cycling culture.

The data used in this study are derived from the survey called “Bike I like you”. The sample of interest is composed of 1,105 individuals, namely those reporting they had access to a bicycle for the home-work trip (see 2.3.6 for a complete analysis of the sample).

2.6.2. Behavioral framework

The choice variables are estimated simultaneously with a comprehensive modeling framework in which latent constructs and individual and household characteristics serve as explanatory variables. There are three simultaneous choice models for the following endogenous outcomes:

- One ordered choice variable representing the frequency of using the bike for leisure
- Two binomial choice variables, including:
  - Using/Not using the bike for commuting
  - Using/Not using the bike for shopping and errands

To identify hidden latent constructs an exploratory factor analysis is carried out. The factor loadings are estimated using principal component analysis with varimax rotation. To establish if the dataset is suitable for exploratory factor analysis, sample adequacy and strength of the inter-correlation of items must be examined. Kaiser-Meyer-Olkin (KMO) measure is used for sample adequacy: KMO values between 0.8 and 1 indicate the sampling is adequate. The Bartlett test of sphericity is used to test the hypothesis that the correlation matrix is an identity matrix, which would indicate that variables are unrelated and therefore unsuitable for structure detection. Furthermore, to examine reliability Cronbach's alpha value is used. A Cronbach's alpha value higher than 0.6 indicates that the dataset is reliable and acceptable.

Six factors were retained based on percentage of total variance in the original variables explained by the factors. However, it is clear from Table 29 that Cronbach's alpha value is higher than 0.6 only for constructs 1, 3 and 4. Therefore, only these two constructs are extracted and included in the analysis.

Table 29. Results of Factor Analysis. Extraction method: principal component analysis, rotation method: varimax with Kaiser normalization.

Item	**Loading**	**Factor**	**Cronbach's alpha**
A1	0.588	1	0.677
A4	0.484		
A6	0.608		
A7	0.719		
A9	0.699		
A11	0.607		
A3	0.761	2	0.412
A5	0.646		
A8	0.505		
C1	0.859	3	0.762
C2	0.812		
C3	0.703		
C4	0.655	4	0.820
C6	0.555		
C7	0.912		
C8	0.871		

The latent variable *Perceived importance of bike infrastructure* uses indicators capturing the appeal of some factors that would facilitate use of the bike, while the latent variable *Attitudes on bikes performance & benefits* expresses the agreement related to generally recognized positive features of bikes. We will not include factor 4 in our behavioral framework as it regards the integration of bicycle use with transit service, and our dependent variables do not explicitly consider it.

Figure 8 presents the conceptual framework of the model. The latent variables (constructs) are represented by the ovals, while the endogenous outcomes (*i.e.*, commute mode choice) considered are identified in the rectangular boxes.

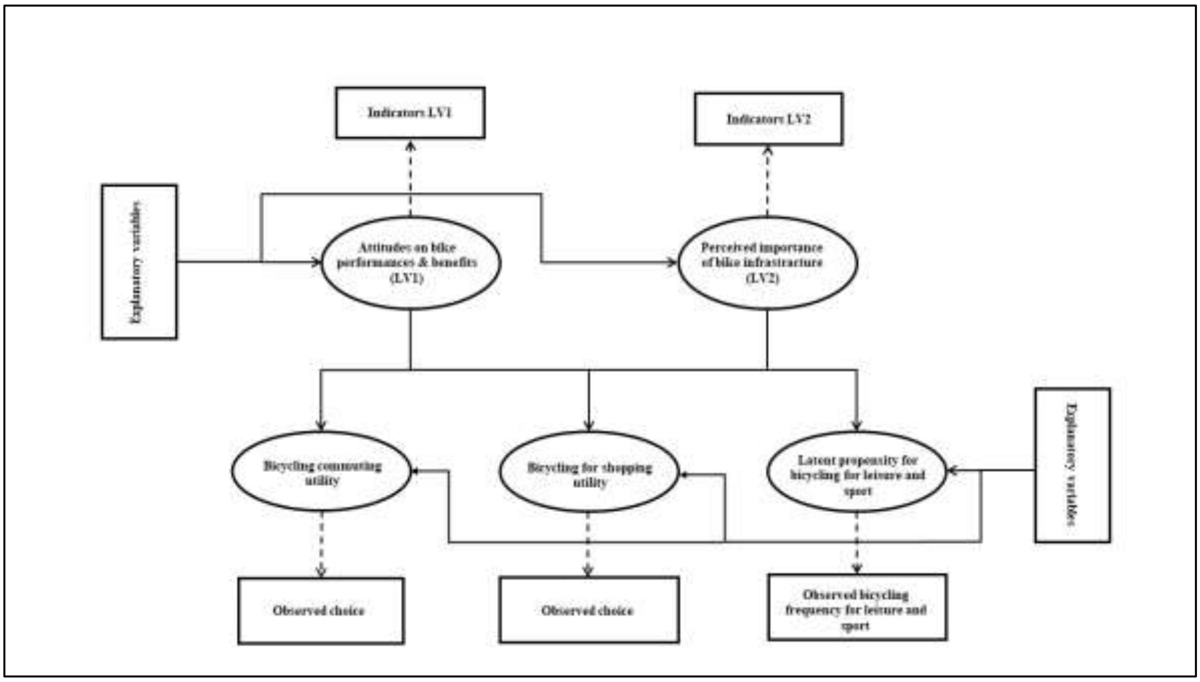


Figure 8. Methodology framework

2.6.3. Modeling framework

The econometric approach taken in this study is a special case of the GHDM model proposed by Bhat (2015). In this approach, a series of sub-models are formulated for different choice dimensions - a binary probit that models the choice to use *vs* not to use the bike for commuting, a binary probit that models the choice to use *vs* not to use the bike for shopping and errands and an ordered probit model of bicycling frequency for leisure and sport, a structural equation model for the latent variables. All the models are econometrically joined together by means of the presence of common random coefficients included in latent variables specification.

Following the framework of ICLV models, the structural equation for each latent variable is specified as a linear function of the respondent's sociodemographic factors:

$$LV_{wq} = \kappa_w + \boldsymbol{\lambda_w x_q} + \omega_{wq} \quad (2.19)$$

where $\boldsymbol{x_q}$ is a vector of individual sociodemographic characteristics; λ is a vector of coefficients associated with these characteristics; and ω_{wq} is a normal distributed error term with zero mean and standard deviation σ_ω. As in the typical ICLV theory, the measurement equation of the indicators is specified as

$$I_{rq} = \delta_r + \zeta_r LV_{wq} + \upsilon_{rq} \quad (2.20)$$

where δ_r is a constant of the r_{th} indicator, ζ_r is the estimated effect of the LV on the r_{th} indicator, and υ_{rq} is a random disturbance with a mean of zero and a standard deviation of σ_I ($\delta_r$and ζ_r are normalized respectively to zero and 1 for one of the indicators of each latent variable for identification purposes). Since the psychometric indicators revealing the latent variables were coded using a Likert scale, we treat them as ordered choices.

Let U_{kq} be the utility that each individual q associates with the alternative of cycling for commuting. The utility function for the alternative k can be written as:

$$U_{kq} = \beta_{k,0} + \boldsymbol{\beta}_{k,LOS}\boldsymbol{LOS}_{kq} + \boldsymbol{\beta}_{k,SE}\boldsymbol{SE}_q + \boldsymbol{\beta}_{k,LV}\boldsymbol{LV}_q + \nu_{kq}\text{ ; alternative k chosen if } u_{kq} > \max u_{dq}; \quad (2.21)$$

where $\boldsymbol{LOS}_{kq}$ is the vector of level-of-service (LOS) characteristics for the mode bicycle , $\boldsymbol{SE}_q$ is a vector of socio-demographics characteristics, $\boldsymbol{LV}_q$ is a vector of latent variables associated with each individual q, $\beta_{k,0}$ is the alternative specific constant, $\boldsymbol{\beta}_{k,LOS}$ is a vector of coefficients associated with LOS characteristics, $\boldsymbol{\beta}_{k,SE}$ is a vector of coefficients associated with socio-demographics attributes and $\boldsymbol{\beta}_{k,LV}$ are the sets of coefficients associated to latent variables. ν_{kq} is a an independent and identically distributed normal error term.

The utility associated with the choice to cycle for shopping and errands is:

$$z_{iq} = \varphi_{i,0} + \boldsymbol{\varphi}_{i,s}\, \boldsymbol{s}_q + \boldsymbol{\varphi}_{i,LV}\boldsymbol{LV}_q + \varepsilon_{qi}\text{ ; alternative k chosen if } z_{iq} > \max z_{jq}; \quad (2.22)$$

s_q is a vector of sociodemographic characteristics associated with individual q, $\boldsymbol{LV}_q$ is a vector of latent variables associated with each individual q, $\varphi_{i,0}$ is a constant term, $\varphi_{i,s}$ is a vector of the effects of the variables s_q on the latent utility and $\varphi_{i,LV}$ is a vector of coefficients associated with the latent variables. Finally, ε_{qi} is a random-error term assumed to be identically and independently normal distributed across individuals q.

Finally, the latent propensity underlying the ordered response observation, that is the bicycling frequency for non-commuting purposes, has been specified as a function of observed and latent cycling variables:

$$y_q^* = \boldsymbol{\alpha x}_q + \boldsymbol{\alpha}_{LV}\boldsymbol{LV}_q + \xi_q,\ y_q = n if \mu_n < y_q^* < \mu_{n+1}\ \mu_{q,0} = -\infty\ and\ \mu_{q,N+1} = \infty \quad (2.23)$$

where $\boldsymbol{x}_q$ is the vector of explanatory variables, $\boldsymbol{LV}_q$ is the vector of individual specific latent variables, $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha_{LV}}$ are the vectors of unknown parameters to be estimated. ξ_q is an independently and identically distributed normal error term.

2.6.3.1. *Model estimation*

Let Ξ represent a vector of the parameters to be estimated, Ξ_{Σ} a vector of all parameters except the variance terms and Ω_q a vector that stacks all the error terms of the structural equations that define the latent variables. Also, define

$a_{kq} = 1$ if individual q chooses to commute by bicycle and 0 otherwise

$b_{iq} = 1$ if individual q chooses to bicycle for shopping and 0 otherwise

c_{nq} = 1 if individual q cycles for leisure and sport with frequency n and 0 otherwise

The conditional likelihood function for an individual q is then:

$$
\begin{aligned}
LL_q(\Xi_{\Sigma})|\Omega_q = \prod_{j=1}^{J}\prod_{k=1}^{K}\prod_{n=1}^{N}\Big\{\big[&D\big(\beta_{k,0} + \boldsymbol{\beta}_{k,LOS}\boldsymbol{LOS}_{kq} + \boldsymbol{\beta}_{k,SE}\boldsymbol{SE}_q + \boldsymbol{\beta}_{k,LV}\boldsymbol{LV}_q - \beta_{d,0} \\
&- \boldsymbol{\beta}_{d,LOS}\boldsymbol{LOS}_{kq} - \boldsymbol{\beta}_{d,SE}\boldsymbol{SE}_q - \boldsymbol{\beta}_{d,LV}\boldsymbol{LV}_q\big)\big] \\
&\times \big[D\big(\varphi_{i,0} + \boldsymbol{\varphi}_{i,s}\,\boldsymbol{s}_q + \varphi_{i,LV}\boldsymbol{LV}_q - \varphi_{i,0} - \boldsymbol{\varphi}_{i,s}\,\boldsymbol{s}_q - \varphi_{i,LV}\boldsymbol{LV}_q\big)\big] \\
&\times \big[D\big(\mu_{n+1} - \boldsymbol{\alpha x}_q - \boldsymbol{\alpha}_{LV}\boldsymbol{LV}_q - \eta_{iq} - \varsigma_{kq}\big) \\
&- D\big(\mu_n - \boldsymbol{\alpha x}_q - \boldsymbol{\alpha}_{LV}\boldsymbol{LV}_q - \eta_{iq} - \varsigma_{kq}\big)\big]\Big\}^{a_{iq}\times b_{kq}\times c_{nq}}
\end{aligned}
\tag{2.24}
$$

where $D(\cdot)$ is the cumulative distribution of the standard normal distribution. Finally, the unconditional likelihood can be computed as:

$$LL_q(\Omega) = \int_{\Omega_q} (\mathrm{LL}_q(\Xi_{\Sigma})|\Omega_q) f_{LV}(\Omega_q) \prod_r f_{I_r}\left(I_{rq}\middle|LV_q(\Omega_q)\right) d\Omega \tag{2.25}$$

The log-likelihood function for the entire data set is:

$$LL(\Omega) = \sum_q ln(LL_q(\Omega)) \tag{2.26}$$

We apply simulation techniques to approximate the integral in (2.25) and maximize the resulting simulated log-likelihood function. All of the parameters in the model are estimated by maximizing the log-likelihood function using PythonBiogeme software (Bierlaire, 2016).

2.6.4. Model results

This section presents a discussion of the model estimation results. To arrive at the final model specification, we tested and examined, with respect to statistical measures of fit, several model specifications. In some cases, variables with marginally significant statistical effect were left in the specification, because of the intuitiveness of the effect of the variable and the ability to explain the phenomenon.

Table 30 shows the results of the determinants of latent constructs. The latent construct *Perception of bike benefits* is affected by age, level of education and level of bike ownership. The age related perception difference can be explained by the fact that, as shown in other studies (Sener *et al.*, 2009) younger individuals are more conscious of the benefits of using the bike. People with a lower level of education are more likely to recognize the positive aspect of cycling. Owning more bikes has a positive and significant influence on this latent construct as well.

The second construct is the *Perceived Importance of bike infrastructure*. The presence of children in the household negatively influences the latent construct, suggesting that the existence of bicycle lockers or safe storage rooms and the presence of a dense bicycle network are not sufficient to encourage these individuals to cycle. The level of income also significantly impacts the latent factor, individuals with a lower income being more likely to recognize the importance of bike infrastructure in the choice to cycle. One possible explanation is that this segment of population, due to their fewer financial resources, value safe paths and secure parking much more highly, as the bicycle is a relatively expensive possession that can be stolen or damaged due to an accident. Further, the number of bicycles in the household has a positive effect on the latent variable.

Table 30. Determinants of latent constructs. "--"in a cell indicates that the variable in the corresponding row does not have a significant impact on the utility of the alternative in the corresponding column.

Explanatory Variables	**LV1 – Perception of bike benefits**		**LV3 - Perceived importance of bike infrastructure**	
	Coeff	R T-stat	Coeff	R T-stat
Age	-0.059	-1.34	--	--
Bachelor's degree or higher	-0.324	-3.92	--	--
Presence of children in the household	--	--	-0.193	-1.79
# of bikes in the household	0.187	3.87	0.166	2.51
Personal Income	--	--	-0.102	-1.59
Mean	2.030	8.34	2.800	13.25
Sigma	0.962	11.52	1.100	10.09

Estimation results for the measurement equations component are given in Table 31. The variables included 9 ordinal indicators, six measuring the *Perception of bike benefits* and three corresponding to the *Perceived Importance of bike infrastructure*. The constants indicate the overall preferences of the respondents but do not have a behavioral interpretation. All the loadings of the latent constructs on the indicators are, as expected, positive and significant. For example, the latent variable *Perception of bike benefits* has a positive effect on the indicators regarding the positive aspects linked to the use of the bike.

Table 31. Impact of Latent Variables on Non-nominal Dependent Variables

Latent Variable	Indicators	*Const*	*T-stat*	*Coef.*	*T-stat*
LV1. Perception of bike benefits	A1. It is a rapid means of transport	-0.012	-0.05	1.66	8.7
	A4. It is not expensive	1.65	6.47	1.25	7.48
	A6. It avoids wasting time looking for parking	0.245	0.95	1.69	8.87
	A7. It is healthy	1.66	4.53	2.2	6.54
	A9. It allows one to appreciate historic centers and increases accessibility to city services	0	n/a	1	n/a
	A11. It contributes to reducing polluting emissions	1.85	5.63	1.53	6.59
LV3. Perceived importance of bike infrastructure	C1. Presence of an extensive network of dedicated bike lanes	-2.59	-2.46	4.38	5.21
	C2. Presence of racks and secure parking for bicycles	-1.71	-2.03	3.38	5.73
	C3. Greater extension of the LTZ or pedestrian zones	0	n/a	1	n/a

Table 32 presents the estimation results for the measurement equation components associated with the binary and ordered variables.

Males exhibit a greater propensity to use the bike for utilitarian purposes than females. The effect here could be attributed to the fact that in general women tend to make more trips for household activities, which sometimes require the transport of goods or passengers (Emond *et al.*, 2009). For this reason, they may prefer the use of more convenient, in terms of travel time, means of transport. Also, for recreational purposes men tend to be more inclined to cycle than women. This result is in line with other studies (Heesch *et al.*, 2012; Menai *et al.*, 2015) showing that males possess extra motivation when it comes to cycling for recreational purposes.

Further, individuals with children are less likely to travel for utilitarian purposes, as they often have to trip chain and complete pick up/drop off tasks, which is burdensome especially if these drop off/pick-up locations are not close to their route. As stated in other studies (Sener *et al.*, 2009; Hamre and Buehler, 2014; Ton *et al.*, 2019) this more complex travel pattern makes biking to work less feasible. We observe the same effect for the propensity to cycle for leisure, but here the interpretation of the result may be different. In fact, it may be that individuals with children have less free time to pursue leisure activities, due to their parental duties.

Consistently with previous research (Parkin *et al.*, 2008; Pinjari *et al.*, 2011), as the number of cars per household rises, the likelihood to commute by bike falls. Instead, the level of bike ownership positively influences the choice to cycle to work, as well as the choices to cycle for shopping and leisure.

A strong determinant in the choice to use active mobility is distance. In agreement with previous studies (Parkin *et al.*, 2008; Portoghese *et al.*, 2011), people who live closer to their workplace are more likely to active commute than those with longer commutes. Interestingly, living in urban areas, positively influences the propensity to cycle for leisure and sport, presumably because of the existence of better biking facilities, such as cycle lanes and a bike-sharing service, as well as the presence of recreation activity locations nearby. From a policy point of view, these outcomes emphasize the importance of the built-environment in the choice to use active mobility.

We also account for observed endogenous effects, and in particular we found that cycling for leisure and sport positively influences the choice to cycle for utilitarian purposes, consistent with the previous literature (Stinson and Bhat, 2004; Park *et al.*, 2011). Note that this is a "true" causal effect because of the presence of the common error terms of the latent variables in the utility functions of choice dimensions.

Regarding the latent variables, they were found to positively influence the choice to cycle, for commuting, shopping and leisure. This means that, regardless of purpose, psychological factors must be taken into account to evaluate individual's propensity to cycle. Furthermore, since all the latent variables come with a positive sign, the indirect effect of the exogenous variables in the latent constructs influence all the outcomes with the same directionality. This means that a change in those variables influence, though with a different weight in each utility function, the propensity to cycle for all purposes.

Table 32. Parameter estimates of the binary and ordered variables. "--"in a cell indicates that the variable in the corresponding row does not have a significant impact on the utility of the alternative in the corresponding column. "n/a" not applicable.

Explanatory variables	**Commuting (binary)**		**Shopping (binary)**		**Leisure and sport (ordinal)**	
	Coeff	*T-stat*	*Coeff*	*T-stat*	*Coeff*	*T-stat*
Constant	-3.540	-9.68	-2.160	-6.71	--	--
Gender (male =1 , female = 0)	0.603	4.85	0.595	6.38	0.586	7.89
Age	--	--	-0.152	-2.85	-0.093	-2.47
Body Mass Index	--	--	--	--	-0.255	-2.66
Children (base: No)	-0.366	-2.74	-0.121	-1.11	-0.155	-1.52
# of components in the household	--	--	--	--	-0.145	-3.18
# of bikes in the household	0.337	3.94	0.173	2.64	0.529	10.81
# of cars in the household	-0.213	-2.42	-0.202	-2.88	-0.119	-2.03
Log commuting distance	-0.314	-5.16	--	--	--	--
Commuting AVG Slope Max	-0.025	-0.70	--	--	--	--
Presence of bike lanes within 400m of home	0.396	3.43	--	--	--	--
Urban (Urban = 1; Suburban and rural = 0)	--	--	--	--	0.241	3.08
Frequency of bicycle use for leisure and sport	0.414	7.25	0.688	15.49	--	--
Loading of latent variables						
LV1 – Perception of bike benefits	0.451	3.81	0.322	3.86	0.32	5.49
LV2 - Perceived importance of bike infrastructure	--	--	--	--	0.129	2.86

2.6.5. Data Fit

The data fit from jointly modeling the three choice dimensions (Table 33) can be assessed by comparing the joint model with an independent model that does not include any error term and the latent psychological variables do not enter into the specification of the utility functions. This model is an independent model because error term correlations across the choice dimensions are ignored. The value of the adjusted likelihood ratio index for the joint model is slightly higher than the independent model, suggesting that the joint model offers a superior goodness-of-fit. The two models can also be compared with a likelihood ratio test. The probability that the adjusted likelihood ratio index between the joint and the independent models could have occurred by chance is close to zero.

Table 33. Model Measures of Goodness-of-Fit

Disaggregate measures of fit	**Model**	
	Joint model	**Independent model**
Log-likelihood value at convergence	-9,319.989	-9,939.426
Loglikelihood of null model	-22,756.385	-22,756.385
Number of parameters (not including constants)	68	62
Adjusted likelihood ratio index	0.590	0.560
Likelihood Ratio Test (LRT) between the Joint and Independent models	χ^2=-2[-9,319.989- (-9,939.426)] = 1,238.874, 6 df, p = 0.000<<0.0001	

2.7. PERCEIVED IMPORTANCE OF FACILITATORS TO CYCLING: A MULTIVARIATE ORDERED APPROACH

2.7.1. Introduction

As seen in 2.2, much research has focused on which determinants influence bike use, but few studies have focused on how different segments of individuals would perceive the implementation of policy measures aimed at encouraging more frequent bike usage. Higher urban densities, the provision of bike facilities or the presence of a bike-sharing service are considered as facilities that can affect individuals' choice to cycle. However, not all those who have access to these facilities travel by bike, as their behavior may depend on their individual characteristics or their perception of the built-environment. In fact, while one person may consider the presence of secure bike parking important, another may feel it is relatively safe to park in an open and unsheltered space.

Improved understanding of the different effects of these factors may support practitioners in developing strategies and interventions in a way that can effectively increase bike use and researchers in gaining a better knowledge of the mechanism underlying the built environment-cycling relationship. Therefore, the major goal of this study was to understand and explore how facilitators to cycling are perceived by different segments of individuals, in view of assessing how to best promote cycling in an urban area.

2.7.2. Data sample

The data for this study is drawn from the survey *Bici Mi Piaci*. The sample comprises 2128 observations. As seen in paragraph 2.3.5, all participants were asked to rate, by means of a 5 point Likert Scale, the importance of specified factors that would encourage them to start cycling or to cycle more often:

- P1. Presence of an extensive network of dedicated bike lanes
- P2. Presence of racks and secure parking for bicycles
- P3. Greater extension of the LTZ or pedestrian zones
- P4. A bike-sharing station close to home or at public transport stops
- P5. If other people cycle
- P6. Dedicated services at work / study (parking, showers, lockers for equipment, *etc.*)
- P7. Combination with public transport services
- P8. Increase in car parking fees

In our analysis, the dependent variables are the factors mentioned above, while the regressors are the individual (gender, age, level of education, type of occupation, *etc.*) and household characteristics (presence of children and number of bicycles) along with some built environment attributes (neighborhood residence characteristics and presence of bike facilities close to home).

2.7.3. Modeling framework

The eight ordinal dependent variables are jointly estimated as common unobserved factors might be present. For this reason, a multivariate ordered probit modeling methodology is adopted in this study.

Let J represent repeated measurements on n different subjects q, where each repeated ordinal observation (indexed by j ∈ J, where J = 8 in our study) is denoted by Y_{qj}. Each observable categorical outcome Y_{qj} and the unobservable latent variable $\tilde{Y}_{qj}$ are connected by:

$$Y_{qj} = r_{qj} \Leftrightarrow \vartheta_{j,r_{qj}-1} < \tilde{Y}_{qj} \leq \vartheta_{j,r_{qj}} \qquad r_{qj} \in 1,\ldots,K_j \tag{2.27}$$

where r_{qj} is a category out of K_j ordered categories (in our case $K_j = 4$) and $\vartheta_j$is a vector of threshold parameters for outcome j with the following restriction:

$$-\infty = \vartheta_{j,0} < \vartheta_{j,1} < \cdots < \vartheta_{j,K_j} = \infty \tag{2.28}$$

The threshold parameters can vary across outcome dimensions $j \in J$ in order to account for differences in the repeated measurements. Given an $n \times p$ matrix X_j of covariates for each $j \in J$, where each x_{qj} is a p-dimensional vector (q-th row of X_j) for subject q and repeated measurement j, the following linear model for the relationship between $\tilde{Y}_{qj}$and the vector of covariates $\boldsymbol{x}_{qj}$ is assumed:

$$\tilde{Y}_{qj} = \boldsymbol{x}_{qj}^T \boldsymbol{\beta}_j + \varepsilon_{qj}, \quad \varepsilon_q = \left(\varepsilon_{q1}, \varepsilon_{q2}, \ldots, \varepsilon_{q8}\right)^T \sim \; N(0, \Sigma), \tag{2.29}$$

where

- $\boldsymbol{\beta}_j$ is a vector of regression coefficients corresponding to outcome j,
- $\varepsilon_{qj}$is an error term with mean zero and multivariate normal distributed with a covariance Σ.

The regression parameters β_j can vary between the repeated measurements j. The errors are assumed to be independent across subjects and orthogonal to the covariates $\boldsymbol{x}_{ij}$. Let the actual observed measurement level for individual q and measurement variable j be m_{qj}. The likelihood function for individual q may be written as follows:

$$LL_q = Pr\left(y_{q1} = m_{q1}, y_{q2} = m_{q2}, \ldots, y_{qJ} = m_{qJ}\right) \tag{2.30}$$

$$LL_q = \int_{v_1 = \vartheta_1^{m_{q1}} - \beta_1 x_{q1}}^{\vartheta_1^{m_{q1}+1} - \beta_1 x_{q1}} \int_{v_2 = \vartheta_2^{m_{q2}} - \beta_2 x_{q2}}^{\vartheta_2^{m_{q2}+1} - \beta_2 x_{q2}} \cdots \int_{v_J = \vartheta_J^{m_{qJ}} - \beta_J x_{qJ}}^{\vartheta_J^{m_{qJ}+1} - \beta_J x_{qJ}} \phi_J\left(v_1, v_2, \ldots, v_J | \boldsymbol{R}\right) dv_1 dv_2 \ldots dv_J \tag{2.31}$$

Where ϕ_J in the above expression represents the standard multivariate normal density function and R are the off-diagonal elements of the covariance matrix $\boldsymbol{\Sigma}$. Computing the high-order I-dimensional rectangular integral in Equation (2.31) could be burdensome. However, the idea behind a recent efficient matrix-based approach, known as the univariate bivariate screening devised by Bhat (2018), has been used to compute the rectangular integral shown above and estimate parameters of the multivariate ordered response model. The mathematical formulations for the method can be found in Bhat (2018). All the parameters are estimated using the GAUSS matrix program language[3].

2.7.4. Model estimation results

Different types of variables were considered in the model specification as explanatory factors that influence the perception of various facilitators to cycling. These included individual characteristics, household characteristics and built environment attributes.

Model results are presented in Table 34. The specification of the model was based on a systematic process of removing statistically insignificant variables and combining variables when their effects were not significantly different.

2.7.4.1. Effects of Individual Attributes

The results indicate the presence of a distinct gender effect. Specifically, females are more likely to place importance on the presence of bike racks and secure parking. This result reveals that women are more concerned about bike theft than men, supporting the hypothesis of a large body of academic research that, on average, females are more concerned with safety issues and tend to avoid risky practices. Instead, an increase in car parking fees would be a stronger incentive to cycle for males. This is not surprising, given that usually women in Italy frequently need to run various errands on the way to or from work (*e.g.* pick up children from school, shopping) and a higher car parking charge might not be sufficient to stimulate use of the bicycle. For the same reason females exhibit a stronger willingness to cycle in the presence of a bike-sharing station close to home or at public transport stops. In this case cycling would limit daily activity patterns to a lesser extent and could help transport intermodality.

Highly educated individuals are, in general, less inclined to use the bike. This is in contrast with other studies (*e.g.* Bhat *et al.*, 2017) that indicate that higher education is linked to increased cycle use. One possible explanation could be that usually university graduates have a more prestigious job, where often a dress code is required, not compatible with bicycle usage. Furthermore, because of their position, they may be expected to use the car, especially in Italy,

[3] The author is grateful to Dr. Chandra R. Bhat and Gopindra S. Nair for their help in code specification. The code was made available by Dr. Bhat when the author was at UT Austin as visiting scholar.

where a strong car-centric culture exists, and its ownership is considered as a symbol of social prestige.

Individuals with a lower income are more likely to recognize the importance of the existence of an extensive bike lane network and the presence of bike parking spots. One possible explanation is that this segment of population, due to their fewer financial resources, value safe paths and secure parking much more highly, as the bicycle is a relatively more expensive possession that can be stolen or damaged due to an accident.

Non-students are more inclined than students to use the bike if other people do so, suggesting the presence of a social norm. On the other hand, students are less likely to consider higher parking fees as an important factor in the willingness to cycle. This may be due to the fact that as they have less access to a car, students are less concerned about increases in parking charges.

Interestingly, the willingness to cycle did not significantly differ across age groups, except for the outcome regarding the presence of facilities at the workplace, with younger respondents valuing them more highly than the older individuals.

Not surprisingly, the results indicate the presence of a distinct effect among people who already cycle (both for recreational and utilitarian purposes) and non-bikers. Specifically, individuals who already use the bike (compared to those who do not) are more likely to place importance on more bicycle lanes and limited traffic zones, presence of a safe parking spot and provision of facilities at workplaces.

2.7.4.2. Effects of Household Attributes

Individuals living in a household with a small number of bikes would have a greater incentive to use the bike if other people did so. This outcome implies that who already possesses a bike does not need to see other people cycling to use it. On the other hand, individuals with less bike access need to be persuaded by other cyclists to use it, suggesting the presence of a social norm component in the decisional process to use the bike. The results also indicate that individuals with no children would have a greater propensity to cycle were there dedicated facilities at the workplace.

2.7.4.3. Effects of Built Environment Attributes

Residential location choice is another important determinant. Individuals who live in suburban areas exhibit a greater willingness to cycle when cycling can be combined with public transport, both in terms of presence of bicycle lockers and covered parking facilities or bike sharing services in proximity to public transport stops. This may reflect the fact that people living in suburban areas feel that cycling is a non-competitive transport alternative but would consider intermodality if an integrated service existed. In particular, the implementation of policies supporting the bike and ride mode would render the public transport service competitive for those trips with a distance from home to the bus/train stop too long to walk but within a competitive cycling distance.

Table 34. Multivariate ordered probit model results

Variables	P1		P2		P3		P4		P5		P6		P7		P8	
	Coeff	*T-stat*	*Coeff*	*T-stat*	*Coeff*	*T-stat*	*Coeff*	*T-stat*	*Coeff*	*T-stat*	*Coeff*	*T-stat*	*Coeff*	*T-stat*	*Coeff*	*T-stat*
Bicyclist (Yes = 1; No = 0)	0.490	7.94	0.301	5.01	0.175	3.92	-0.134	-2.83	--	--	0.158	3.77	--	--	--	--
Gender (male = 1; female = 0)	--	--	-0.104	-2.05	--	--	-0.113	-2.98	--	--	--	--	--	--	0.179	3.88
Male * Bicyclist	-0.198	-3.06	-0.120	-1.58	--	--	--	--	--	--	--	--	--	--	--	--
Age	--	--	--	--	--	--	--	--	--	--	-0.897	-3.63	--	--	--	--
Bachelor's degree or higher education level	--	--	-0.154	-4.27	-0.060	-1.40	-0.264	-5.90	-0.138	-3.04	-0.219	-4.52	-0.271	-5.88	--	--
Student	--	--	--	--	--	--	--	--	-0.266	-2.83	--	--	--	--	-0.230	-2.15
Income	-0.068	-2.48	-0.064	-2.49	--	--	--	--	--	--	--	--	--	--	--	--
Presence of children (Yes = 1; No = 0)	--	--	--	--	--	--	--	--	--	--	-0.127	-2.89	--	--	--	--
Children * Bicyclist	--	--	--	--	--	--	--	--	--	--	--	--	-0.115	-2.85	-0.081	-1.71
# of bicycles in the household	--	--	--	--	--	--	-0.062	-2.95	-0.062	-2.98	--	--	--	--	--	--
Flexible working hours	--	--	--	--	--	--	--	--	--	--	0.139	2.82	--	--	--	--
Presence of bike lanes within 400m of home	--	--	-0.067	-2.00	--	--	-0.074	-2.23	--	--	--	--	-0.074	-2.23	--	--
Urban	--	--	--	--	--	--	-0.152	-3.18	--	--	-0.080	-1.59	-0.077	-1.47	--	--
Threshold																
Threshold1	-1.954	-21.20	-2.150	-22.18	-1.342	-24.69	-1.892	-24.52	-0.952	-17.78	-2.137	-14.72	-1.879	-25.27	-0.155	-3.38
Threshold2	-1.499	-18.41	-1.588	-19.20	-0.883	-18.62	-1.372	-19.42	-0.456	-9.04	-1.675	-11.85	-1.372	-19.85	0.291	6.38
Threshold3	-0.946	-12.32	-1.037	-13.05	-0.253	-5.88	-0.821	-12.41	0.122	2.47	-1.078	-7.75	-0.842	-12.87	0.892	18.83
Threshold4	-0.357	-4.80	-0.384	-4.98	0.300	6.81	-0.261	-4.01	0.650	12.69	-0.452	-3.29	-0.235	-3.66	1.266	24.46
Correlation terms																
P1	n/a	n/a														
P2	0.829	83.38	n/a	n/a												
P3	0.561	27.78	0.564	29.52	n/a	n/a										
P4	0.511	25.13	0.600	33.61	0.513	29.16	n/a	n/a								
P5	0.198	7.42	0.241	9.25	0.198	8.65	0.314	14.40	n/a	n/a						
P6	0.461	20.99	0.522	26.76	0.360	16.96	0.486	25.87	0.376	17.34	n/a	n/a				
P7	0.481	22.57	0.508	25.58	0.392	18.77	0.637	42.12	0.347	15.58	0.587	35.02	n/a	n/a		
P8	0.155	5.39	0.174	6.09	0.282	11.28	0.280	11.01	0.294	13.26	0.194	7.20	0.249	9.60	n/a	n/a

2.7.4.4. *Model Goodness of fit*

Model goodness-of-fit measures are shown in Table 35. In addition to the joint model, an independent ordered probit model system was estimated by setting all correlation terms to zero. The performance of the joint model was then assessed comparing goodness-of-fit metrics. The value of the adjusted likelihood ratio index for the joint model is slightly higher than that for the independent model, suggesting that the joint model offers a superior goodness-of-fit than the independent model. The χ^2 test statistic of the likelihood ratio test (LRT) between the joint and independent models is statistically significant at any degree of confidence.

All the error correlation terms are significant (Table 34), suggesting that a multivariate ordered response model that accommodates error correlations is appropriate in this particular context. The error correlation terms indicate the presence of significant unobserved attributes that simultaneously affect the dependent variables considered in the study. For example, it is found that the error correlation for a better bicycle network and the existence of more bike parking spots is positive and significant (0.829 with a t-stat of 83.38). This indicates that unobserved attributes, that contribute to increasing the latent utility of the importance of the presence of an extended bicycle network, are positively correlated with unobserved attributes that contribute to the latent utility of the importance of more safe parking spots. One reason for this correlation may be the existence of some unobserved behavior traits, such as safety concerns about road obstacles and theft, that could impact the willingness to use the bike.

Table 35. Model Goodness of fit

Goodness of fit measures	*Joint model*	*Independent model*
Log likelihood at convergence	-20,630.45	-23,004.74
Log likelihood at constants	-23,190.73	-23,190.73
Number of parameters	95	67
Adjusted likelihood ratio index	0.1063	0.0051
Likelihood Ratio Test (LRT) between the Joint and Independent models	χ^2=-2[-23,004.74 - (-20,630.45)] = 4,748.58, 28 df, p = 0.000	

PART 3 - EVALUATION OVER TIME OF TRAVEL DEMAND MANAGEMENT STRATEGIES

3.1.INTRODUCTION

To reduce car dependence in favor of more sustainable alternative modes of travel requires promoting appropriate policies that ensure an efficient organization of the transportation system. Public transport has a major role to play in reducing private car use. The introduction of new routes for improving accessibility, density, and frequency of public transport services, so as to be able to compete effectively with the private car, may be one solution for reducing vehicle kilometers traveled. Service quality is perceived as an important determinant of users' travel demand.

Therefore, analyzing travel behavior is not a simple task for car use (Sanjust *et al.*, 2015), as it embraces individuals' attitudes, perceptions, habits, and knowledge. The so-called voluntary travel behavior change (VTBC) programs were conceived to take these aspects into account (Ampt, 2003). These programs use information and communication tools to encourage people to change their travel behavior, and their implementation has proven to be effective in reducing private car use (Brög *et al.*, 2009). Analysis of several projects conducted over the past 20 years has pinpointed the strengths and weaknesses of these measures (see 1.2.3). In particular, information provision is just one of the elements of the process, and alone, without any intervention in the choice context, it may well be ineffective. For these reasons, the implementation of an integrated system of structural and information measures could yield better results and enhance the effect of each measure.

Another aspect that warrants further investigation concerns the long-term effects, given that individuals' socio-economic (SE) characteristics may change over time. and consequently, their psycho-attitudinal and motivational characteristics, which are known to strongly influence travel behavior and which soft measures focus upon. Some authors have found that behavior changes persist and may even increase over time (Taylor and Ampt, 2003) if appropriately monitored and reinforced.

Addressing the question of when soft policy measures are effective, the development of useful programs requires longitudinal panel studies that examine behavioral changes (Richter *et al.*, 2011). Longitudinal data are also required for constructing integrated choice latent variable (ICLV) models. As explained in 1.3.9.2 these models do not support the derivation of policies that aim to change travel behavior by simply changing the value of a latent variable, so carefully designed (longitudinal) experiments are needed to provide a more solid foundation for deriving policy implications related in particular to latent variables in ICLVs (Chorus and Kroesen, 2014).

Notwithstanding this requirement, the majority of recent papers infer policy implications that are not adequately supported by the data used for ICLV estimation, whereas others, despite having longitudinal data at hand, simply employ structural equations (Bamberg *et al.*, 2003; Jariyasunant *et al.*, 2015).

Given this background, the aim of the first part of this work is to analyze the short-term effect on travel mode choice of introducing a new sustainable form of transport into the choice set (hard measure) when implementing a VTBC program (soft measure). The transport context chosen for this experiment is a corridor linking the city center of Cagliari (Italy) to a university/hospital complex, where a new light rail route went into service in February 2015. A survey called "Cittadella Mobility Styles" was designed. It included the new light rail line in the choice set of a large-scale VTBC program conducted in the metropolitan area of Cagliari. In particular, an attempt was made to overcome the critical issue concerning evaluation of the measure by creating a control group, so as to disentangle the effect of the structural measure from that of personalized information provision.

In the second part of the research we evaluate, on the one hand, the long-term effects on travel mode choice of the implementation of a new light rail line, on the other we investigate if any changes in the psycho-attitudinal factors and/or in socio-economic characteristics exist after implementation of those measures. In particular, the objective of the study is to analyze whether these changes in individual characteristics are able to affect mode choice from a modeling perspective, through the specification and estimation of hybrid choice models that use, for the same sample, the data collected for these two moments in time. This second part attempts to provide a contribution to the research, analyzing, from a statistical and modeling perspective, the evolution of travel behavior over time as well as individuals' intrinsic characteristics (socioeconomic and attitudinal) following a change in the context characteristics. The results of model estimation were not intended to derive policy implications, but rather to understand whether the criticism raised about these models can actually be supported, for the case at hand, by a scientific result.

The remainder of the chapter is organized as follows. Section 3.2 provides a review of TDM strategies on university campuses. Section 3.3 describes the study context and data collection. Section 3.4 presents the results of behavioral change a few months after the implementation of the new light rail line. Section 3.5 presents an exploratory analysis of the data gathered in the third wave survey. In the same section the ICLV model devised is presented, followed by discussion of the model results.

3.2. TRAVEL DEMAND MANAGEMENT STRATEGIES ON UNIVERSITY CAMPUSES

This section provides a review of the travel demand management measures implemented in different university campuses worldwide.

One of the most widely implemented TDM strategies at universities campus is parking management. Two main approaches have been adopted: political and economic (Shoup, 2008). The political approach relies upon rules and regulations (*e.g.* the restriction of the number of parking spots), while the economic approach is based on pricing parking at market value. Different campuses in the USA are restricting parking permits only to people who do not have a viable alternative mode of transportation to the university (Isler *et al.*, 2005). Other campuses are stopping providing free parking spaces (Balsas, 2003; Isler *et al.*, 2005). Barla *et al.* (2012) evaluated the potential for reducing the commute mode share of cars at Université Laval in Quebec City (Canada) using stated preference data. They found that the cost of parking diminishes the probability of commuting by car, with clear differences across professional status and income groups. In the context of the University of Idaho, USA, Delmelle and Delmelle (2012) reported that increasing the price of parking for students on the campus is a disincentive to driving. Rotaris and Danielis (2014) tested different hypothetical transport policies at the University of Trieste, in Italy, and highlight three policies that lead to a decrease in car use and are also considered as socially and economically efficient: subsidizing bus fares, a mix of bus subsidies with parking restrictions and both increased parking prices and restrictions. Cruz *et al.* (2017) found that the effective control of illegal parking and the banning of this practice could help to reduce the number of cars in the university campus area of Coimbra, Portugal, by approximately 10%.

Other policies involve the implementation of rideshare programs, such as carpooling incentives. Aoun *et al.* (2013) proposed for the American University of Beirut, Lebanon, a dynamic taxi-sharing service which combines the higher vehicle occupancy of a shared taxi with the reliability and comfort of a private taxi at the reduced cost of a public transport fare. Erdoğan *et al.* (2015) indicated that providing priority parking and cheaper parking options to rideshare program members would help to increase interest in ridesharing and discourage single occupancy vehicle trips.

Often universities work in collaboration with transit agencies to offer students special discounts or free transit passes. These programs have different potential benefits, such as increased student transit ridership, reduced demand for campus parking, use of off-peak transit capacity and improved transit agency performance (Yu and Beimborn 2018). In USA more than 50 colleges and universities provides fare-free transit for over 800,000 people (Brown *et al.*, 2001; Brown *et al.*, 2003; Han *et al.*, 2019). Another example is the U-Pass or Reduced-Fare Ticket measures implemented in Canada and Germany. Upon enrolment students, have the possibility of paying a sum additional to their tuition fee, to allow them to use all the public transport lines involved in the agreement for free. For example, Letarte *et al.* (2016) showed an increase of public transit share (+18%) and a decrease of car share at the University of Ottawa, Canada, after the program was launched. In Germany, such arrangements benefit one third of the 1.9 million German students (De Witte, 2006).

Some studies have focused on providing information on available transport alternatives. Rose (2008) analyzed the effects of a voluntary travel behavior change program targeted incoming first-year students at the Clayton Campus of Monash University in Melbourne, Australia. He found a significant effect of the program in terms of reducing single occupant commuting and increasing public transport use (up 5.9%).

In the field of soft measures, the possibility of using new technologies to influence students' travel behavior habits is gaining increasing attention. Jariyasunant *et al.* (2015) designed a computational travel feedback system, Quantified Traveler, in which feedback about movements (carbon emissions, calories burned, travel time and cost) is used to change travelers' mode or trip choice. In an experiment conducted at the University of California Berkley, USA, they found a statistically significant decrease in the average distance driven, the average reduction being 39 kilometers or 33% lower than the first week. Di Dio *et al.* (2015) developed a smartphone app called "TrafficO_2" that aims to nudge commuters towards more sustainable mobility by providing monetary incentives for each responsible choice. In a test conducted with a selected sample of students at Palermo University, Italy, they observed a reduction by almost half of the carbon dioxide equivalent emissions when compared to their previous habits.

Only a few studies have investigated changes in cognitive factors after the implementation of TDM measures at Universities. Heath and Gifford (2002) used an expanded version of the TPB to evaluate the change after implementation of U-pass program in the university of British Columbia, Canada. Similarly, Bamberg *et al.* (2003) investigated the changes in the constructs of the TPB after introduction of a prepaid bus ticket at the University of Giessen in Germany. Jariyasunant *et al.* (2015) evaluated the efficacy of the Quantified Travel platform not only in terms of travel behavior, but also in terms of change in the cognitive factors.

3.3. STUDY CONTEXT AND DATA COLLECTION

The transport context chosen for this study is a corridor linking the city center of Cagliari (Italy) to a university/hospital complex (Cittadella Universitaria), where in February 2015 a new light rail route (METROCAGLIARI) went into service. Unexpectedly, in September 2015, the public transport agency of the metropolitan area of Cagliari introduced a new bus route ("University Express"), that connects one of Cagliari's largest residential areas with the complex. Note that the light rail line serves a fairly short corridor, and for some people is more convenient, in terms of travel time, to use the bus because of its more extensive network. The Cittadella is a major trip attractor, thus a large number of people could be intercepted. The number of people potentially attracted daily to the Cittadella is just over 10,200: 1,784 university and hospital employees (17.5%), 7,872 students (77.2%) and 580 (5.7%) for hospital admissions, medical examinations, patients' visitors, *etc.*

The program, called Cittadella Mobility Style, comprised four macro-phases:

- first wave survey: a preliminary survey to capture the travel patterns of individuals traveling the Cittadella;
- Personalized Travel Plan (PTP): creation and delivery of a PTP containing suggestions for a sustainable travel alternative to the Cittadella (designed on the basis of the information collected with the first wave survey);
- second wave survey: assessment of the effectiveness of the measure and monitoring by administering a second questionnaire after deliver of the PTP;
- third wave survey: conducted two years later the second one to evaluate the effectiveness of the measure in the long-term effects and to detect any change in individuals' characteristics

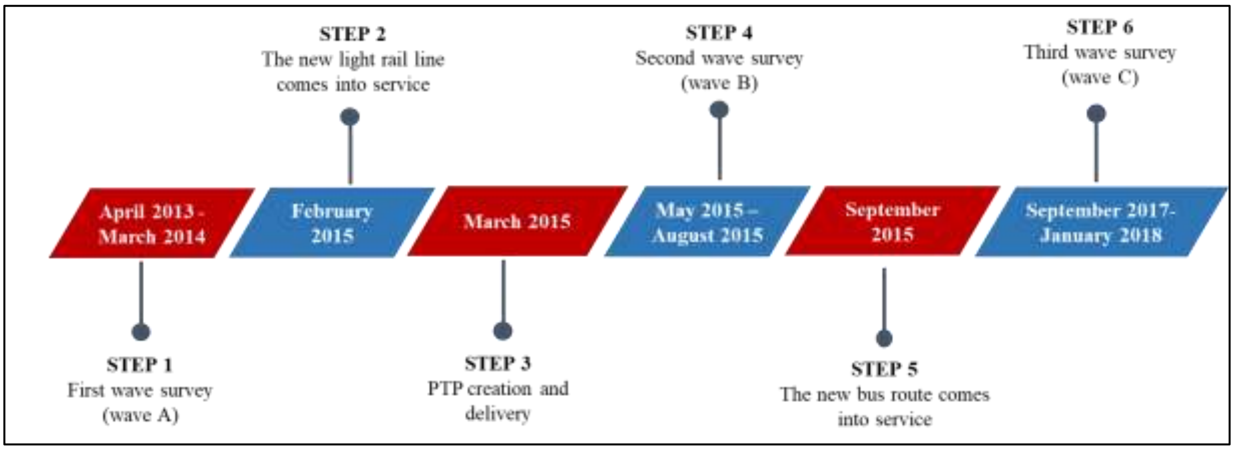


Figure 9. The program Cittadella Mobility Styles

3.3.1. First wave survey

The first survey was implemented in April 2013. The aim was to intercept as many university students, university and hospital staff and visitors to the hospital for medical examinations, admissions, *etc.*, as possible. Potential candidates were contacted both via mailing lists provided by the university and hospital (8,847 invitation mails were sent) and through a promotional campaign conducted via traditional communication channels (postcards, press conferences, TV and daily newspapers) and social media, inviting them to complete an on-line questionnaire hosted on the project's web site (the web site recorded 6,402 visits during the promotional campaign period). The possibility of winning an iPad or a €100 gift voucher was offered as an incentive to filling complete the questionnaire.

Figure 10. Promotional postcard

The objective of the first wave survey was to gather detailed information about commuters and individuals travelling occasionally to the Cittadella, so as to be able to define, with a high degree of personalisation, the information to be presented in the PTP delivered to the selected target. In particular, the questions concerned:

- Trip description:
    - trip origin and frequency
    - mode used
    - trip characteristics (walking time, waiting time, time aboard, *etc.*)
    - fare
- Personal perceptions, propensities, attitudes, beliefs:
    - attachment to car
    - dislike of public transport
    - propensity to use the light rail
    - pro-environmental behavior
    - environmental awareness
    - level of confidence with information systems
- Personal Information:
    - age
    - gender
    - education

- occupation
- income
- household structure
- car ownership

Questions concerning personal perception, propensities and attitudes were identified as follows. We selected two focus groups, in order to understand the peculiarities, benefits/disadvantages, motivation and limits associated with use of the private car and of public transport in general. The first group was made up of habitual public transport users, the second of car drivers. Both groups were invited to discuss identical topics, so as to capture the differences associated with diverse travel habits. Analysis of the results enabled to identify, under the supervision of a team of environmental psychologists, 36 items, measured by means of the 5-point Likert scale.

The questionnaire was filled in by 2,886 individuals; 2,163 questionnaires were completed (74.9%) and 723 partially completed (25.1%). Regarding individual characteristics, the analysis revealed a larger percentage of females than males (59.4% vs 40.6%). In terms of age distribution, the sample was composed as follows 18-30 (71.4%), 31-40 (10.3%), 41-80 (18.3%). For age and gender, the intercepted sample is representative of the statistical population. The majority are highly-educated (68.4% had a high school education, 28.3% an undergraduate or higher degree). The sample consists mainly of students (69.4%) who do not have children (86.1%). Students living away from home account for 36.2% of the sample. Average number of household members is 2.4. As for personal monthly income, 61.9% stated they earned less than € 1,000 a month 25.1% € 1.000-2.000, 10.5% € 2.000-3.000, 2.5% >€ 3.000. Analysis of trip purpose revealed a prevalence of trips for study purposes (69.5%), followed by work (18.0%) and trips to the hospital for medical treatment or visits to patients (9.9%). Comparison of these percentages with the statistical population shows them to match perfectly those detected at an aggregate level. The most widely used mode of travel stated by respondents was public transport with a percentage of around 52.7% (which reflects the large number of students in the sample), followed by car as driver, 37.7%, car as passenger 7.7% and motorbike/moped 2%. Cross analysis of travel mode and trip purpose shows that the majority of work trips are made by car as driver (75.6%). A similar situation is observed for hospital trips for medical treatment or visits to patients, car users accounting for 43.2% and 55.1% respectively. The majority of students (66.9%) travel by public transport, as they are practically obliged to do so, 74.3% stating they had no alternative means of transport.

Table 36. Data collected in the first wave survey and socio-economic characteristics

Variables	**Total**	**By car as driver**	**By car as passenger**	**By motorbike**	**By bus**
Questionnaires completed	2,163	816 (37.7%)	166 (7.7%)	42 (1.9%)	1,139 (52.7%)
Socio-economic characteristics					
Male	40.6%	49.4%	30.7%	73.8%	34.5%
Age					
18-30	71.4%	47.1%	71.1%	47.6%	89.7%
31-40	10.3%	17.5%	7.8%	14.3%	5.3%
41-80	18.3%	35.4%	21.1%	38.1%	5.0%
Level of education					
Low (High school and lower)	71.7%	56.6%	74.7%	64.3%	82.2%
Medium (Undergraduate and Master degree)	20.2%	26.5%	18.7%	19.0%	16.0%
High (Higher than Master degree)	8.1%	16.9%	6.6%	16.7%	1.8%
Employment status					
Student	69.4%	45.0%	68.1%	40.5%	88.2%
Employed	28.1%	52.8%	22.9%	57.1%	10.0%
Unemployed or retired	2.5%	2.2%	9.0%	2.4%	1.8%
Children	13.9%	27.6%	15.7%	26.2%	3.4%
Monthly personal income					
Low (<1,000 euro)	61.9%	43.8%	72.3%	42.8%	74.1%
Medium (1,000-2,000 euro)	25.1%	35.7%	19.9%	28.6%	18.1%
High (>2,000 euro)	13.0%	20.6%	7.8%	28.6%	7.8%
Trip characteristics					
Purpose					
Study	69.5%	48.4%	59.0%	47.6%	87.0%
Work	18.0%	36.3%	9.7%	40.5%	5.3%
Medical treatment	6.3%	7.5%	16.3%	4.8%	4.1%
Visiting patients	3.6%	5.4%	7.8%	0.0%	1.8%
Other	2.6%	2.4%	7.2%	7.1%	1.8%
Travel distance (average) [km]	-	10.72	14.04	7.62	9.2
Travel time aboard (average) [min]	-	17.68	19.04	11.93	27.54
Walking time from origin to PT-stop (average - bus only) [min]	-	-	-	-	5.19
Number of changes (bus only)					
0	-	-	-	-	80.1%
1	-	-	-	-	17.2%
2	-	-	-	-	2.7%

3.3.2. The PTP

The PTP is one of the key components of a VTBC program. The choice to opt for a personalized approach was based on the findings of several works that demonstrated personalized information to be more effective in evoking travel behavior change than mass communication, as car drivers cannot simply ignore it (Gärling and Fujii, 2009).

Analysis of the questionnaire responses, particularly place of residence and means of transportation used to reach the Cittadella, made it possible to identify the target population. We considered as potential users all car drivers living within the metropolitan area of Cagliari, along the light rail corridor. These numbered 544 out of a total of 2,163 car drivers. The PTP were not delivered to all 544 car drivers identified, as a sub-sample of 20% (109 users) was randomly chosen as a control group for evaluating the effectiveness of the VTBC program. In this way it was possible to disentangle the effect of the information (soft) measure from the structural (hard) measure and *vice versa* and at the same time assess the combined effect of the hard + soft measures.

The travel alternative was created on the basis of the travel patterns reported by respondents to the preliminary questionnaire. We analyzed all the travel information and quantitative feedback on weekly use of the car, distance traveled and journey time. To create the PTP using the light railway as an alternative means of travel, we took into account the departure station, walking time and distance to the nearest station, waiting time at the station, journey time and distance and fare paid. The alternatives were processed using Citilabs CUBE software.

Special attention was devoted to the graphics of the PTP and to the information contained therein as it has been demonstrated (Gaker and Walker, 2011) that accuracy is important, and that the information needs to be conveyed in a manner that is easily understood.

The PTP contained the following information (Figure 11):

- a practical map showing the route to reach the light railway station;
- a detailed description of the actual individual and environmental effects of the travel behavior adopted. In fact, the PTPs presented a daily, weekly and annual evaluation of the key factors that come into play in travel choice: monetary cost, journey time, calories burned and CO_2 emissions. We also indicated the disadvantages associated with car use with respect to the light rail, so as to offer new viewpoints and food for thought regarding mobility and its effects, with a view to evoking travel behavior change;
- personalized slogans and other useful information on sustainable travel in general and specifically on use of the light railway;
- links that provide useful information on mobility.

The PTP was e-mailed to the participants at the address indicated in the preliminary survey. The information contained therein could be read from any mobile device (pc, tablet, smartphone), so that participants always had the information on hand.

The cost of the marketing campaign, staff and creation of the PTP was around € 30,000.00.

Figure 11. The Personalized Travel Plan

3.3.3. Second wave survey

Once the light rail service had become operational (February 2015), PTPs were delivered to the selected target. A few months later a second wave survey was implemented to give participants time to process the information and try the proposed alternative. The aim of the second survey was to:

- monitor users’ behavior after PTP delivery and the introduction of the light rail service;
- obtain feedback on the design and layout of the PTP and on the information contained;
- collect the necessary information for modeling the phenomenon.

The questions in the second wave survey concerned:

- Trip description:
  - trip origin and frequency
  - mode used
  - trip characteristics (walking time, waiting time, time aboard, *etc.*)
  - fare
- The light rail service:
  - level of knowledge about the light rail service
  - motivation for using or not using the light rail service

- Information about PTPs:
    - design and layout of the PTP
    - quality of information provided
    - quantity of information provided

The second questionnaire was sent to all individuals who accurately completed the first one, not simply to the potential light rail users in the target. The reason for this was not to preclude the possibility of analysing the behavior of those users who were not included in the target. So, the results of the second wave survey were analyzed distinguishing three different categories:

- PTP: participants who received the PTP;
- CG: participants who formed part of the control group;
- No PTP: participants who did not receive the PTP and were not included in the CG. These were users of bus services or car drivers who did not benefit from using the light rail service as they did not live along the corridor served by light railway.

Fully completed questionnaires numbered 740, 34.2% of the total sent. Unfortunately, with surveys comprising different phases spread over time, a large number of the initial participants may fall by the wayside. On the other hand, in this way it is possible to collect detailed information from those who made the effort to fill in both questionnaires, and to examine the phenomenon in depth.

Analysis of the socio-economic characteristics of respondents to the second wave survey, limited to those who traveled to the Cittadella at least once (516 individuals), revealed that the sample reproduces both the first wave survey sample and the population gravitating towards the Cittadella. Table 37 summarizes individuals' socio-economic characteristics, split into the three categories mentioned above.

Table 37. Data collected in the second wave survey

Variables	PTP	CG	No PTP	Total
Questionnaires sent	435	109	1,619	2,163
Questionnaires completed	169 (38.9%)	45 (41.3%)	526 (32.5%)	740 (34.2%)
Participants traveling to the Cittadella	113 (66.9%)	29 (64.4%)	374 (71.1%)	516 (69.7%)
Socio-economic characteristics				
Male	42.5%	41.4%	36.4%	38.0%
Age				
18–30	39.8%	48.3%	78.6%	68.4%
31–40	17.7%	17.2%	9.4%	11.6%
41–80	42.5%	34.5%	12.0%	20.0%
Level of education				
Low (High school and lower)	44.3%	51.8%	78.3%	69.4%
Medium (Undergraduate and Master degree)	32.7%	24.1%	16.1%	20.2%
High (Higher than Master degree)	23.0%	24.1%	5.6%	10.4%
Employment status				
Student	38.1%	37.9%	78.3%	67.2%
Employed	59.3%	62.1%	19.5%	30.6%
Unemployed or retired	2.7%	0.0%	2.1%	2.1%
Children	27.4%	31.0%	8.8%	14.1%
Monthly personal income				
Low (<1,000 euro)	42.5%	48.3%	65.2%	59.3%
Medium (1,000-2,000 euro)	34.5%	34.5%	24.1%	26.9%
High (>2,000 euro)	23.0%	17.2%	10.7%	13.8%

3.3.4. Third wave survey

The third wave was conducted two years after the second, following the introduction of the new bus service and light rail to assess 1) whether the measures had long term effects and 2) to detect any change in individuals' SE and psycho-attitudinal characteristics.

This survey was designed such that the data collected were perfectly comparable with those gathered in the first wave and included the same questions that appeared in the first questionnaire. However, having analyzed the responses to the second wave, it was deemed useful to include some questions about long term travel behavior. As in the two years that had passed, some of the participants might no longer travel to the Cittadella (students had graduated, contracts finished, *etc.*) the participants were asked whether they used the light rail to travel to other destinations/for other reasons and the trip characteristics (frequency, purpose, *etc.*). The third questionnaire was also e-mailed to all 2,163 individuals who accurately completed the first one, 522 (24.1%) questionnaires were completed and 464 (88.9%) of respondents traveled to the Cittadella. Table 38 summarizes individuals' socio-economic characteristics, split into the three categories mentioned above.

Table 38. Data collected in the third wave survey

Variables	**PTP**	**CG**	**No PTP**	**Total**
Questionnaires sent	435	109	1,619	2,163
Participants traveling to the Cittadella	115 (26.4%)	28 (25.7%)	321 (19.8%)	464 (21.5%)
Socio-economic characteristics				
Male	48.7%	35.7%	37.7%	40.3%
Age				
18–30	26.1%	32.1%	62.0%	51.3%
31–40	17.4%	14.3%	16.5%	16.6%
41–80	56.5%	53.6%	21.5%	32.1%
Level of education				
Low (High school and lower)	31.3%	35.7%	39.3%	37.1%
Medium (Undergraduate and Master degree)	32.2%	35.7%	50.5%	45.0%
High (Higher than Master degree)	36.5%	28.6%	10.3%	17.9%
Employment status				
Student	16.5%	14.3%	42.7%	34.5%
Employed	80.9%	82.1%	48.3%	58.4%
Unemployed or retired	2.6%	3.6%	9.0%	7.1%
Children	34.8%	35.7%	14.0%	20.5%
Monthly personal income				
Low (<1,000 euro)	24.3%	32.1%	57.6%	47.8%
Medium (1,000-2,000 euro)	47.8%	57.1%	32.4%	37.7%
High (>2,000 euro)	27.8%	10.7%	10.0%	14.4%

3.4. ANALYSIS OF SECOND WAVE SURVEY

Analysis of the results was limited to those users who participated in the first wave survey and reported traveling to the Cittadella at least once (516 individuals), the reason being that only these individuals could be compared to understand whether they had changed their travel habits after implementation of the hard and soft measures. Note that the hard measure involved all individuals (PTP, CG and No PTP), whereas the effect of the soft measure is divided into two categories: 1) one related to mass communication and advertising campaign of the light rail service, which involved all individuals 2) the second related to the personalized measure represented by the PTP, which involved all those who had received and read the PTP. Hence it is worth noting that all the individuals, in addition to the hard measure, were intercepted by a generalized soft measure which could have had an effect, though not quantifiable, on travel behavior.

3.4.1. PTP results

Note that 54 out of the 113 participants (47.8%) in the PTP group who had traveled to the Cittadella at least once after the introduction of the light rail service, stated they had not received/read the PTP. Hence, from now on, they will be included in the control group as they have the same characteristics (they belonged to the potential light rail users' target but did not receive/read any personalized information).

The overall effect of evaluating the measure was not distorted by this assumption as not having received/read the PTP neutralizes the effect of the personalized measure. Thus, these individuals are only liable to be influenced by mass communication and the hard measure. Treating these individuals as PTP recipients would, conversely, have resulted in erroneously evaluating the effect of the personalized measure as it was not actually delivered.

It was decided to send the PTP via e-mail as most people are now familiar with modern technology and it enabled the measure to be implemented on a large scale. Moreover, as 65% of the sample of the first wave survey possessed a smartphone, the PTP was always on hand making it possible to test the proposed alternative without any prearrangement. However, this was not sufficient to lead people to read the PTP: direct contact, though costlier and affecting sample size, seems to still be the best way to attract people's interest.

Special attention was devoted to assessing the effectiveness of the PTP both in terms of design and layout and quality of the information contained. In particular, analysis of the responses of participants who read the PTP, revealed that the information was well-organized (97%), not excessive (85%) and the colours used facilitated reading (95%). The suggested alternative was explained clearly and links to the maps were considered useful for understanding it. Therefore, we can reasonably claim that the objective of providing participants with a graphically pleasing personalized plan and a clear description of the suggested alternative has been achieved.

As far as the PTP is concerned, the respondents stated that it also provided information on aspects they were unaware of or unable to quantify such as CO_2 emissions and kcal burned (75%), so they were able to appreciate the benefits of using the light rail (71%). This shows that the PTP was successful in delivering precise information about the alternative available to car drivers: in fact, 65% stated that the PTP provided them with information that enabled them to make a decision.

Analysis of the second wave questionnaire indicated that 47.5% had tried the travel alternative suggested in the PTP, 28.8% stated they would try it in the future while 23.7% stated they had not tried it and did not intend to do so in the future.

3.4.2. Behavior Change Analysis

To evaluate any changes in travel behavior we analyzed the modal share observed in the second wave survey (Figure 12 a). First of all, 49% of respondents had chosen to use the light railway, of whom 41% already traveled by public transport, 7% were car drivers and 1% used other means of transport.

The diagram (b) of Figure 12 quantifies the effect of each measure implemented on the switch from car driver to light rail user. Roughly half (46%) the participants who read the PTP (27 individuals) changed their travel behavior against just 34% of the control group, that also included those who had not received/read the PTP, demonstrating the greater effectiveness of a personalized measure compared to mass communication. Another significant aspect is the duration of the use of the light rail: 27.9% who read the PTP and 21.7% of control group became a frequent user (the ratio of light rail use frequency to total frequency was greater than 0.5). Considering this result, we can reasonably claim that, all other conditions being equal (hard measure and mass communication), the personalized measure can contribute by 12% to behavior change for those participants in the potential light rail user target. This is in line with the findings of the few studies reported in the literature that quantify the contribution of VTBC programs to behavior change as between 5 and 15% (see 1.2.3.1). However, it should be stressed that in these cases the effects of behavior change may well have been overestimated as monitoring was limited to a very short period of time.

Nevertheless, it is important to note that mass communication and introduction of a new line of light service also played an important role in travel behavior change.

A moderate percentage (23%) of participants who were not included in the potential light rail user target (No PTP) also changed their behavior, but the percentage is lower than the other two cases, as was to be expected.

Lastly, we evaluated which information contained in the PTP had the greatest effect on travel behavior change. Respondents' answers were measured on a 5-point Likert scale. Participants considered all the information provided important for deciding whether to change their travel behavior: CO_2 emissions (3.67 AVG), journey time (3.56 AVG), cost (3.44 AVG), calories burned (3.26 AVG) in that order. Though the results for journey time and costs were predictable, surprisingly CO_2 emissions were considered the most important. As mentioned previously, people are usually unaware of this aspect or are unable to quantify emissions. This information was considered significant by 51.7% of those who changed their travel behavior.

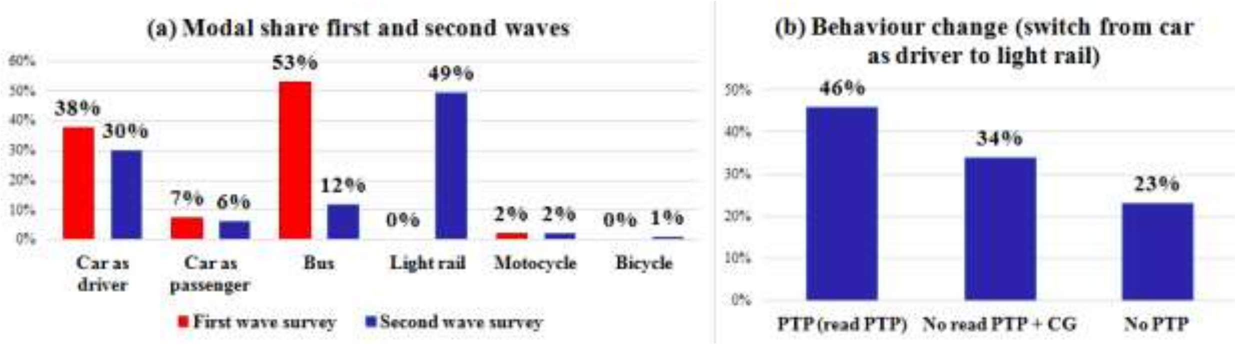


Figure 12 a) Comparison of modal share first and second survey waves, and (b) behavior change (switch from car as driver to light rail).

3.5.ANALYSIS OF THIRD WAVE SURVEY

Only a total of 350 people participated in all three surveys. However only 61% (215 individuals) who traveled to the Cittadella were included in the sample analyzed for assessing travel behavior. The sample analyzed for studying psycho-attitudinal factors and for hybrid model estimation comprised instead 149 individuals, as out of the 215 identified, we eliminated those who could not use at least two of the alternative travel modes considered (car, bus and light-rail). Thus, all those included in the final sample had access to a car for the commute to the Cittadella (Table 39). Note that, since the sample is now composed by only 149 individuals, it is not be possible to assess the effect of the information measure. In fact, the paucity of responses did not allow to perform significant analysis regarding this topic.

Table 39. Data collection

	First survey wave (wave A)	**Second survey wave (wave B)**	**Third survey wave (wave C)**
People contacted	8,847	2,163	2,163
Questionnaires completed	2,163 (24.4%)	740 (34.2%)	522 (24.1%)
Participants traveling to the Cittadella	2,163 (24.4%)	516 (23.8%)	464 (21.4%)

Individuals who participated in all three waves 350 (16.1%)

Individuals who traveled to Cittadella 215 (9.9%)

Individuals who had available at least two alternatives among car, bus and light rail 149 (6.9%)

Table 40 provides an overview of the socio-economic variables and their descriptive characteristics in the first (2013) and third wave (2017). Socio-economic characteristics included gender, age, occupation, educational level, household composition and household income.

Clearly the number of students decreased significantly in the third wave, due to the fact that the majority of them had graduated (percentage of graduates passed from 26.8% to 40.3%) and had started work. Consequently, respondents’ levels of income have also increased, and as a result we also observe an increase in car ownership in the third wave, in line with other studies (Dargay, 2007; Oakil *et al.*, 2016). Moreover, some participants no longer live with their parents, proven by a slight decrease in the average number of household members.

Table 40. Socio-economic characteristics between first wave and third wave survey

Variables	First wave (wave A)			Third wave (wave C)		
	N.	*%*	*AVG*	*N.*	*%*	*AVG*
Tot	149	-	-	149	-	-
Gender (male)	66	44.3%	-	66	44.3%	-
Age			35.1			39.1
Age 18_30	73	49.0%	-	60	40.3%	-
Age 31_40	23	15.4%	-	30	20.1%	-
Age 41_60	51	34.2%	-	48	32.2%	-
Over 60	2	1.3%	-	11	7.4%	-
Employment status						
Student	71	47.7%	-	38	25.5%	-
Employed	76	51.0%	-	102	68.5%	-
Unemployed or retired	2	1.3%	-	9	6.0%	-
Level of education						
Low (High school and lower)	81	54.4%	-	49	32.9%	-
Medium (graduates or master degree)	40	26.8%	-	60	40.3%	-
High (Higher than master degree)	8	5.4%	-	40	26.8%	-
Number of household members	-	-	2.94	-	-	2.8
Children	35	23.5%	-	33	22.1%	-
Own car	118	79.2%	-	138	92.6%	-
Number of cars per household	-	-	1.9	-	-	1.8
Personal income per month						
Income € 0 - 1,000	76	51.0%	-	59	39.6%	-
Income € 1,000 - 2,000	50	33.6%	-	66	44.3%	-
Income € 2,000 - 4,000	18	12.1%	-	21	14.1%	-
Income > € 4,000	5	3.4%	-	3	2.0%	-

Note that 26.1 % of the sample changed trip origins, and the number people who moved from a location close to or distant from the metro corridor and *viceversa* is equal to 9.3%.

Analyzing the responses to the third questionnaire, it was possible to examine travel behavior, and hence modal share, of all the respondents who traveled to the Cittadella following implementation of the measures. Figure 13 shows the modal share for the first and third waves. Generally, there has been a slight decrease (-6.0%) in the use of the private car for traveling to the

Cittadella between the two waves. Interestingly there is a sharp drop in the number of people using the bus and an increase in travelers by light rail. This is in line with the findings of other works, which argue that a rail transit service is able to attract significantly more passengers than an express bus service.

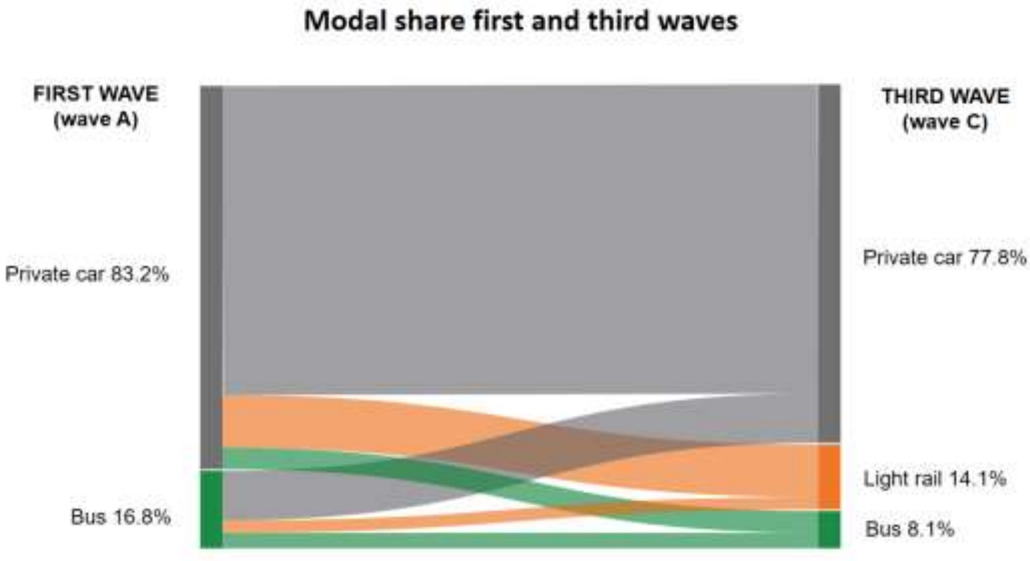


Figure 13. Modal share first and third waves

Examining the data in detail, we observe that 16.1% individuals changed from private car to bus and light rail, but 10.7% changed from bus to private car. This increase in private car use could be attributed to a change in the SE characteristics of the sample, especially the average increase in income and a higher level of car ownership (note that in the meantime many students had received research grants for furthering their education or for medical specialisation or are now working) which are certainly two determinants of travel mode choice.

3.5.1. Evolution of psycho-attitudinal factors

One important aspect of the study concerned the definition and subsequent evaluation, in two different moments of time, wave A and wave C, of those psycho-attitudinal factors that are able to impact on travel behavior and/or *viceversa*. Table 41 shows the questions asked, along with summary statistics of the responses in both waves of the survey. T-tests were used to detect any significant relationship between the scores of the two waves. Note that no significant differences were detected in attitudes between who dropped out in wave 3 and who participated in the other waves.

Table 41. Psycho-attitudinal factors between first wave and third wave survey

ITEMS To what extent do you agree with the following statements? (Assign a score from 1=not at all to 5=very much)		**Wave A** *Mean*	**Wave A** *St dev*	**Wave C** *Mean*	**Wave C** *St dev*	**Diff**	**T-stat**
ATTACHMENT TO THE CAR	A1. The car is the most convenient means of transport in terms of trip time	3.37	1.22	3.56	1.11	-0.19	-1.39
	A2. The car offers a high level of comfort (comfort, privacy, flexibility, *etc.*) that other forms of transport do not provide	3.90	1.09	4.03	0.96	-0.13	-1.07
	A3. The car is the only means of transport compatible with daily commitments (work, school runs, shopping, *etc.*)	3.05	1.29	3.33	1.26	-0.27	-1.86
	A4. Driving is a pleasurable experience	2.86	1.31	2.76	1.31	0.10	0.66
	A5. Driving gives a feeling of freedom that other means of transport cannot provide	2.98	1.28	3.01	1.29	-0.03	-0.18
	A6. Car use is a habit: one does not consider available alternatives every time	3.50	1.24	3.45	1.37	0.05	0.33
	A7. Owning a nice car is a sign of prestige and a status symbol	1.83	1.13	1.91	1.18	-0.08	-0.6
	A8. The car is a means of self-expression and a reflection of personal taste	1.72	1.00	1.81	1.05	-0.09	-0.73
AVERSION TO PUBLIC TRANSPORT	B1. Travel times are too long	3.52	1.11	3.44	1.14	0.08	0.62
	B2. Services are not reliable in that they do not guarantee regularity and certainty of travel times	3.44	1.09	2.92	1.25	0.52	3.81
	B3. Comfort is poor (overcrowding, carrying bulky goods, *etc.*)	3.47	1.10	3.17	1.18	0.30	2.28
	B4. The service is not compatible with daily commitments (work, school runs, shopping, *etc.*)	3.36	1.16	3.45	1.22	-0.09	-0.63
	B5. Traveling on public transport is not a pleasurable experience	2.69	1.14	2.48	1.11	0.21	1.65
	B6. Public transport is unpopular because people do not like depending on others to get around	2.99	1.27	3.03	1.24	-0.04	-0.28
	B7. Only those who do not have alternatives use public transport as they are obliged to do so.	2.84	1.27	2.41	1.14	0.43	3.07
	B8. Public transport use is commonly associated with modest social and economic condition	2.03	1.24	1.91	1.13	0.13	0.93
WILLINGNESS TO USE THE LIGHT RAIL	C1. I would use the light rail if travel times were shorter	4.23	1.09	3.86	1.25	0.37	2.72
	C2. I would use the light rail if fares were cheaper	4.33	0.93	3.83	1.22	0.50	4.01
	C3. I would use the light rail if CO_2 emissions were reduced	4.30	0.97	3.96	1.14	0.34	2.79
	C4. I would use the light rail if it was less stressful than driving	4.52	0.81	4.31	0.97	0.21	2.01
	C5. I would use the light rail if the network was extended and the number of lines increased	4.67	0.67	4.56	0.85	0.11	1.29
	C6. I would use the light rail if the service was free	4.31	1.15	3.93	1.35	0.39	2.68
	C7. I would use the light rail if there was a free wi-fi service on board	3.94	1.23	3.41	1.40	0.54	3.51

PRO ENVIRONMENTAL BEHAVIOR	D1. I unplug electronic devices when they are not in use (*e.g.* TV, phone charger, *etc.*)	3.98	1.23	3.82	1.16	0.15	1.11
	D2. I use low-energy light bulbs	4.36	0.94	4.40	0.90	-0.04	-0.38
	D3. I do not waste water	4.20	1.03	4.30	0.98	-0.10	-0.87
	D4. I buy local fruit and vegetables, which are not transported by plane or lorries	3.95	1.15	4.12	1.05	-0.17	-1.36
	D5. When shopping, I use my own reusable bag instead of the plastic bag provided by the supermarket	4.46	0.98	4.40	0.91	0.06	0.55
	D6. I use public transport to deliberately reduce the air pollution caused by car use	2.98	1.40	3.04	1.27	-0.06	-0.39
	D7. For short trips, I cycle or walk, rather than taking the car	3.97	1.23	4.03	1.15	-0.06	-0.44
ENVIRONMENTAL AWARENESS	E1. It is very important to be aware of how one's own actions can impact the environment	4.56	0.73	4.50	0.85	0.05	0.58
	E2. Environmental awareness is a very important personal characteristic	4.49	0.75	4.42	0.90	0.07	0.77
	E3. Human activities are seriously abusing the environment and its resources	4.62	0.71	4.67	0.72	-0.05	-0.57
	E4. Pro-environmental behavior is very satisfying	4.32	0.90	4.33	0.89	-0.01	-0.06
	E5. Daily use of the car is one the most environmentally harmful human activities	3.91	0.97	3.91	1.07	0	0
	E6. Using public transport for daily trips helps considerably to improve our environment	4.23	0.84	4.20	0.90	0.03	0.33

At an aggregate level, some differences were detected in the answers to psycho-attitudinal questions, with some average responses being significantly different over the two waves. Examining the psycho-attitudinal factors measured, it emerges that during the time elapsed between the two surveys:

1. Users reported greater attachment to their cars, consistent with the greater number of car owners. In particular, there was a significant difference in the indicator A3, suggesting that the perception of the car as the only means of transport compatible with daily commitments is greater in spite of the introduction of the light rail and the new bus line. This is due to the limited coverage of the light rail service in the urban area and to the fact that the new bus route only serves a residential corridor, where few shops or public offices exist. Moreover, note that in the third wave some students had started work, and therefore have less flexible working hours than students.

2. By contrast, the sample were less unwilling to use public transport, in keeping with the findings of a well-controlled field experiment conducted in Sweden by Pedersen *et al.* (2011). It is interesting to observe a significant difference in the indicators B2 and B7. The first, related to the regularity of the public transport service, scored lower in wave C, indicating that after introduction of the light rail, that runs along a rail corridor and is not slowed by vehicle traffic, the sample consider the public transport service more reliable in terms of punctuality of the service and fixed travel times. The second, concerning what kind of people use public transport, also scored lower in the third wave. This is because the light rail is used not only by people who cannot afford a car but also by high-income earners, who appreciate the benefits in terms of travel time and a greater level of comfort than the bus.

3. Individuals were less willing to use the light rail. A possible explanation for this result is that before it went into service, the expectations surrounding the new alternative stirred people's

curiosity, mainly because, as found in other works, the light rail could be perceived to offer a higher service quality than the bus.

4. For pro-environmental behavior and environmental awareness statements, on average we did not find any significant differences. This result shows a weak correlation between these two constructs and the introduction of a structural measure in the choice context.

Interestingly, there are some differences in psycho-attitudinal factors across population segments. For instance, taking into account the construct *Attachment to the car* several distinctions can be observed in the answers provided by men and women across the two waves, confirming that the psycho-attitudinal factors can change over time and this variation may depend on individuals' characteristics.

A confirmatory factor analysis was performed prior to modeling choices in order to identify one or more latent dimensions (called factors or components) underpinning a set of items or variables. Table 42 shows the factors with the linked items.

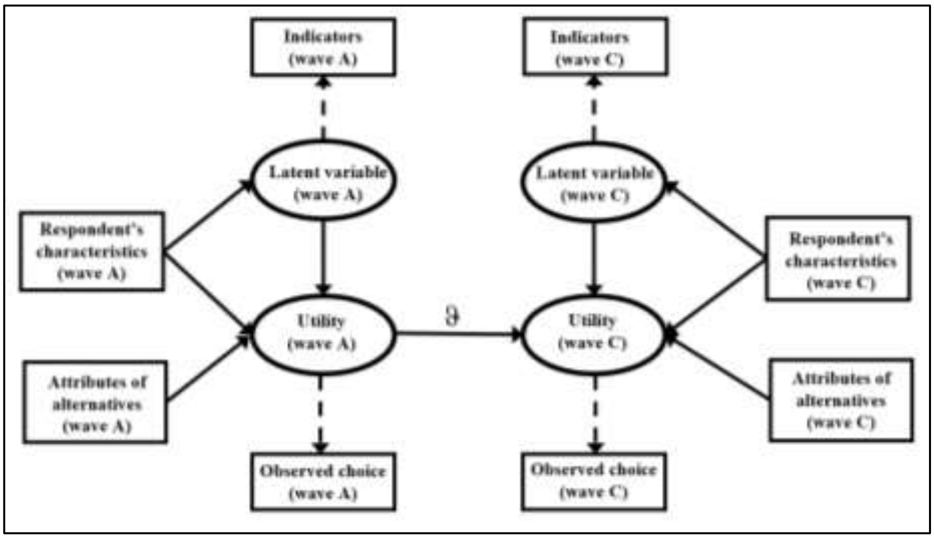
Table 42. Factor analysis

Attachment to the car
A1. A2. A3. A4. A5.
Unwillingness to use public transport
B1. B2. B3. B4. B5.
Willingness to use the light rail
C1. C2. C3. C4. C5.
Pro-environmental behavior
D1. D2. D3. D4. D5.
Environmental awareness
E1. E2. E3. E4. E5. E6.

3.5.2. Model specification

The model framework we used is a latent variable model jointly estimated on a two-wave panel dataset as shown in Figure 14. The discrete choice model (DCM) in our ICLV model is a multinomial logit that incorporates latent variables to measure individual attitudes.

Figure 14. Model framework



Because we collected data in two different time waves, before and after implementation of the new light rail, the hybrid structure for each wave is jointly estimated to control for scale differences between the two datasets, to detect any differences in individual preferences and attitudes between the two waves.

As in typical discrete choice models, we define U_{jq}^{w} the utility that each individual q associates with alternative j in the first (w = A) and third wave (w = C) respectively. The discrete part of the joint discrete choice model can be specified as:

$$U_{jq}^{A} = ASC_j + \boldsymbol{\beta}_{j\boldsymbol{LOS}}\boldsymbol{LOS}_{jq}^{A} + \boldsymbol{\beta}_{j\boldsymbol{SE}}\boldsymbol{SE}_{q}^{A} + \beta_{jLV}^{A} LV_q^{A} + \varepsilon_{jq}^{A} \tag{3.1}$$

$$U_{jq}^{C} = \theta\left(ASC_j + \boldsymbol{\beta}_{j\boldsymbol{LOS}}\boldsymbol{LOS}_{jq}^{C} + \boldsymbol{\beta}_{j\boldsymbol{SE}}\boldsymbol{SE}_{q}^{C} + \beta_{jLV}^{C} LV_q^{C} + \varepsilon_{jq}^{C}\right) \tag{3.2}$$

where SE_q^w is a vector of individual socio-economic characteristics, $LOS_{jn}^{w}$is a vector of travel mode alternative attributes, $LV_q^w$is a latent variable, $\beta_{j\boldsymbol{LOS}}$, $\beta_{j\boldsymbol{SE}}$,β_{jLV}^{w} are vectors of coefficients associated with the variables and ASC_j^w are the alternative constants. ε_{jq}^{w} are the independently and identically distributed Gumbel error terms for each wave and $\theta = \sigma^A/\sigma^C$ is the scale parameter that yields the same variance in both wave utilities.

Following the framework of hybrid choice models, we model each psychological construct as a latent variable that depends on the socio-economic characteristics of each individual q:

$$LV_q^w = \kappa^w + \boldsymbol{\lambda}^w \boldsymbol{SE}_q^w + \omega_q^w \tag{3.3}$$

where κ^w is the intercept, λ^w is the vector of the coefficients associated with the socio-economic characteristics and ω_q^w is the normally distributed error term, with zero mean and standard deviation σ_ω^w. SE_q^w can be different from the socio-economic characteristics included in the discrete choice model and all coefficients are allowed to vary between waves.

The measurement equation of the discrete choice model is defined by a dummy variable that takes the value one if the alternative chosen has the highest utility, zero otherwise:

$$y_{jq}^{w} = \begin{cases} 1 if U_{jq}^{w} = max_i\{U_{iq}^{w}\} \\ 0 \; otherwise \end{cases} \quad \forall\, i \in D^{w}(q) \tag{3.3}$$

where $D^{w}(q)$ is the set of alternatives available to individual q in each wave.

The measurement equation of the latent variable is given by a set of R indicators according to the following expression:

$$I_{rq}^{w} = \gamma_r^{w} + \alpha_r^{w} LV_q^{w} + \upsilon_{rq}^{w} \tag{3.4}$$

where γ_r^{w} is the intercept, α_r^{w} is the coefficient associated with the latent variable and υ_{rn}^{w} is an error term that can have any distribution Q (assumed to have zero mean and standard deviation σ_υ^{w}). All the coefficients are also allowed to vary between the first and third wave.

Indicators are expressed in a five-point numerical scale, so the measurement equation of the indicators is expressed as an ordered logit:

$$P(I_{rq}^{w} = 1) = \frac{1}{1 + e^{[\gamma_r^{w} + \alpha_r^{w} LV_q^{w} - \eta_1]}}$$

$$P(1 < I_{rq}^{w} < 5) = \frac{1}{1+e^{[\gamma_r^{w} + \alpha_r^{w} LV_q^{w} - \eta_I]}} - \frac{1}{1+e^{[\gamma_r^{w} + \alpha_r^{w} LV_q^{w} - \eta_{I-1}]}} \tag{3.5}$$

$$P(I_{rq}^{w} = 5) = 1 - \frac{1}{1 + e^{[\gamma_r^{w} + \alpha_r^{w} LV_q^{w} - \eta_{I-1}]}}$$

where η_I are thresholds defined respectively as $\eta_1 = 0, \eta_2 = \eta_1 + \delta_1, \eta_3 = \eta_2 + \delta_2, \eta_4 = \eta_3 + \delta_3$.

Because we assumed that ε_{jq}^{w} is i.i.d. Gumbel across alternatives, in each wave, the probability that decision-maker q chooses alternative j is given by:

$$P_{jq}^{A}(\omega_q^{A}) = \frac{exp\left(ASC_j + \beta_{j\boldsymbol{LOS}}\boldsymbol{LOS}_{jq}^{A} + \beta_{jSE}\boldsymbol{SE}_q^{A} + \beta_{jLV}^{A} LV_q^{A}(\omega_q^{A})\right)}{\sum_{j\in D_q} exp\left(ASC_j + \beta_{j\boldsymbol{LOS}}\boldsymbol{LOS}_{jq}^{A} + \beta_{j\boldsymbol{SE}}\boldsymbol{SE}_q^{A} + \beta_{jLV}^{A} LV_q^{A}(\omega_q^{A})\right)} \tag{3.6}$$

$$P_{jq}^{C}(\omega_q^{C}) = \frac{\exp\left(\theta\left(ASC_j + \beta_{j\boldsymbol{LOS}}\boldsymbol{LOS}_{jq}^{C} + \beta_{j\boldsymbol{SE}}\boldsymbol{SE}_q^{C} + \beta_{jLV}^{C} LV_q^{C}(\omega_q^{C})\right)\right)}{\sum_{j\in D_q} \exp(\theta(ASC_j + \beta_{j\boldsymbol{LOS}}\boldsymbol{LOS}_{jq}^{C} + \beta_{j\boldsymbol{SE}}\boldsymbol{SE}_q^{C} + \beta_{jLV}^{C} LV_q^{C}(\omega_q^{C})))} \tag{3.7}$$

The joint probability for an individual q making the choice j for each period w is the integral over the distribution ω_q:

$$P_{jq}^{w} = \int_{\omega} P_{jq}^{w}(\omega_q^{w}) f_{LV}^{w}(\omega_q) \prod_{r} f_{Ir}^{w}\left(I_{rq} \middle| LV_q^{w}(\omega_q^{w})\right) d\omega \tag{3.8}$$

where $f_{LV}^{w}(\omega_q)$ and $f_{Ir}^{w}\left(I_{rq} \middle| LV_q^{w}(\omega_q)\right)$ are the distribution of the latent variable and the indicators, respectively. All the models were estimated using PythonBiogeme (Bierlaire, 2016).

3.5.3. Model results

Three modes were considered to be available for the trip to Cittadella: 1) private car, 2) bus, 3) light rail. Travel times and costs for each mode were determined for each commuter, based on the location of the person's home and work. We simulated, for each individual, the values of the attributes of the non-chosen available alternatives using Citlabs CUBE software. Travel time was differentiated as walk time (for private car, bus and light rail) and in-vehicle time (for private car, bus and light rail). Walking time enters separately for car, public transport and light rail. In-vehicle time enters separately for car, public transport and light rail. The cost was entered specific in the utility function of car, bus and light rail. Before estimating the joint hybrid choice model, we tested various discrete choice models to gain a better knowledge of the phenomenon. The most relevant estimation results of the best specifications are summarized in Table 43.

The estimation of the discrete part alone is shown in the first column of Table 43. The sign of all level of service (LOS) coefficients is in line with microeconomic theory. The frequency attribute, which measures the number of times/year the trip to the Cittadella is made, was positive when incorporated into the utility function of the car, indicating a habit effect in car use. In terms of socio-economic characteristics, not surprisingly, the level of personal income, the number of cars per household and car ownership positively affect the utility of the car mode. The scale parameter (θ) that allows for heteroskedasticity between waves was not significantly different from one, meaning that the two datasets have the same variance.

Several hybrid choice models are estimated for measuring the effect of individual characteristics on mode choice, accounting for attitudes and perceptions collected at two different moments in time (first and third wave surveys). Importantly only the latent variable *Attachment to the car* was found relevant for the purpose of the study. We also tested a number of interaction terms between LOS variables (*i.e.* car travel time and public transport (PT) travel time) and the latent variable *Attachment to the car* but none of the results turned out to be statistically significant.

The last four columns in Table 43 show the results of estimating two hybrid choice models that also include the latent variable *Attachment to The Car*. In model AC1 we used the indicators of the psychological factor collected only in the first wave while model AC2 was estimated using indicators of the psychosocial factor collected in both the first and third waves. This kind of specification allowed us to quantify the error committed using the indicators of the psychological factor collected in the first wave, for both waves.

As can be observed all the coefficients of the discrete part in the two hybrid choice models are in line with the results of the DCM alone.

The latent variable was positive in both models for each wave, indicating that people who are more attracted by the car for its high level of comfort, flexibility and shorter travel time have a lower level of disutility associated with the private vehicle.

As the coefficients of the latent variable model were introduced specific for each wave, we were able to take into account whether 1) the *Attachment to the car* changed over time, even if the analysis of indicator values (Table 41) did not detect significant changes, and 2) *Attachment to the car* affects mode choice in different ways. We found that, for both models, the coefficients associated with the latent variable were not significantly different over waves, showing that the impact of the psychological construct remained stable over time, even after introduction of the new light rail and bus line. There is an intuitive explanation for this phenomenon: a certain number of people during the timeframe of the survey started to use the car, developing a stronger car dependence, so those who showed less attachment to the car due to the introduction of the light rail and the new bus route were replaced by these.

The structural model for the first wave indicates that men have a higher level of attachment to the car. This result is in line with the findings of other works, showing that men, generally, tend to be more car dependent. Also, the number of cars per driver in the household positively impacts car attachment, in keeping with the results of the discrete part of the model.

Table 43. Model results. n/a not applicable.

	DCM alone		ATTACHMENT TO THE CAR			
			ICLV - AC1		ICLV - AC2	
Attributes	**values**	**R t-test**	**values**	**R t-test**	**values**	**R t-test**
Discrete part						
Constant car	-5.05	-3.21	-4.14	-1.50	-4.65	-1.65
Constant bus	-3.61	-2.85	-2.16	-1.28	-2.24	-1.31
Car attributes						
Travel time	-0.08	-1.70	-0.07	-2.02	-0.07	-2.05
Travel Cost	-0.34	-1.39	-0.22	-0.99	-0.22	-0.98
Walking Time from/to parking area	-0.02	-0.40	-0.02	-0.68	-0.01	-0.58
Bus attributes						
Travel Time	-0.05	-2.66	-0.04	-2.47	-0.01	-2.53
Walking Time from/to Bus stop	-0.06	-1.70	-0.06	-1.84	-0.06	-1.80
Light rail attributes						
Travel Time	-0.10	-3.39	-0.07	-2.22	-0.07	-2.23
Walking Time from/to Light rail station	-0.18	-2.82	-0.12	-1.50	-0.12	-1.57
Bus and Light rail attributes						
Cost	-0.81	-1.87	-0.63	-1.54	-0.60	-1.45
Socio-economic characteristics (specific to car)						
Personal income	0.28	1.30	0.22	1.01	0.24	1.06
Number of cars per driver in the household	0.87	1.73	0.45	0.84	0.48	1.00
Car ownership (Yes = 1; No = 0)	0.85	2.03	0.68	1.35	0.68	1.38
Other characteristics of the trip (specific to car)						
Frequency of trips from origin to Cittadella	0.003	1.57	0.002	1.23	0.002	1.31
Scale factor θ (R t-test against 1)	1.30	1.03	2.10	0.79	2.09	0.79
LV attachment to car wave A	n/a	n/a	0.40	1.49	0.52	1.28
LV attachment to car wave C	n/a	n/a	0.20	0.74	0.31	2.08
Latent variable model						
Latent variable Attachment to the CAR wave A structural equation						
Intercept	n/a	n/a	3.74	10.50	3.83	13.19
Standard deviation of error term	n/a	n/a	0.80	5.95	0.78	5.78
Gender (Man = 1; Woman = 0)	n/a	n/a	0.33	2.00	0.30	1.90
Number of cars per driver in the household	n/a	n/a	0.44	1.40	0.46	1.48
Worker dummy (Yes = 1; No = 0)	n/a	n/a	-0.70	-1.90	-0.78	-2.63
Student dummy (Yes = 1; No = 0)	n/a	n/a	-0.36	-1.02	-0.45	-1.73
Latent variable Attachment to the CAR wave C structural equation						
Intercept	n/a	n/a	n/a	n/a	4.01	13.53
Standard deviation of error term	n/a	n/a	n/a	n/a	0.54	3.47
Age	n/a	n/a	n/a	n/a	-0.01	-1.69
Number of cars per driver in the household	n/a	n/a	n/a	n/a	0.28	1.39
Initial log-likelihood	-251.16		-1,563.60		-2,846.83	
Final log-likelihood	-134.74		-1211.96		-2,278.60	
Adjusted ρ^2	0.41		0.20		0.18	

Noteworthy is the fact that the explanatory variables relevant in the structural equation of the first wave were no longer relevant in the structural equation of the third wave, except for the number of cars per driver in the household. This result shows that the explanatory variables related to the latent variable can change over time, validating the criticism raised in Chorus and Kroesen (2014), whereby using cross-sectional data does not allow for within-person comparisons.

Table 44 shows the probability of choosing car and elasticity of demand for car with respect to cost by car. The disaggregate direct elasticity is computed according to the following expression:

$$E_{c_{jq}}^{j} = \frac{\partial P_q(j)}{\partial U c_{jq}} \frac{c_{jq}}{P_q(j)} \tag{3.9}$$

where P_q is the choice probability that individual q chooses the car and c_{jq} is the cost, expressed in euros, associated with the alternative car for individual q. The results are similar in all models, since the latent variable does not modify the marginal utilities.

However, the latent variable affects overall utility, hence choice probability. The model AC1 slightly underestimates the choice probability in the third wave, compared to model AC2, and this effect is more pronounced when the analysis is performed for those categories which we found relevant for explaining the *Attachment to the car*. Nevertheless, the difference in choice probability is quite small, and this might be viewed as an additional proof of the stability in the overall level of the latent attitude between the two waves.

Table 44. Probability and elasticity

	Probability of choosing car		**Elasticity of the demand respect to cost by car**	
	MODEL AC1	MODEL AC2	MODEL AC1	MODEL AC2
First wave	83.3%	83.2%	-0.075	-0.075
Female	83.5%	83.1%	-0.075	-0.076
Student	81.0%	81.1%	-0.084	-0.083
Worker	85.4%	85.0%	-0.064	-0.065
Third wave	77.9%	78.3%	-0.197	-0.194
Female	75.9%	77.2%	-0.206	-0.194
Student	67.9%	68.4%	-0.315	-0.309
Worker	81.9%	82.7%	-0.151	-0.144

CONCLUSIONS

The primary objective of this thesis is to explore the methodological processes able to guarantee that sustainable mobility policies seeking to reduce car use are accompanied by a quantitative assessment of the effects that this generates in the transport context. In particular, we focused on the role and weight of psycho-attitudinal variables (attitudes, perceptions, habits, etc.) in the process of choosing a sustainable mode of transport. Addressing these aspects is far from straightforward and has motivated the development of the models presented in this thesis.

ASSESSMENT AND FORECASTING OF BIKE USE IN AN URBAN CONTEXT. The first part of the thesis presents the findings of a study focusing on unraveling the linkage among psycho-attitudinal factors related to bike use and the choice to cycle. We used different econometric approaches that allow to jointly model multiple outcomes of mixed types. The data used are drawn from a survey conducted in Sardinia (Italy), where cycling is mainly considered as a form of exercise and recreation.

Our research findings indicate that is essential to consider people's psychological characteristics in order to develop better bike promoting strategies. From the results of the first modeling application it emerged that the choice to travel to work by bike, besides level-of-service and network characteristics, such as travel time and topography, is significantly affected by latent inertia. The positive effect of the latent variable inertia in the bike alternative suggests that, like other modes of transport, bike commuting is habit forming. So, it appears that the frequency of past cycling behavior is a crucial factor in the mode choice process that leads to prefer commuting by bike rather than by other means of transport. These findings provide further support for policy initiatives like the implementation of programs in which streets are temporarily closed to traffic allowing access to individuals for leisure activities, which are based on the assumption that convincing people to take up cycling could lead to an increase in cycling for work and non-work purposes.

In the second application we specifically studied whether psycho-attitudinal factors vary among people with different cycling experience through the estimation of an integrated choice latent variable model with a generalized ordered probit choice kernel, where we allowed the thresholds themselves to be a function of both objectives and psycho-attitudinal variables. The results indicate that all three latent variables, related to the perceptions of the positive aspects of biking, positively influence the propensity to cycle, supporting the idea of a relationship between attitudes and the cycling experience. In the third application we conducted a multivariate analysis to explore the interplay among psycho-attitudinal factors related to cycling and the choices to cycle to work, for shopping and for leisure. In doing so, we estimated a multivariate ICLV model, that allows to jointly model a large set of mixed data outcomes. We found that both explanatory variables and cycling perceptions, as captured by our latent variables, are similar for the three outcomes (commute, shopping and leisure). Further, through the presence of common error terms in the utility functions of choice dimensions, we account for observed endogenous effects, and in particular we found that cycling for leisure positively influences the choice to cycle for utilitarian purposes. Outcomes of the second and third application reinforce the idea that promoting cycling through the implementation of awareness campaigns and educational programs, intended to improve peoples' perceptions of the bike mode, can persuade them to consider the bike as an alternative means of transport to private motorized vehicles. Further, investments aimed at supporting use of the bike for leisure (*e.g.* cycle routes) may increase the number of people who choose to use the bike as an alternative means of transport for commuting or shopping.

In the last application, we explored how facilitators to cycling were perceived by different segments of individuals, in view of assessing how to best promote cycling in an urban area. To perform this analysis, we estimate a multivariate ordered probit. Results indicate that how people perceive the implementation of policy measures aimed to encourage more frequent cycling, depends on their socio-demographic characteristics. Hence, a holistic approach with a variety of activities is needed, as improvements in cycling infrastructure may not be enough. Changes in driving culture and promotional campaigns targeted at specific population segments would be the best approach to promote this green mode of transport.

Nevertheless, the current study also contains some limitations that need to be stressed and potentially addressed by future research. First, the sample was not representative of the whole population, but was composed predominantly of public sector employees, whose socio-demographic characteristics and work-related factors are fairly homogeneous. Understanding the decision to commute by bicycle of the entire population, without excluding any segment of workers, is crucial for policy makers who intend to implement effective strategies for promoting bike use. Another limitation is that all the latent variables included in our analysis were related to bike use, when it is plausible that other psycho-social factors, such as a green lifestyle or a strong attachment to the car, influence commute mode choice.

Lastly, it should be stressed that because this research adopted revealed preference data, we only provided associations between latent variables and outcome variables. As suggested by Chorus and Kroesen, 2014, latent variables and choice variables are likely to have influenced each other over time, and only the use of longitudinal data may help in establishing causality. Notwithstanding this limitation, cross-sectional studies, like ours, which allow to jointly model a large set of mixed data outcomes can lead to more accurate policy assessment of strategies designed to encourage bike use.

EVALUATION OVER TIME OF TRAVEL DEMAND MANAGEMENT STRATEGIES. The second part of the work attempted to assess in quantitative terms the role that a combination of hard (i.e., new mode of sustainable transport—light railway line in the choice set) and soft measures (differentiated with respect to the degree of personalization—mass communication versus personalized travel planning) can have in travel mode choice (switch from car driver to light rail).

A comparison was carried out of the data collected in three survey waves before and after a new light railway line went into service. The research analyzed the short- and long-term additional effect of the personalized soft measure versus the combination of hard and soft (mass communication) measures usually implemented when a new service becomes operational.

First of all, this study brought to light the importance of, and at the same time the difficulties encountered in, collecting longitudinal data. Unfortunately, with surveys comprising different phases spread over time a large number of the initial participants may fall by the wayside, in spite of a major promotional campaign. In the case at hand however, note that the metropolitan area of Cagliari is relatively small (population 431,819), and for this reason it is difficult to intercept a large number of individuals to ensure, after processing the responses, that sufficiently large samples are obtained. Clearly this problem has repercussions on the results obtained for behavior change, as in numerical terms they may appear irrelevant. Moreover, the time gap between the survey and implementation of the light rail (2 months and 2 years) may influence attitudes towards a certain means of transport and be a limitation when analyzing the results.

Next, we evaluated the overall effect of all the implemented measures in the short term. We described in detail the methodology and procedure adopted for implementing the sensitization campaign (soft measure). Validation of the results depends both on the methodological framework and on how the different phases of the campaign are conducted, so as to be able to distinguish the contribution of hard from soft measures and a combination thereof. For the design of the personalized soft measures (VTBC program) the key factors that distinguish successful VTBC programs were used, in an attempt to overcome those problems that generally affect this kind of measure (*e.g.*, lack of a control group). For assessing changes in travel behavior, the modal shares observed in the first and second wave surveys were compared. Considering the results, it can reasonably be claimed that the awareness campaign (soft measure + VTBC program) in the specific context contributes by 12% to behavior change. This is in line with the findings of the few studies reported in the literature that quantify the contribution of VTBC programs to behavior change as 5 to 15%. Further, the combination of hard and soft measures achieved a change in travel behavior of 34%, when the measure is not personalized, and 46% with the VTBC program. This confirms the importance of adopting TDM measures on a large scale, with a combination of measures that changes the choice context and information measure, both generalized and personalized.

Finally, we attempted to assess, in quantitative terms, the evolution of psycho-attitudinal factors over time before and after implementation of the new travel alternative and their role in the change process. The first important result obtained concerns confirmation of considering data gathered before and after implementation of policy measures, also with respect to those psycho-social attributes that could play a crucial role in change processes and might vary over time. If these are not constantly and sufficiently measured, they would deprive us of the opportunity both to measure them in modeling terms and analyze and evaluate whether they influence travel behavior or *viceversa* are affected by behavior. Even more so when modeling results are intended to provide indications for intervention policies aimed at incentivizing preferably car drivers to change their travel behavior.

From a modeling perspective, our results show that the explanatory variables used in the structural equation of the latent variable changed between the two waves, accounting for within-person comparisons. As discussed by Chorus and Kroesen (2014) the use of cross-sectional data would not have allowed us to understand this kind of aspect and would have led us to derive inappropriate policy implications from the model. However, we are also aware that the estimation of the ICLV model allowed us to understand that the latent factor *attachment to the car* remains unchanged, even after introduction of the new bus route and light rail line. But, if these kinds of variables are stable over time, this means for policy makers that the implementation of a structural measure does not suffice to significantly impact individuals' cognitive factors. These findings support the idea of other studies that only the presence of a strong shock in the choice context (such as the prohibition of a mode alternative) or the constant provision over years of personalized information (in our study we implemented the soft measure only once, a few weeks after the new light rail line went into service), which focus on those factors that could diminish this emotional attachment, are able to trigger a shift in people's psycho-attitudinal characteristics. Because of the small sample size, the study did not allow to define a rule about the implications of ICLV, but it does provide a cue to investigate this issue in different contexts using a robust sample. If other studies focused on the same aspects of our work and found the same conclusions, these could become the correct answer to the criticism surrounding ICLVs.

REFERENCES

Aarts, H., Verplanken, B., & Van Knippenberg, A. (1997). Habit and information use in travel mode choices. Acta Psychologica, 96(1-2), 1-14.

Abou-Zeid, M., & Ben-Akiva, M. (2011). The effect of social comparisons on commute well-being. Transportation Research Part A: Policy and Practice, 45(4), 345-361.

Abraham, J. E., McMillan, S., Brownlee, A. T., & Hunt, J. D. (2002, January). Investigation of cycling sensitivities. In 81st Annual Meeting of the Transportation Research Board. Transportation Research Board, Washington, DC.

Abrahamse, W., Steg, L., Vlek, C., & Rothengatter, T. (2005). A review of intervention studies aimed at household energy conservation. Journal of environmental psychology, 25(3), 273-291.

Ajzen, I. (1991). The theory of planned behavior. Organizational behavior and human decision processes, 50(2), 179-211.

Akar, G., & Clifton, K. (2009). Influence of individual perceptions and bicycle infrastructure on decision to bike. Transportation Research Record: Journal of the Transportation Research Board, (2140), 165-172.

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

Aldred, R., Elliott, B., Woodcock, J., & Goodman, A. (2017). Cycling provision separated from motor traffic: a systematic review exploring whether stated preferences vary by gender and age. Transport reviews, 37(1), 29-55.

Ampt, E. Voluntary Household Travel Behaviour Change: Theory and Practice. Paper presented at the 10th International Association of Travel Behaviour Research Conference, Lucerne, Switzerland, August 15–16 2003.

Ampt, L. (1999). Grass routes: from travel blending to living neighbourhoods. Traffic engineering & control, 40(10).

Anable, J. (2005). 'Complacent car addicts' or 'aspiring environmentalists'? Identifying travel behaviour segments using attitude theory. Transport policy, 12(1), 65-78.

Aoun, A., Abou-Zeid, M., Kaysi, I., & Myntti, C. (2013). Reducing parking demand and traffic congestion at the American University of Beirut. *Transport Policy*, 25, 52-60.

Arroyo, R., Ruiz, T., Casquero, D., & Mars, L. (2018). Trip characteristics analysis of the effects of a travel behavior change program. Transportation Research Record, 2672(47), 146-158.

Aultman-Hall, L., Hall, F. L., & Baetz, B. B. (1997). Analysis of bicycle commuter routes using geographic information systems: implications for bicycle planning. Transportation research record, 1578(1), 102-110.

Bahamonde-Birke, F., & de Dios Ortúzar, J. (2014). Is sequential estimation a suitable second best for estimation of hybrid choice models?. Transportation Research Record, 2429(1), 51-58.

Balsas, C. J. (2003). Sustainable transportation planning on college campuses. Transport Policy, 10(1), 35-49.

Baltes, M. R. (1996). Factors influencing nondiscretionary work trips by bicycle determined from 1990 US census metropolitan statistical area data. Transportation research record, 1538(1), 96-101.

Bamberg, S., Ajzen, I., & Schmidt, P. (2003). Choice of travel mode in the theory of planned behavior: The roles of past behavior, habit, and reasoned action. Basic and applied social psychology, 25(3), 175-187.

Barbour, N., Zhang, Y., & Mannering, F. (2019). A statistical analysis of bike sharing usage and its potential as an auto-trip substitute. Journal of Transport & Health, 12, 253-262.

Barla, P., Lapierre, N., Daziano, R. A., & Herrmann, M. (2012). Reducing Automobile Dependency on Campus: Evaluating the Impact TDM Using Stated Preferences. Cahier de recherche/Working Paper, 3.

Beirão, G., & Cabral, J. S. (2007). Understanding attitudes towards public transport and private car: A qualitative study. Transport policy, 14(6), 478-489.

Ben-Akiva, M. E., Lerman, S. R., & Lerman, S. R. (1985). Discrete choice analysis: theory and application to travel demand (Vol. 9). MIT press.

Ben-Akiva, M., Bolduc, D., & Bradley, M. (1993). Estimation of travel choice models with randomly distributed values of time. Transportation Research Record, 1413, 88-97.

Ben-Akiva, M., McFadden, D., Gärling, T., Gopinath, D., Walker, J., Bolduc, D., Börsch-Supan, A., Delquié, P., Larichev, O., Morikawa, T., Polydoropoulou, A. & Rao, V. (1999). Extended framework for modeling choice behavior. Marketing letters, 10(3), 187-203.

Ben-Akiva, M., McFadden, D., Train, K., Walker, J., Bhat, C., Bierlaire, M., & Daly, A. (2002). Hybrid choice models: progress and challenges. Marketing Letters, 13(3), 163-175.

Bergantino, A. S., Bierlaire, M., Catalano, M., Migliore, M., & Amoroso, S. (2013). Taste heterogeneity and latent preferences in the choice behaviour of freight transport operators. Transport Policy, 30, 77-91.

Berggren, M., Graves, A., Pickus, H., & Hand Wirtis, L. (2012). The 20s Bikeway: Clinton to Steele. Portland State University.

Bergström, A., & Magnusson, R. (2003). Potential of transferring car trips to bicycle during winter. Transportation Research Part A: Policy and Practice, 37(8), 649-666.

Bhat, C. R. (1998). Accommodating variations in responsiveness to level-of-service measures in travel mode choice modeling. Transportation Research Part A: Policy and Practice, 32(7), 495-507.

Bhat, C. R. (1998). Analysis of travel mode and departure time choice for urban shopping trips. Transportation Research Part B: Methodological, 32(6), 361-371.

Bhat, C. R. (2000). Incorporating observed and unobserved heterogeneity in urban work travel mode choice modeling. Transportation science, 34(2), 228-238.

Bhat, C. R. (2011). The maximum approximate composite marginal likelihood (MACML) estimation of multinomial probit-based unordered response choice models. Transportation Research Part B: Methodological, 45(7), 923-939.

Bhat, C. R. (2015). A new generalized heterogeneous data model (GHDM) to jointly model mixed types of dependent variables. Transportation Research Part B: Methodological, 79, 50-77.

Bhat, C. R., & Dubey, S. K. (2014). A new estimation approach to integrate latent psychological constructs in choice modeling. Transportation Research Part B: Methodological, 67, 68-85.

Bhat, C. R., & Guo, J. Y. (2007). A comprehensive analysis of built environment characteristics on household residential choice and auto ownership levels. Transportation Research Part B: Methodological, 41(5), 506-526.

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

Bhat, C. R., Dubey, S. K., & Nagel, K. (2015). Introducing non-normality of latent psychological constructs in choice modeling with an application to bicyclist route choice. Transportation Research Part B: Methodological, 78, 341-363.

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

Bierlaire, M. (2018). Estimating choice models with latent variables with PandasBiogeme (No. REP_WORK).

Böcker, L., Prillwitz, J., & Dijst, M. (2013). Climate change impacts on mode choices and travelled distances: a comparison of present with 2050 weather conditions for the Randstad Holland. Journal of Transport Geography, 28, 176-185.

Bonsall, P. (2009). Do we know whether personal travel planning really works?. Transport Policy, 16(6), 306-314.

Börjesson, M., & Eliasson, J. (2012). The value of time and external benefits in bicycle appraisal. Transportation Research Part A: Policy and Practice, 46(4), 673-683.

Boyd, J. H., & Mellman, R. E. (1980). The effect of fuel economy standards on the US automotive market: an hedonic demand analysis. Transportation Research Part A: General, 14(5-6), 367-378.

Braun, L. M., Rodriguez, D. A., Cole-Hunter, T., Ambros, A., Donaire-Gonzalez, D., Jerrett, M., & de Nazelle, A. (2016). Short-term planning and policy interventions to promote cycling in urban centers: Findings from a commute mode choice analysis in Barcelona, Spain. Transportation Research Part A: Policy and Practice, 89, 164-183.

Brög, W., Erl, E., Ker, I., Ryle, J., & Wall, R. (2009). Evaluation of voluntary travel behaviour change: Experiences from three continents. Transport policy, 16(6), 281-292.

Brown, B. B., Werner, C. M., & Kim, N. (2003). Personal and contextual factors supporting the switch to transit use: Evaluating a natural transit intervention. Analyses of Social Issues and Public Policy, 3(1), 139-160.

Brown, J., Hess, D. B., & Shoup, D. (2001). Unlimited access. *Transportation*, 28(3), 233- 267.

Brownstone, D., & Train, K. (1998). Forecasting new product penetration with flexible substitution patterns. Journal of econometrics, 89(1-2), 109-129.

Buehler, R. (2012). Determinants of bicycle commuting in the Washington, DC region: The role of bicycle parking, cyclist showers, and free car parking at work. Transportation research part D: Transport and Environment, 17(7), 525-531.

Buehler, R., & Pucher, J. (2012). Cycling to work in 90 large American cities: new evidence on the role of bike paths and lanes. Transportation, 39(2), 409-432.

Cairns, S., Sloman, L., Newson, C., Anable, J., Kirkbride, A., & Goodwin, P. (2008). Smarter choices: assessing the potential to achieve traffic reduction using 'soft measures'. Transport Reviews, 28(5), 593-618.

Cao, X., Mokhtarian, P. L., & Handy, S. L. (2009). Examining the impacts of residential self-selection on travel behaviour: a focus on empirical findings. Transport reviews, 29(3), 359-395.

Cardell, N. S., & Dunbar, F. C. (1980). Measuring the societal impacts of automobile downsizing. Transportation Research Part A: General, 14(5-6), 423-434.

Carver, C. S., Scheier, M. F., & Weintraub, J. K. (1989). Assessing coping strategies: a theoretically based approach. Journal of personality and social psychology, 56(2), 267.

Caulfield, B., Brick, E., & McCarthy, O. T. (2012). Determining bicycle infrastructure preferences–A case study of Dublin. Transportation research part D: transport and environment, 17(5), 413-417.

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

Cervero, R., Sarmiento, O. L., Jacoby, E., Gomez, L. F., & Neiman, A. (2009). Influences of built environments on walking and cycling: lessons from Bogotá. International Journal of Sustainable Transportation, 3(4), 203-226.

Cherchi, E., Meloni, I., & de Dios Ortúzar, J. (2014). The latent effect of inertia in the modal choice. In 13th International Conference on Travel Behavior Research (pp. 517-534). Lulu.com.

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

Cialdini, R. B. (2001). Harnessing the science of persuasion. Harvard business review, 79(9), 72-81.

Cialdini, R. B. (2009). Influence: Science and practice (Vol. 4). Boston: Pearson education.

Cruz, L., Barata, E., Ferreira, J. P., & Freire, F. (2017). Greening transportation and parking at University of Coimbra. International Journal of Sustainability in Higher Education, 18(1), 23-38.

Daganzo, C. (2014). Multinomial probit: the theory and its application to demand forecasting. Elsevier.git modeling. Journal of choice modelling, 21, 25-29.

Daly, A., Hess, S., Patruni, B., Potoglou, D., & Rohr, C. (2012). Using ordered attitudinal indicators in a latent variable choice model: a study of the impact of security on rail travel behaviour. Transportation, 39(2), 267-297.

Dargay, J. (2007). The effect of prices and income on car travel in the UK. Transportation Research Part A: Policy and Practice, 41(10), 949-960.

Davies, N. (2012). What are the ingredients of successful travel behavioural change campaigns?. Transport Policy, 24, 19-29.

Daziano, R. A., & Bolduc, D. (2013). Incorporating pro-environmental preferences towards green automobile technologies through a Bayesian hybrid choice model. Transportmetrica A: Transport Science, 9(1), 74-106.

de Dios Ortúzar, J., & Willumsen, L. G. (2011). Modelling transport. John wiley & sons.

De Geus, B. (2007). Cycling to work: psychosocial and environmental factors associated with cycling and the effect of cycling on fitness and health indexes in an untrained working population. Vrije Universiteit, Department of human physiology and sports medicine.

de Leon, A. R., & Chough, K. C. (2013). Analysis of mixed data: An overview. In Analysis of Mixed Data (pp. 29-40). Chapman and Hall/CRC.

De Witte, A., Macharis, C., Lannoy, P., Polain, C., Steenberghen, T., & Van de Walle, S. (2006). The impact of "free" public transport: The case of Brussels. Transportation Research Part A: Policy and Practice, 40(8), 671-689.Delmelle, E. M., & Delmelle, E. C. (2012). Exploring spatio-temporal commuting patterns in a university environment. Transport Policy, 21, 1-9.

DEFRA, A. (2008). Framework for pro-environmental behaviours. department for environment, food and rural affairs. British Government, London, 76.

dell'Olio, L., Ibeas, A., Bordagaray, M., & Ortúzar, J. D. D. (2013). Modeling the effects of pro bicycle infrastructure and policies toward sustainable urban mobility. Journal of Urban Planning and Development, 140(2), 04014001.

Delmelle, E. M., & Delmelle, E. C. (2012). Exploring spatio-temporal commuting patterns in a university environment. Transport Policy, 21, 1-9.

DeMaio, P. (2009). Bike-sharing: History, impacts, models of provision, and future. Journal of public transportation, 12(4), 3.

De Witte, A., Hollevoet, J., Dobruszkes, F., Hubert, M., & Macharis, C. (2013). Linking modal choice to motility: A comprehensive review. Transportation Research Part A: Policy and Practice, 49, 329-341.

Di Dio, S., Casto, B. L., Micari, F., Rizzo, G., & Vinci, I. (2015). Mobility, data, and behavior: the TrafficO2 case study. In Handbook of Research on Social, Economic, and Environmental Sustainability in the Development of Smart Cities (pp. 382-406). IGI Global.

Dickinson, J. E., Kingham, S., Copsey, S., & Hougie, D. J. P. (2003). Employer travel plans, cycling and gender: will travel plan measures improve the outlook for cycling to work in the UK?. Transportation research part D: transport and environment, 8(1), 53-67.

Dill, J., & Carr, T. (2003). Bicycle commuting and facilities in major US cities: if you build them, commuters will use them. Transportation Research Record, 1828(1), 116-123.

Dill, J., & Voros, K. (2007). Factors affecting bicycling demand: initial survey findings from the Portland, Oregon, region. Transportation Research Record: Journal of the Transportation Research Board, (2031), 9-17.

Domencich, T. A., & McFadden, D. (1975). Urban travel demand-a behavioral analysis (No. Monograph).

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

Emond, C. R., Tang, W., & Handy, S. L. (2009). Explaining gender difference in bicycling behavior. Transportation Research Record, 2125(1), 16-25.

Erdoğan, S., Cirillo, C., & Tremblay, J. M. (2015). Ridesharing as a green commute alternative: A campus case study. *International Journal of Sustainable Transportation*, 9(5), 377-388.

Eriksson, L., Garvill, J., & Nordlund, A. M. (2006). Acceptability of travel demand management measures: The importance of problem awareness, personal norm, freedom, and fairness. Journal of environmental psychology, 26(1), 15-26.

Eriksson, L., Garvill, J., & Nordlund, A. M. (2008). Acceptability of single and combined transport policy measures: The importance of environmental and policy specific beliefs. Transportation Research Part A: Policy and Practice, 42(8), 1117-1128.

Ettema, D., & Nieuwenhuis, R. (2017). Residential self-selection and travel behaviour: what are the effects of attitudes, reasons for location choice and the built environment?. Journal of transport geography, 59, 146-155.

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

Ferdous, N., Eluru, N., Bhat, C. R., & Meloni, I. (2010). A multivariate ordered-response model system for adults' weekday activity episode generation by activity purpose and social context. Transportation research part B: methodological, 44(8-9), 922-943.

Fernández-Heredia, Á., Monzón, A., & Jara-Díaz, S. (2014). Understanding cyclists' perceptions, keys for a successful bicycle promotion. Transportation research part A: policy and practice, 63, 1-11.

Fishman, E. (2016). Bikeshare: A review of recent literature. Transport Reviews, 36(1), 92-113.

Foerster, J. F. (1979). Mode choice decision process models: a comparison of compensatory and non-compensatory structures. Transportation Research Part A: General, 13(1), 17-28.

Fraser, S. D., & Lock, K. (2011). Cycling for transport and public health: a systematic review of the effect of the environment on cycling. European journal of public health, 21(6), 738-743.

Friman, M., Larhult, L., & Gärling, T. (2013). An analysis of soft transport policy measures implemented in Sweden to reduce private car use. Transportation, 40(1), 109-129.

Friman, M., Maier, R., & Olsson, L. E. (2019). Applying a motivational stage-based approach in order to study a temporary free public transport intervention. Transport Policy, 81, 173-183.

Fujii, S., & Taniguchi, A. (2006). Determinants of the effectiveness of travel feedback programs—a review of communicative mobility management measures for changing travel behaviour in Japan. Transport policy, 13(5), 339-348.

Fujii, S., Bamberg, S., Friman, M., & Gärling, T. (2009). Are effects of travel feedback programs correctly assessed?. *Transportmetrica*, *5*(1), 43-57.

Gaker, D., Walker, J. L. Insights on Car-Use Behaviors from Behavioral Economics. In Auto Motives: Understanding Car Use Behaviours (Lucas, K., Blumenberg, E., Weinberger, R. eds.), Emerald Group Publishing Limited, Bingley, 2011, pp. 107–120.

Gao, J., Ettema, D., Helbich, M., & Kamphuis, C. B. (2019). Travel mode attitudes, urban context, and demographics: do they interact differently for bicycle commuting and cycling for other purposes?. Transportation, 1-23.

Gardner, G. (1998). Transport implications of leisure cycling. Crowthorne, Berks: Transport Research Laboratory.

Gärling, T., & Fujii, S. Travel Behavior Modification: Theories, Methods, and Programs. The Expanding Sphere of Travel Behaviour Research, 2009, pp. 97–128.

Gärling, T., & Schuitema, G. (2007). Travel demand management targeting reduced private car use: effectiveness, public acceptability and political feasibility. Journal of Social Issues, 63(1), 139-153.

Garrard, J., Rose, G., & Lo, S. K. (2008). Promoting transportation cycling for women: the role of bicycle infrastructure. Preventive medicine, 46(1), 55-59.

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

Glerum, A., Stankovikj, L., Thémans, M., & Bierlaire, M. (2013). Forecasting the demand for electric vehicles: accounting for attitudes and perceptions. Transportation Science, 48(4), 483-499.

Goetzke, F., & Rave, T. (2011). Bicycle use in Germany: explaining differences between municipalities with social network effects. Urban studies, 48(2), 427-437.

Gollwitzer, P. M. (1993). Goal achievement: The role of intentions. European review of social psychology, 4(1), 141-185.

Graham-Rowe, E., Skippon, S., Gardner, B., & Abraham, C. (2011). Can we reduce car use and, if so, how? A review of available evidence. Transportation Research Part A: Policy and Practice, 45(5), 401-418.

Greene, W. H., & Hensher, D. A. (2003). A latent class model for discrete choice analysis: contrasts with mixed logit. Transportation Research Part B: Methodological, 37(8), 681-698.han

Greene, W. H., & Hensher, D. A. (2010). Ordered choices and heterogeneity in attribute processing. Journal of Transport Economics and Policy (JTEP), 44(3), 331-364.

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

Gunn, H., & Bates, J. (1982). Statistical aspects of travel demand modelling. Transportation Research Part A: General, 16(5-6), 371-382.

Guo, J., Bhat, C., & Copperman, R. (2007). Effect of the built environment on motorized and non motorized trip making: substitutive, complementary, or synergistic?. Transportation Research Record: Journal of the Transportation Research Board, (2010), 1-11.

Habib, K. M. N., & Zaman, M. H. (2012). Effects of incorporating latent and attitudinal information in mode choice models. Transportation Planning and Technology, 35(5), 561-576.

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

Hamre, A., & Buehler, R. (2014). Commuter mode choice and free car parking, public transportation benefits, showers/lockers, and bike parking at work: Evidence from the Washington, DC Region. Journal of Public Transportation, 17(2), 4.

Han, D., Yu, J., Beimborn, E., Jin, Z., & Tan, W. (2019). Elements of Successful Universal Student Transit Pass Programs from Planning to Implementation: A Benchmark Study. Transportation Research Record, 0361198119834916.

Handy, S. L., Boarnet, M. G., Ewing, R., & Killingsworth, R. E. (2002). How the built environment affects physical activity: views from urban planning. American journal of preventive medicine, 23(2), 64-73.

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

Handy, S., Cao, X., & Mokhtarian, P. (2005). Correlation or causality between the built environment and travel behavior? Evidence from Northern California. Transportation Research Part D: Transport and Environment, 10(6), 427-444.

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

Harden, B. (2008). For bicyclists, a widening patchwork world. Washington Post, August, 31(2008), A01.

Hausman, J. A., & Wise, D. A. (1976). A conditional probit model for qualitative choice: discrete decisions recognizing interdependence and heterogenous preferences.

Heath, G. W., Brownson, R. C., Kruger, J., Miles, R., Powell, K. E., & Ramsey, L. T. (2006). The effectiveness of urban design and land use and transport policies and practices to increase physical activity: a systematic review. Journal of physical activity and health, 3(s1), S55-S76.

Heath, Y., & Gifford, R. (2002). Extending the theory of planned behavior: Predicting the use of public transportation 1. Journal of Applied Social Psychology, 32(10), 2154-2189.

Heesch, K. C., Sahlqvist, S., & Garrard, J. (2012). Gender differences in recreational and transport cycling: a cross-sectional mixed-methods comparison of cycling patterns, motivators, and constraints. International Journal of Behavioral Nutrition and Physical Activity, 9(1), 106.

Heinen, E., & Buehler, R. (2019). Bicycle parking: a systematic review of scientific literature on parking behaviour, parking preferences, and their influence on cycling and travel behaviour. Transport Reviews, 1-27.

Heinen, E., Maat, K., & Van Wee, B. (2011). The role of attitudes toward characteristics of bicycle commuting on the choice to cycle to work over various distances. Transportation research part D: transport and environment, 16(2), 102-109.

Heinen, E., Maat, K., & van Wee, B. (2013). The effect of work-related factors on the bicycle commute mode choice in the Netherlands. Transportation, 40(1), 23-43.

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

Hess, S., Daly, A., & Batley, R. (2018). Revisiting consistency with random utility maximisation: theory and implications for practical work. Theory and Decision, 84(2), 181-204.

Howard, C., & Burns, E. K. (2001). Cycling to work in Phoenix: Route choice, travel behavior, and commuter characteristics. Transportation Research Record, 1773(1), 39-46.

Hsieh, H. S., Kanda, Y., & Fujii, S. (2017). Reducing car use by volitional strategy of action and coping planning enhancement. Transportation research part F: traffic psychology and behaviour, 47, 163-175.

Huang, X., Cao, X. J., Yin, J., & Cao, X. (2017). Effects of metro transit on the ownership of mobility instruments in Xi'an, China. Transportation Research Part D: Transport and Environment, 52, 495-505.

Hunt, J. D., & Abraham, J. E. (2007). Influences on bicycle use. Transportation, 34(4), 453-470.

ISFORT. 15° Rapporto sulla mobilità in Italia. ISFORT, www.isfort.it, 2018.

Isler, E., Hoel, L. A., & Fontaine, M. D. (2005). Innovative parking management strategies for Universities: accommodating multiple objectives in a constrained environment. In Annual Meetings of the Transportation Research Board, Washington, DC.

ISTAT (2014). I tempi della vita quotidiana. <https://www.istat.it/it/files/2016/11/Report_Tempidivita_2014.pdf>

Jariyasunant, J., Abou-Zeid, M., Carrel, A., Ekambaram, V., Gaker, D., Sengupta, R., & Walker, J. L. (2015). Quantified traveler: Travel feedback meets the cloud to change behavior. Journal of Intelligent Transportation Systems, 19(2), 109-124.

Jensen, A. F., Cherchi, E., & Mabit, S. L. (2013). On the stability of preferences and attitudes before and after experiencing an electric vehicle. Transportation Research Part D: Transport and Environment, 25, 24-32.

Kamargianni, M., & 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.

Kamargianni, M., Dubey, S., Polydoropoulou, A., & Bhat, C. (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.

Kristensen, J. P., & Marshall, S. (1999). Mobility management to reduce travel: the case of Aalborg. Built Environment (1978-), 138-150.

Krizek, K. J., & Johnson, P. J. (2006). Proximity to trails and retail: effects on urban cycling and walking. Journal of the American Planning Association, 72(1), 33-42.

Kroesen, M., & Handy, S. (2014). The relation between bicycle commuting and non-work cycling: results from a mobility panel. Transportation, 41(3), 507-527.

Lawson, A. R., Pakrashi, V., Ghosh, B., & Szeto, W. Y. (2013). Perception of safety of cyclists in Dublin City. Accident Analysis & Prevention, 50, 499-511.

Legambiente (2017). 1st Report on Bike Economy and Urban Cycling in Italy.

Letarte, L., Pouliot, S., & Waygood, E. O. D. (2016). The Impacts of Universal Bus Pass on University Student Travel Behavior (No. 16-1498).

Li, X., Zhang, Y., Zhang, R., Lu, Y., & Xie, S. (2018). Overcoming Barriers to Cycling: Exploring Influence Factors of Cyclists' Preference in Free-Floating Bikesharing (No. 18-02707).

Litman, T. (2003). The Online TDM Encyclopedia: mobility management information gateway. Transport Policy, 10(3), 245-249.

Lois, D., Moriano, J. A., & Rondinella, G. (2015). Cycle commuting intention: A model based on theory of planned behaviour and social identity. Transportation research part F: traffic psychology and behaviour, 32, 101-113.

Loukopoulos, P. (2007). A classification of travel demand management measures. In Threats from Car Traffic to the Quality of Urban Life: Problems, Causes and Solutions (pp. 273-292). Emerald Group Publishing Limited.

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

Maldonado-Hinarejos, R., Sivakumar, A., & 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(6), 1287-1304.

Mandel, B., Gaudry, M., & Rothengatter, W. (1997). A disaggregate Box-Cox Logit mode choice model of intercity passenger travel in Germany and its implications for high-speed rail demand forecasts. The Annals of Regional Science, 31(2), 99-120.

Maness, M., & Cirillo, C. (2016). An indirect latent informational conformity social influence choice model: Formulation and case study. Transportation Research Part B: Methodological, 93, 75-101.

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

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

McFadden, D. (1974). The measurement of urban travel demand. Journal of public economics, 3(4), 303-328.

McFadden, D. (1978). Modelling the choice of residential location. Spatial Interaction Theory and Residential Location (Karlquist A. Ed., pp. 75-96).

McFadden, D. (1981). Econometric models of probabilistic choice. In Structural Analysis of Discrete Data with Econometric Applications. C. Manski and D. McFadden, Eds. MIT Press, Cambridge, MA.

McFadden, D. (1986). The choice theory approach to market research. Marketing science, 5(4), 275-297.

McFadden, D., & Train, K. (2000). Mixed MNL models for discrete response. Journal of applied Econometrics, 15(5), 447-470.

McKelvey, R. D., & Zavoina, W. (1975). A statistical model for the analysis of ordinal level dependent variables. Journal of mathematical sociology, 4(1), 103-120.

Meloni, I., & di Teulada, B. S. (2015). I-Pet Individual Persuasive Eco-travel Technology: A tool for VTBC program implementation. Transportation Research Procedia, 11, 422-433.

Meloni, I., Sanjust, B., Sottile, E., & Cherchi, E. (2013). Propensity for voluntary travel behavior changes: An experimental analysis. Procedia-social and behavioral sciences, 87, 31-43.

Menghini, G., Carrasco, N., Schüssler, N., & Axhausen, K. W. (2010). Route choice of cyclists in Zurich. Transportation research part A: policy and practice, 44(9), 754-765.

Mokhtarian, P. L. (2016). Presenting the Independence of Irrelevant Alternatives property in a first course on logit modeling. Journal of choice modelling, 21, 25-29.

Möser, G., & Bamberg, S. (2008). The effectiveness of soft transport policy measures: A critical assessment and meta-analysis of empirical evidence. Journal of Environmental Psychology, 28(1), 10-26.

Motoaki, Y., & Daziano, R. A. (2015). A hybrid-choice latent-class model for the analysis of the effects of weather on cycling demand. Transportation Research Part A: Policy and Practice, 75, 217-230.

Moudon, A. V., Lee, C., Cheadle, A. D., Collier, C. W., Johnson, D., Schmid, T. L., & Weather, R. D. (2005). Cycling and the built environment, a US perspective. Transportation Research Part D: Transport and Environment, 10(3), 245-261.

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

Muñoz, B., Monzon, A., & Lois, D. (2013). Cycling habits and other psychological variables affecting commuting by bicycle in Madrid, Spain. Transportation research record, 2382(1), 1-9.

Muñoz, B., Monzon, A., & López, E. (2016). Transition to a cyclable city: Latent variables affecting bicycle commuting. Transportation Research Part a: Policy and Practice, 84, 4-17.

Nadal, L. (2007). Bike sharing sweeps Paris off its feet. Sustainable transport, (19).

Namgung, M., & Jun, H. J. (2019). The influence of attitudes on university bicycle commuting: Considering bicycling experience levels. International Journal of Sustainable Transportation, 13(5), 363-377.

Nankervis, M. (1999). The effect of weather and climate on bicycle commuting. Transportation Research Part A: Policy and Practice, 33(6), 417-431.

Nikitas, A. (2018). Understanding bike-sharing acceptability and expected usage patterns in the context of a small city novel to the concept: A story of 'Greek Drama'. Transportation research part F: traffic psychology and behaviour, 56, 306-321.

Nkurunziza, A., Zuidgeest, M., Brussel, M., & Van Maarseveen, M. (2012). Examining the potential for modal change: Motivators and barriers for bicycle commuting in Dar-es-Salaam. Transport policy, 24, 249-259.

Oakil, A. T. M., Manting, D., & Nijland, H. (2016). Determinants of car ownership among young households in the Netherlands: The role of urbanisation and demographic and economic characteristics. Journal of transport geography, 51, 229-235.

Park, H., Lee, Y. J., Shin, H. C., & Sohn, K. (2011). Analyzing the time frame for the transition from leisure-cyclist to commuter-cyclist. Transportation, 38(2), 305-319.

Parkin, J., Wardman, M., & Page, M. (2008). Estimation of the determinants of bicycle mode share for the journey to work using census data. Transportation, 35(1), 93-109.

Paulssen, M., Temme, D., Vij, A., & Walker, J. L. (2014). Values, attitudes and travel behavior: a hierarchical latent variable mixed logit model of travel mode choice. Transportation, 41(4), 873-888.

Pedersen, T., Friman, M., & Kristensson, P. (2011). Affective Forecasting: Predicting and Experiencing Satisfaction With Public Transportation 1. Journal of Applied Social Psychology, 41(8), 1926-1946.

Pendyala, R. M., Kitamura, R., Chen, C., & Pas, E. I. (1997). An activity-based microsimulation analysis of transportation control measures. Transport Policy, 4(3), 183-192.

Pinjari, A. R., Eluru, N., Bhat, C. R., Pendyala, R. M., & Spissu, E. (2008). Joint model of choice of residential neighborhood and bicycle ownership: accounting for self-selection and unobserved heterogeneity. Transportation Research Record, 2082(1), 17-26.

Pinjari, A. R., Pendyala, R. M., Bhat, C. R., & Waddell, P. A. (2007). Modeling residential sorting effects to understand the impact of the built environment on commute mode choice. Transportation, 34(5), 557-573.

Pinjari, A. R., Pendyala, R. M., Bhat, C. R., & Waddell, P. A. (2011). Modeling the choice continuum: an integrated model of residential location, auto ownership, bicycle ownership, and commute tour mode choice decisions. Transportation, 38(6), 933.

Portoghese, A., Spissu, E., Bhat, C. R., Eluru, N., & Meloni, I. (2011). A copula-based joint model of commute mode choice and number of non-work stops during the commute. International Journal of Transport Economics/Rivista internazionale di economia dei trasporti, 337-362.

Prato, C. G., Bekhor, S., & Pronello, C. (2012). Latent variables and route choice behavior. Transportation, 39(2), 299-319.

Prochaska, J. O., & DiClemente, C. C. (1982). Transtheoretical therapy: toward a more integrative model of change. Psychotherapy: theory, research & practice, 19(3), 276.

Pucher, J., & Buehler, R. (2006). Why Canadians cycle more than Americans: a comparative analysis of bicycling trends and policies. Transport Policy, 13(3), 265-279.

Pucher, J., & Buehler, R. (2009). Cycling for a few or for everyone: The importance of social justice in cycling policy. World Transport Policy & Practice, 15(1).

Pucher, J., Dill, J., & Handy, S. (2010). Infrastructure, programs, and policies to increase bicycling: an international review. Preventive medicine, 50, S106-S125.

Pucher, J., Komanoff, C., & Schimek, P. (1999). Bicycling renaissance in North America?: Recent trends and alternative policies to promote bicycling. Transportation Research Part A: Policy and Practice, 33(7), 625-654.

Revelt, D., & Train, K. (1998). Mixed logit with repeated choices: households' choices of appliance efficiency level. Review of economics and statistics, 80(4), 647-657.

Richter, J., Friman, M., Gärling, T. Soft Transport Policy Measures: Gaps in Knowledge. International Journal of Sustainable Transportation, Vol. 5, No. 4, 2011, pp. 199–215.

Rietveld, P., & Daniel, V. (2004). Determinants of bicycle use: do municipal policies matter?. Transportation Research Part A: Policy and Practice, 38(7), 531-550.

Rodríguez, D. A., & Joo, J. (2004). The relationship between non-motorized mode choice and the local physical environment. Transportation Research Part D: Transport and Environment, 9(2), 151-173.

Rose, G. (2008). Encouraging sustainable campus travel: Self-reported impacts of a university TravelSmart initiative. Journal of Public Transportation, 11(1), 5.

Rotaris, L., & Danielis, R. (2014). The impact of transportation demand management policies on commuting to college facilities: A case study at the University of Trieste, Italy. Transportation Research Part A: Policy and Practice, 67, 127-140.

Rye, T. (2002). Travel plans: do they work?. Transport Policy, 9(4), 287-298.

Ryley, T. (2006). Use of non-motorised modes and life stage in Edinburgh. Journal of transport geography, 14(5), 367-375.

Sacks, D. W. (1994). Greenways as alternative transportation routes: a case study of selected greenways in the Baltimore, Washington area (Doctoral dissertation, M. Sc. Thesis, Towson State University, Towson, MD).

Saelens, B. E., & Handy, S. L. (2008). Built environment correlates of walking: a review. Medicine and science in sports and exercise, 40(7 Suppl), S550.

Sanjust, B., Meloni, I., & Spissu, E. (2015). An impact assessment of a travel behavior change program: a case study of a light rail service in Cagliari, Italy. Case Studies on Transport Policy, 3(1), 12-22.

Santos, G., Maoh, H., Potoglou, D., & von Brunn, T. (2013). Factors influencing modal split of commuting journeys in medium-size European cities. Journal of Transport Geography, 30, 127-137.

Scagnolari, S., Walker, J., & Maggi, R. (2015). Young drivers' night-time mobility preferences and attitude toward alcohol consumption: A Hybrid Choice Model. Accident Analysis & Prevention, 83, 74-89.

Schwanen, T., & Mokhtarian, P. L. (2005). What affects commute mode choice: neighborhood physical structure or preferences toward neighborhoods?. Journal of transport geography, 13(1), 83-99.

Schwartz, S. H. (1977). Normative influences on altruism. Advances in experimental social psychology, 10(1), 221-279.

Sears, J., Flynn, B., Aultman-Hall, L., & Dana, G. (2012). To bike or not to bike: seasonal factors for bicycle commuting. Transportation Research Record: Journal of the Transportation Research Board, (2314), 105-111.

Sener, I. N., Eluru, N., & Bhat, C. R. (2009). An analysis of bicycle route choice preferences in Texas, US. Transportation, 36(5), 511-539.

Sener, I., Eluru, N., & Bhat, C. (2009). Who are bicyclists? Why and how much are they bicycling?. Transportation Research Record: Journal of the Transportation Research Board, (2134), 63-72.

Shoup, D. (2008). The politics and economics of parking on campus. The implementation and effectiveness of transport demand management measures: an international perspective, 121.

Soto, J. J., Cantillo, V., & Arellana, J. (2018). Incentivizing alternative fuel vehicles: the influence of transport policies, attitudes and perceptions. Transportation, 45(6), 1721-1753.

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

Srinivasan, S., & Bhat, C. R. (2005). Modeling household interactions in daily in-home and out-of-home maintenance activity participation. Transportation, 32(5), 523-544.

Steg, L. (2005). Car use: lust and must. Instrumental, symbolic and affective motives for car use. Transportation Research Part A: Policy and Practice, 39(2-3), 147-162.

Steg, L., & Tertoolen, G. (1999). Sustainable transport policy: the contribution from behavioural scientists. Public money and management, 19(1), 63-69.

Steg, L., & Vlek, C. (1997). The role of problem awareness in willingness-to-change car use and in evaluating relevant policy measures. Traffic and transport psychology. Theory and application.

Steg, L., & Vlek, C. (2009). Encouraging pro-environmental behaviour: An integrative review and research agenda. Journal of environmental psychology, 29(3), 309-317.

Stern, P. C., Dietz, T., Abel, T., Guagnano, G. A., & Kalof, L. (1999). A value-belief-norm theory of support for social movements: The case of environmentalism. Human ecology review, 81-97.

Stinson, M. A., & Bhat, C. R. (2003). Commuter bicyclist route choice: Analysis using a stated preference survey. Transportation Research Record, 1828(1), 107-115.

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

Stinson, M. A., & Bhat, C. R. (2005). A comparison of the route preferences of experienced and inexperienced bicycle commuters. In TRB 84th Annual Meeting Compendium of Papers (No. 05-1434).

Stopher, P. R. Voluntary Travel Behavior Change. In Handbook of Transport Strategy, Policy and Institutions (Button, K. J., Hensher, D.A. ed.), Emerald Group Publishing Limited, Bingley, 2005, pp. 561–579.

Stopher, P., Clifford, E., Swann, N., & Zhang, Y. (2009). Evaluating voluntary travel behaviour change: Suggested guidelines and case studies. Transport Policy, 16(6), 315-324.

Swait, J., & Louviere, J. (1993). The role of the scale parameter in the estimation and comparison of multinomial logit models. Journal of marketing research, 30(3), 305-314.

Taniguchi, A., Suzuki, H., & Fujii, S. (2007). Mobility management in Japan: Its development and meta-analysis of travel feedback programs. Transportation research record, 2021(1), 100-109.

Taylor, D., & Mahmassani, H. (1996). Analysis of stated preferences for intermodal bicycle-transit interfaces. Transportation Research Record, 1556(1), 86-95.

Taylor, M. A., & Ampt, E. S. (2003). Travelling smarter down under: policies for voluntary travel behaviour change in Australia. Transport policy, 10(3), 165-177.

Tertoolen, G., Van Kreveld, D., & Verstraten, B. (1998). Psychological resistance against attempts to reduce private car use. Transportation research part A: Policy and practice, 32(3), 171-181.

Terza, J. (1985). Ordered Probit: A Generalization. Communications in Statistics – A. Theory and Methods, 14, 1–11.

Thigpen, C. G., Driller, B. K., & Handy, S. L. (2015). Using a stages of change approach to explore opportunities for increasing bicycle commuting. Transportation Research Part D: Transport and Environment, 39, 44-55.

Thorhauge, M., Haustein, S., & Cherchi, E. (2016). Accounting for the theory of planned behaviour in departure time choice. Transportation Research Part F: Traffic Psychology and Behaviour, 38, 94-105.

Ton, D., Duives, D. C., Cats, O., Hoogendoorn-Lanser, S., & Hoogendoorn, S. P. (2019). Cycling or walking? Determinants of mode choice in the Netherlands. Transportation Research Part A: Policy and Practice, 123, 7-23.

Tørnblad, S. H., Kallbekken, S., Korneliussen, K., Mideksa, T. K. Using Mobility Management to Reduce Private Car Use: Results from a Natural Field Experiment in Norway. Transport Policy, Vol. 32, 2014, pp. 9–15.

Train, K. E. (2009). Discrete choice methods with simulation. Cambridge university press.

Train, K. E., McFadden, D. L., & Goett, A. A. (1987). Consumer attitudes and voluntary rate schedules for public utilities. The Review of Economics and Statistics, 383-391.

Triandis, H. C. (1979). Values, attitudes, and interpersonal behavior. In Nebraska symposium on motivation. University of Nebraska Press.

Un-Habitat Core Team (2013). Planning and Design for Sustainable Urban Mobility: Global Report on Human Settlements 2013. UN-HABITAT, Routledge.

United States Environmentale Protection Agency (2017). <https://www.epa.gov/greenvehicles/fast-facts-transportation-greenhouse-gas-emissions>

Verplanken, B., Aarts, H., Van Knippenberg, A. D., & Moonen, A. (1998). Habit versus planned behaviour: A field experiment. British journal of social psychology, 37(1), 111-128.

Vij, A., & Walker, J. L. (2014). Hybrid choice models: The identification problem. In Handbook of Choice Modelling. Edward Elgar Publishing.

Vij, A., Carrel, A., & Walker, J. L. (2013). Incorporating the influence of latent modal preferences on travel mode choice behavior. Transportation Research Part A: Policy and Practice, 54, 164-178.

Vredin Johansson, M., Heldt, T., & Johansson, P. (2005). Latent variables in a travel mode choice model: Attitudinal and behavioural indicator variables. Department of Economics, Uppsala University.

Walker, J., & Ben-Akiva, M. (2002). Generalized random utility model. Mathematical social sciences, 43(3), 303-343.

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

Wardman, M., Tight, M., & Page, M. (2007). Factors influencing the propensity to cycle to work. Transportation Research Part A: Policy and Practice, 41(4), 339-350.

Williams, R. (2006). Generalized ordered logit/partial proportional odds models for ordinal dependent variables. The Stata Journal, 6(1), 58-82.

Winters, M., Brauer, M., Setton, E. M., & Teschke, K. (2010). Built environment influences on healthy transportation choices: bicycling versus driving. Journal of urban health, 87(6), 969-993.

Winters, M., Buehler, R., & Götschi, T. (2017). Policies to promote active travel: evidence from reviews of the literature. Current environmental health reports, 4(3), 278-285.

Witlox, F., & Tindemans, H. (2004). The application of rough sets analysis in activity-based modelling. Opportunities and constraints. Expert Systems with Applications, 27(4), 585-592.

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

Yamamoto, T. (2009). Comparative analysis of household car, motorcycle and bicycle ownership between Osaka metropolitan area, Japan and Kuala Lumpur, Malaysia. Transportation, 36(3), 351-366.

Yáñez, M. F., Raveau, S., & Ortúzar, J. D. D. (2010). Inclusion of latent variables in mixed logit models: modelling and forecasting. Transportation Research Part A: Policy and Practice, 44(9), 744-753.

Yu, J., & Beimborn, E. (2018). College Student Transit Pass Programs (No. Project J-7, Topic SA-42).

Zahran, S., Brody, S. D., Maghelal, P., Prelog, A., & Lacy, M. (2008). Cycling and walking: Explaining the spatial distribution of healthy modes of transportation in the United States. Transportation research part D: transport and environment, 13(7), 462-470.

LIST OF TABLES

LIST OF FIGURES