University of Basel

Faculty of Business and Economics

Master Thesis

Master of Science in Economics and Public Policy

Substituting the Traditional? Examining Usage Patterns of E-Bikers through Discrete Choice Modeling

Noémie Kipfer

2018-050-658

Sonnmattweg 8

4410 Liestal

noemie.kipfer@stud.unibas.ch

Supervisor

Prof. Dr. Beat Hintermann

June 7, 2024

Abstract

This master thesis explores which modes of transport are being replaced by e-bike users and for which specific stages e-bikes are used by using a discrete choice modeling approach and a binary logistic regression. Incorporating these substitution patterns, the potential reduction in CO_2 emission in Switzerland's transport sector due to the e-bike is calculated. First, this thesis analyzes the substitution patterns due to the e-bike using stated preference data for specific previously GPS-tracked stages of the E-Biking in Switzerland Project (EBIS). The e-bike mainly substitutes stages with the traditional bike. At the kilometer-level, the greatest substitution occurs for public transport (41.4%). Second, to gain a more nuanced understanding of the substitution of traditional modes of transport across socio-demographic groups, a Multinomial logit model (MNL) and a Mixed multinomial logit model (MMNL) are estimated. The results show that there are distinct preferences for substituted modes across individuals. These heterogeneous substitution preferences can be seen across age groups, genders, and languages, accounting for regional and cultural aspects, as well as urbanity levels. Stage-specific attributes such as travel time, travel cost, and the slope of a stage decrease the utility of all individuals for a substituted mode. Using these insights, it is shown that the e-bike has already decreased CO_2 emissions by up to 12.7%. The additional potential to decrease CO_2 emissions is estimated to be up to 10.3% across the population of Switzerland. This nuanced knowledge about different socio-demographic groups' substituted transportation choices can help inform policymakers to implement group-specific mobility strategies in order to reduce CO_2 in Switzerland's private transport sector.

Abbreviations

AIC	Akaike Information Criterion
ARE	Federal Office of Spatial Development
ASC	Alternative specific constant
BIC	Bayesian Information Criterion
DCM	Discrete choice model
EBIS	E-Biking in Switzerland Project
HH	Household
iia	Independence of irrelevant alternatives
iid	Independent and identically distributed
km	Kilometer
MMNL	Mixed multinomial logit model
MNL	Multinomial logit model
MOBIS	Mobility behaviour in Switzerland Study
MTMC 2021	Transport and Mobility Microcensus of Switzerland 2021
pt	Public transport
RP	Revealed preference
SP	Stated preference
VTT	Value of travel time
WTP	Willingness-to-pay

Contents

List of Figures

List of Tables

1 Introduction

The popularity of bicycles with electric assistance (e-bikes)[1] has surged globally, presenting a significant shift in mobility patterns, also in Switzerland (Velosuisse, 2023). Their growth is fueled by advancements in battery technology, increasing environmental awareness, and the desire for efficient, cost-effective mobility solutions. The adoption of e-bikes is not just an addition to the transportation landscape but has also been a subject of increasing interest, with studies indicating considerable potential for e-bikes to substitute travel with other modes (Andersson et al., 2021; Bigazzi and Wong, 2020; Kroesen, 2017; Moser et al., 2018; Reck et al., 2022; Sun et al., 2020). Furthermore, research has explored the role of e-bikes in overcoming barriers to bicycle use, shedding light on usage patterns and the potential for e-biking among different demographic groups, including the younger population (de Haas et al., 2021; Goel et al., 2021; Plazier et al., 2017; Yu et al., 2022).

Transport behavior, in general, is often subject to market inefficiencies due to various external costs and benefits linked to different transportation methods. Examples include traffic congestion, accidents, or environmental issues such as CO_2 emissions, of which the latter is the main interest of this thesis. Generally, these externalities can be described as external effects of the economic activity of an agent on society that are not considered by the agent, i.e., do not enter her utility function (Pigou, 1924). One way to reduce these externalities can be replacing trips or modes with higher external costs for other modes with lower external costs. The adoption of e-bikes presents a sustainable alternative to motorized transport, offering a reduction of CO_2 emissions for each trip in general (see, e.g., Sacchi and Bauer, 2023) and being able to have significant environmental benefits in terms of reduction of CO_2 (Bucher et al., 2019; McQueen et al., 2020; Moser et al., 2018; Philips et al., 2020, 2022). It is crucial to point out the relevance of the e-bike to more sustainable mobility strategies and its effect on external costs in the mobility sector when more individuals are switching from, e.g., cars and pt to choosing e-bikes as their transport mode of choice.

However, e-bike usage and its substitution effect have not been fully understood, especially in Switzerland. Reck et al. (2022) is calculating the emission effect per km based on the choice prediction of a mixed multinomial logit model. Nevertheless, this is only a prediction and knowledge about preferences ranking across the actual substitution has not been available so far. Also, a calculation of the impact of the e-bike concerning the potential of emission savings across the population of Switzerland has only been done so far by Bucher et al. 2019 and Moser et al. 2018. The former included commuting journeys in Switzerland.

[1] Note, that e-bikes are in this thesis defined as (s-)pedelecs, i.e., a bicycle with functional pedals which are assisted by an electric motor with speed up to 45 km/h.

They based their hypothetical scenarios, especially on different weather. Still, they do not know in particular which mode the e-bike substitutes. Moser et al. (2018) conducted a field experiment in Switzerland by giving an e-bike trial up to one year to 144 participants. They were mainly interested in the induced habitual behavior of the e-bike trial but did not track the actual travel behavior of the participants. Therefore, data on actually made trips are missing so far to conduct a more refined analysis of the underlying substituted mode of the e-bike and consequently its effect on CO_2 emissions.

To be able to determine the impact of the e-bike on CO_2 emissions in the mobility sector of Switzerland, several questions must be answered. First, knowing the substituted mode to calculate the emission savings is crucial. For example, if e-bikes mainly substitute trips that would have been made with a traditional bike before rather than with a car, the net environmental benefits could even be negative. This also applies to the case of when the e-bike induces a lot of new trips that would not have been made before. Second, the preference for the substituted modes can vary across individuals. As it is impossible to track and survey a whole population, it is necessary to account for different preferences of specific socio-demographic groups for the substituted modes. One can imagine that an individual being older than 60 years and traveling by e-bike would not substitute the same mode as a young individual traveling by e-bike. Last, not all trips are substituted; E-bike users still travel by car, pt, bike, or walk. Therefore, the probability of using an e-bike for a specific trip by a particular individual should be estimated.

Therefore, the primary question this master thesis aims to answer is:

> "What is the estimated impact of e-bike usage on CO_2 emissions reduction in Switzerland's private transport sector compared to a world without an e-bike adoption?"

To examine this overall research question, the thesis will process along three sub-questions as follows:

1. Which modes are being substituted by the e-bike?
2. Who is replacing which mode of transport with the e-bike?
3. Which stages are being substituted by the e-bike?

To examine these substitution patterns, the data is mainly sourced from the E-Biking in Switzerland Project (EBIS). The study is particularly interested in the behavior of e-bike users. The research team of the University of Basel and the ETH Zurich gathered GPS-tracking data on travel behavior via the app "Catch-my-Day" as well as stage-contextual and socio-demographic information of the participants through additional surveys during the study period between September 2022 and July 2023.

An overview of how these research questions will be approached is provided in Figure 1. The first research question will be addressed using percentage calculations based on the intermediate survey of the EBIS sample. It provides an indication of the underlying substituted modes of transport. In particular, for specific GPS-tracked e-bike stages in EBIS, additional insights on the trip purpose and the modes of transport used before owning an e-bike for that specific stage were gathered in an intermediate survey. Therefore, the substituted mode shares will be calculated, and a first insight into the potential emission reduction of these stages is provided (see the upper part of Figure 1).

Nevertheless, the second research question must be addressed in order to gain a more profound understanding of the user groups and their preferences for the substituted modes of transport. A discrete choice modeling method, which is widely applied in the transport mode choice literature (Ben-Akiva and Bierlaire, 1999; Brownstone, 2001), will be used in order to achieve this. The method allows for the comparison of choices among different modes of transport and the examination of trade-offs between different attributes, such as e.g., travel time across the modes of transport (Train, 2003). It accounts for the complex decision-making process of individuals and the simultaneous consideration of several alternatives of modes of transport (Ben-Akiva and Lerman, 1985). Also, it is able to predict probabilities of choices across substituted modes of transport for additional datasets (Train, 2003), which makes it a valuable approach for incorporating these substitution findings in subsequent models.

The collected data of the intermediate survey from EBIS is suitable to estimate a Multinomial logit model (MNL) and Mixed multinomial logit model (MMNL), as it accounts for travel time, travel cost, the availability of mobility tools, trip purposes, and socio-demographic information. The 4567 choice observations allow for the assessment of trade-offs between the substituted modes of transport and the elasticities to changes in stage-specific attributes. The stage-specific attributes travel time, travel cost, and slope for the non-chosen alternatives are added from different sources, as well as the weather variables influencing the substituted mode choice, as can be seen in the upper part of Figure 1. Subsequently, the estimated preferences for the substituted modes of transport of the EBIS sample are used to predict the substituted mode of transport for additional comparable GPS-tracked stages of EBIS. This transfer is necessary, to address the third sub-research-question of this thesis.

The third question on which trips are being replaced by the e-bike is approached with a binary logistic regression framework. For those stages detected to be done with the e-bike and not

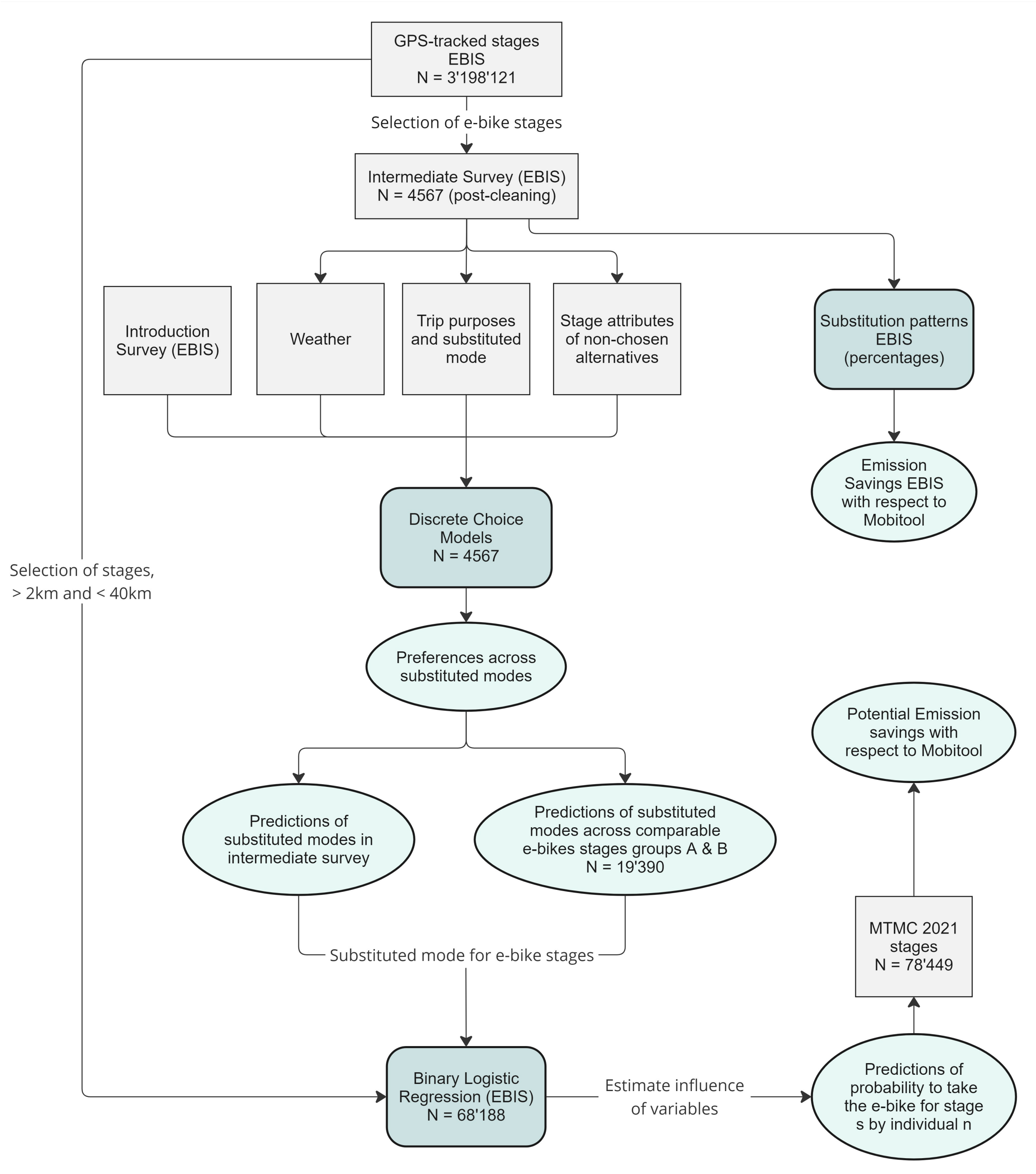


Figure 1: Overview of research design

queried in the intermediate survey, I use the predictions of the DCM to increase the number of included substituted stages in order to make the binary logistic regression model more robust and reliable. The estimated preference parameter of the individuals being part of the DCM sample, lets us predict the substituted mode of transport for the other GPS-tracked e-bike stages of these individuals. The mode of transport with the highest prediction of being chosen was taken as the substituted mode and imputed into the binary logistic regression. This procedure can be observed in the middle to lower section of the process flow chart in Figure 1.

The probability of conducting a stage with the e-bike is estimated with the depending variable on whether a trip was made with the e-bike (= "switched") or not for the EBIS participants. These predictions on the determinants of the choice to travel with the e-bike for a specific stage can be transferred to the Transport and Mobility Microcensus of Switzerland 2021 (MTMC 2021) sample. The MTMC 2021 contains a representative sample of the population of Switzerland, including information about the socio-demographic characteristics of households and individuals, their mobility tools, and their travel behavior (Federal Office for Spatial Development ARE, 2023). In transferring these predictions, the overall research question on the determination of the potential reduction in CO_2 emissions of Switzerland's transport sector due to the e-bike is assessed.

The findings of this master thesis contribute to the broader field of sustainable transportation and environmental policy, especially on the role of e-bikes in the mobility landscape. Understanding whether e-bikes are replacing non-motorized forms of transport like walking and cycling or cutting down on the use of cars and public transport (pt) depending on various socio-demographic factors is crucial to gaining a deeper insight and knowledge into the travel behavior of individuals in Switzerland. Furthermore, the knowledge about substituted transportation choices of different demographic groups can influence broader targeted and group-specific mobility or promotional strategies for e-bikes.

The results from the DCM show different preferences for the substituted modes of transport across socio-demographic groups of the participants of the E-Biking in Switzerland Project. In the Multinomial logit model (MNL), an increase in travel time is perceived negatively across all modes of transport but not equally in magnitude. An increase in travel cost also negatively impacts the choice probability for the substituted mode. The slope only influences the choice probability of the substituted mode bike, but not on walk. Age, language, and urbanity level of the place of residence are shown as the main drivers of different preferences across the modes of transport. Gender, Household (HH) size, and employment status have differing influences depending on the model specification. Furthermore, the Mixed multinomial logit model (MMNL) states clearly that the participants have significantly heterogeneous preferences for the substituted modes of transport walk, bike, and car but not for pt. Travel time of the substituted mode of transport is also perceived differently across the e-bikers in the sample.

Using these preferences, the probability of choosing the e-bike for a stage was estimated and transferred to the representative sample of the MTMC 2021. The results show that the mobility sector of Switzerland benefits from CO_2 emission reduction of 10.3% under the assumption that everyone would own an e-bike and would behave analogously to the individuals in the sample of the EBIS project.

The rest of this master thesis is organized as follows. The following Chapter 2 provides a literature review on e-bike adoption and substitution patterns as well as the potential of the e-bike on emission reduction in the transport sector. In Chapter 3, the data and data sources used in this thesis are described. Chapter 4 provides the methodological approach and the empirical strategy for the analysis in Chapter 5. The estimation results of the analysis are discussed in Chapter 6, as well as discussing limitations to the results and providing suggestions for further research. The thesis is concluded in Chapter 7.

2 Literature Review

E-bikes offer a way to make mobility more sustainable, reduce external costs, and provide an efficient alternative to cars and pt (Bucher et al., 2019; McQueen et al., 2020; Philips et al., 2022). Therefore, studies have been interested in the role of the e-bike, especially in the last ten years when sales increased in European countries (Bigazzi and Wong, 2020). This literature review systematically examines the findings concerning shifting transport mode choices due to the uptake of the e-bike. It delves into the factors influencing e-bike adoption and substitution patterns, including motivational and attitudinal aspects and topographical, such as the slope, and socio-demographic characteristics. Furthermore, this literature review focuses on the methodology of DCM within transportation studies. This approach is discussed in more detail in Chapter 4. Together with the binary logistic regression analysis, it is the primary method used in this thesis to understand and predict substitution patterns due to the e-bike. A crucial part will also be the findings from other studies concerning the environmental implications of the e-bike. The section will be finalized with a critical assessment of methodological gaps and how this thesis will contribute to the literature.

E-bike adoption: Motivation of e-bike ownership and usage

The motivation behind e-bike ownership and usage asked in surveys to e-bike users can already give insights into potential substitution effects (Dill and Rose, 2012; Jones et al., 2016; Popovich et al., 2014; Wolf and Seebauer, 2014). The reduction in physical demand compared to the traditional bike is a primary motivational factor (Dill and Rose, 2012; Popovich et al., 2014; Jones et al., 2016). Also, individuals who would not consider the traditional bike as part of their actual choice set of modes of transport are more open to using an e-bike instead of other transportation modes (Jones et al., 2016). However, the e-bike is not a general substitution for the ownership of a car as a mobility tool, but rather extending the choice set (MacArthur et al., 2014; Wolf and Seebauer, 2014).

The experimental approach of e-bike trial periods offers a further understanding of substitution of other modes of transport (Andersson et al., 2021; Cairns et al., 2017; Fyhri and Fearnley, 2015; Moser et al., 2018). For example, Moser et al. (2018) showed that habitual aspects influence e-bike uptake and usage by Swiss citizens. Trying the e-bike for a defined period can decrease car and motorbike usage, even in the long-term (Moser et al., 2018). Andersson et al. (2021) underline these findings: E-bike trials not only increased the number of e-bike trips and the respective distance traveled among Sweden employees, but also traditional cycling. The increase was due to a decreased car distance of 14 km per person and day, corresponding to 37% of car distance traveled (Andersson et al., 2021). Analogous findings among e-bike trial experiments were found in Norway and the UK (Cairns et al., 2017; Fyhri and Fearnley, 2015).

The shortcoming of these trials is that mostly no actual travel behavior was tracked, but only surveys were conducted on what the participants recalled from their trials. Furthermore, due to the cost intensity, samples are relatively small (up to 100 participants), which makes it hard to find reliable and generalizable effects of the treatment group. This generalization problem is even more accentuated when a potential selection effect exists when recruiting participants from one specific firm, as in the case of Cairns et al. (2017) and Andersson et al. (2021).

E-bike mode substitution

Using larger datasets can overcome the small sample problem to estimate robust substitution patterns. In the Netherlands, several studies were done based on a yearly national mobility survey. But, in terms of substitution patterns, the findings do not seem clear on what the e-bike does primarily substitute (de Haas et al., 2021; Kroesen, 2017; Lee et al., 2015). On the one hand, Lee et al. (2015) conducted a survey including a one-day travel diary of 217 e-bike users in the Netherlands. They concluded from percentage-calculations, that mainly trips with the car get substituted by the e-bike by up to 40%, while admitting that it does depend on the elderly people overrepresented in the sample. On the other hand, de Haas et al. (2021) and Kroesen (2017) showed different substitution effects. Kroesen (2017) applied a structural equation model and concluded that the e-bike strongly reduces traditional bike usage, followed by car and pt. Furthermore, e-bike ownership even reduced the ownership of traditional bikes while car ownership did not decline. Similarly, de Haas et al. (2021) concluded that mainly traditional bike trips get substituted followed by car and pt trips, but not as high as in Lee et al. (2015). But, for the trip purpose *Commute*, the car got substituted most frequently (de Haas et al., 2021). This finding is in line with Sun et al. (2020), who additionally propose that e-bike substitutes cars for shopping purposes.

Denmark is analogous to the Netherlands characterized by a prominent cycling culture. The results of a survey Haustein and Møller (2016) concerning mode substitution of the e-bike in Denmark showed that 64% of individuals agreed that they use the e-bike on trips when they would have used the traditional bike before, which underlines the findings in the Netherlands.

In general, these surveys are able to provide helpful insights into motivation and substitution effects. But this stated preference (SP) data is often prone to bias of the participants (Ben-Akiva et al., 1994; Morikawa, 1989). This is especially problematic when trying to estimate long-term effects such as Moser et al. (2018). Furthermore, often only a snapshot of one-day travel behavior is used to calculate substitution effects, which might not be representative of repeated travel behavior. Revealed preference (RP) data has the advantage that it tracks real travel behavior and is, therefore, prone to recall or social desirability bias (Morikawa, 1989). This data type has also been used by studies among trials of e-bikes to gather GPS-tracking data (Cairns et al., 2017; Andersson et al., 2021) but also on e-bike users that already purchase and regularly used an e-bike (Plazier et al., 2017). For example, Plazier et al. (2017) gathered GPS-tracking data as well as contextual information on 24 e-bike commuters in the Netherlands over two weeks. Among this small sample, the main substituted mode of transport was the car, even for longer trips. Their main explanation was an association of the e-bike with a positive experience overall. However, 24 participants are a very small sample which was observed across only two weeks. Therefore, these results might not be as robust as the results of a large sample (on the individual level and on the observation level) study.

GPS-tracking is a valuable tool to gather large datasets on travel behavior to predict mode substitution (Bigazzi and Wong, 2020; Reck, 2021; Reck et al., 2022). For example, Reck et al. (2022) used a MMNL with a focus on shared micro-mobility modes such as e-scooters as well as e-bikes to predict substitution modes in Switzerland. They showed that the substitution of a certain mode is dependent on the distance traveled. The shorter the distance, the more walking-km are substituted by personal and shared e-bikes, while for longer trips, car and pt are more often replaced (Reck et al., 2022). E-bike-sharing systems showed, in general, that these are not an alternative to cars (Bieliński et al., 2021; Fukushige et al., 2021).

Socio-demographic variables

In a previous section, we saw that different socio-demographic groups differ in their motivation for e-bike adoption, which is why it makes sense to include them in mode choice models as well. For example, gender is a common variable in transportation mode choice. The available literature mainly distinguishes between males and females and does not include other genders in the analyses. Thus, this thesis is also only distinguishing between males and females. The literature on the motivation of e-bike uptake and mode substitution already shows some significant differences across gender (see, e.g., Lee et al., 2015; Wolf and Seebauer, 2014). Also Woodward et al. (2021) highlights that the e-bike is used differently between male and female individuals. Age is shown to be an important determinant for e-bike ownership as well (Lee et al., 2015; MacArthur et al., 2014; Rérat, 2021; Sun et al., 2020; Van Cauwenberg et al., 2018; Wolf and Seebauer, 2014). While older people and female individuals are more likely to purchase an e-bike due to reduced physical exertion, younger e-bike users primarily buy an e-bike for utilitarian reasons and to replace car trips (Lee et al., 2015; MacArthur et al., 2014; Rérat, 2021; Sun et al., 2020; Van Cauwenberg et al., 2018; Wolf and Seebauer, 2014).

Also, e-bikers at the age of retirement use their car less but mainly replace the traditional bike (Lee et al., 2015; Sun et al., 2020).

Concerning the urbanity level of the place of residence, Rérat (2021) proposes that the e-bike also allows individuals from suburban and rural areas in Switzerland to be able to cycle more often. These areas are, in general, more motorized than the average, which lets the e-bike denote a less CO_2 emission-intensive alternative (Rérat, 2021). Also, Sun et al. (2020) shows that in the Dutch people living in non-urban areas are more likely to reduce car usage due to the ownership of an e-bike.

Trip purpose

The employment status is related to mode choice in general (Scheiner and Holz-Rau, 2007; Vega and Reynolds-Feighan, 2008), which makes the substitution effects of the e-bike on commuting trips a well-studied subject (Casier and Witlox, 2022; Dill and Rose, 2012; Fitch et al., 2022; Fyhri and Fearnley, 2015; Gao et al., 2021; Lee et al., 2015; MacArthur et al., 2018; Wolf and Seebauer, 2014). Among early adopters of e-bikes in Austria, the main purpose of using the e-bike was for leisure trips as the age was comparably high. Therefore, more emission-intensive modes of transport were barely substituted for commuting trips (Wolf and Seebauer, 2014). On the other hand, Fyhri and Fearnley (2015) concluded that e-bikes would have a greater effect on commuting mode choice than when doing a leisure-related trip. Mainly, two distinct user groups concerning trip purposes and mode substitution can be determined: There are utilitarian purposes such as commuting, which makes the e-bike a valuable alternative to traditional bikes and cars (MacArthur et al., 2018), and there are leisure purposes, as the e-bike also allows older people to still use a bike for, e.g., running errands (de Haas et al., 2021; Dill and Rose, 2012; Lee et al., 2015). Among commuters Fitch et al. (2022) studied the substitution effects of an e-bike-lending program of six months at Google. An increase of, on average, approximately two days per week of bike commuting was detectable and explained by a significant decrease in car and motorbike commuting trips (Fitch et al., 2022).

Trip- and alternative-specific attributes

Many trip-related attributes important for mode choice concerning cyclists have been identified in the literature. The literature highlights travel time, travel distance, slope, and, especially in route choice, cycling lanes, and traffic volume (see, e.g., Ha et al., 2020; Yang et al., 2018; Menghini et al., 2010). Travel time can be incorporated into discrete choice models as such but also measured as Value of travel time (VTT) (Truong and Hensher, 1985;

Bates, 1987). In transport economics, VTT is defined as the monetary value associated with travel time changes. Typically, a positive value is associated with it, suggesting a positive WTP for travel time savings. In other words, an individual is willing to accept higher costs to avoid an increase in travel time (DeSerpa, 1971). Also, the value of travel time is perceived differently across modes of transport as shown by, e.g., Schmid et al. (2021). Slope is particularly important for the modes of transport traditional bike, e-bike, and walking. For example, Meister et al. (2023) shows that an increase in slope does have a negative impact on the utility of the individuals using an e-bike or a traditional bike. However, using an e-bike lets individuals perceive slope as less exhausting compared to the traditional bike (Meister et al., 2023). Due to the scope of this thesis, infrastructural elements such as cycling lanes and traffic volumes will not be included.

Discrete choice models and substitution patterns in transport mode choice

DCM are a widely used approach in the transport mode choice literature and is suitable to look at the trade-offs between different discrete travel alternatives (Ben-Akiva and Bierlaire, 1999; Brownstone, 2001). Some studies are looking into the substitution patterns occurring from a specific mode of transport (Gao et al., 2023; Rahman and Baker, 2018; Reck et al., 2022). For example Rahman and Baker (2018) assessed the induced mode switch behavior due to a new flyover in Bangladesh by estimating a MNL. Reck et al. (2022) estimated substitution patterns concerning micro-mobility services based on predictions of a MMNL including a large dataset of GPS-tracked trips in Switzerland. They estimated that the personal e-bike substitutes mainly car trips (32%), followed by walking (26%), biking, and pt (21% each). On the km-level car was substituted with 41% and pt with 31%, followed by bike (17%) and walk (9%). Gao et al. (2023) used a discrete choice approach to estimate the transport mode substitution concerning bike-sharing systems in Shanghai. They calculated probabilities of travelers in Shanghai choosing different transport modes compared to bike-sharing systems. To the best of the author's knowledge, there is no study that investigates the effects of the e-bike on the mode substitution rates of other modes of transport using a discrete choice model approach.

Binary logistic regression and substitution patterns

Within the studies of transportation mode choice concerning, not only the MNL and MMNL are used to predict substitution patterns, but also binary logistic regressions for two discrete choices of alternatives, for example between the modes car and pt (Ahmed et al., 2020; Miletić et al., 2017; Puan et al., 2019; Witchayaphong et al., 2020; Youssef et al., 2021), between e-bikers and non e-bikers (Casier and Witlox, 2022; Jahre et al., 2019) or between active travel modes and non-active travel modes (Henning et al., 2020; Piatkowski et al., 2015). Casier and Witlox (2022) showed that weather conditions, trip time, and financial support by the employer for commuting by active travel modes were important determinants for using an e-bike in general. Nevertheless, some studies used the approach concerning the choice of traditional bike or not (Hu et al., 2021; Mohiuddin et al., 2022; Sears et al., 2012), which is the same concept as the choice between taking the e-bike or not. Similarly, Fearnley (2022) employed this approach in order to gain more insights into the substitution of other modes of transport due to e-scooters. However, a binary logistic approach concerning the alternative e-bike and another mode of transport has only been made so far by Ton and Duives (2021) between car and e-bike and by Bieliński et al. (2021) concerning (non-)usage of an e-bike sharing system. This master thesis aims to fill this gap by applying a binary logistic regression to examine the choice between taking the e-bike or taking another mode of transport.

Environmental impact of e-bikes

As the e-bike does have an impact on modal shift and is able to substitute the car for some trips, consequently, there is a vast literature on the e-bikes' impact on externalities in the transport sector, especially CO_2 emissions (Astegiano et al., 2019; Bucher et al., 2019; Goodman et al., 2019; McQueen et al., 2020; Philips et al., 2020, 2022; Piatkowski et al., 2015; Winslott-Hiselius and Svensson, 2017). Different approaches were used to induce these emission savings. For example, GPS data gathered from e-bike sharing systems showed that these shared e-bike providers help reduce emissions (Fukushige et al., 2023; Raposo and Silva, 2022). Also, the usage of personal e-bikes makes its contribution to reducing CO_2 emissions, as seen in Philips et al. (2022). They accounted for different urbanity levels of domiciles and socio-demographic characteristics in order to calculate CO_2 reductions due to the e-bike adaption (Philips et al., 2022). They estimate a maximum capability to reduce 24.4 million tons of CO_2 per year if everyone had an e-bike and would replace the travel mode car. This reduction corresponds to 0.58 tons per year per person, i.e., up to 50% in the transport sector of the UK (Philips et al., 2020, 2022). Also Bucher et al. (2019) predicted the greenhouse gas emission reductions in Switzerland due to the e-bike. Their prediction is rooted in energy demand connected to different weather scenarios. They estimate a reduction of up to 17.5% of the fossil fuel-based emissions of commuting-related trips due to the e-bike.

Scenario analysis based on historical adoption rates can also show the potential of e-bikes in terms of emission reduction (Astegiano et al., 2019), but they do not account for preferences across the population for specific modes of transport and for the fact that some are more likely to adopt to the usage of e-bikes as seen in, e.g., Kroesen (2017).

While these approaches are based on predictions, surveys can also help to gain findings on the potential of CO_2 emissions (McQueen et al., 2020; Winslott-Hiselius and Svensson, 2017). Winslott-Hiselius and Svensson (2017) conducted a web-based survey across e-bike users in Sweden. They show that mainly the car could be substituted, which is the major driver for an emission reduction of an average CO_2 emission reduction of 8.2kg per week and individual. This accounts for approximately 14–20% of the average total CO_2 per person from transportation (Winslott-Hiselius and Svensson, 2017). McQueen et al. (2020) are analogously using survey data and estimate a reduction of about 12% of CO_2 emissions caused in transportation in North America due to the e-bike. They explain this reduction with a main modal shift from cars to e-bikes. However, it is important to note that the e-bike does not only substitute trips, it can also induce new trips. This, in turn, does increase the CO_2 emissions again (Lee et al., 2015).

Discussion and conclusion

To gain insights into user preferences and usage patterns for underlying substituted modes of transport, often only percentages across the survey samples are calculated and compared in two-paired sample analysis (see, e.g., Bucher et al., 2019; Kroesen, 2017). This approach comes short as this method does not account for the complex decision-making process of individuals and the simultaneous consideration of several alternatives. Also, we do not understand which factors influence the substituted modes of transport the most. Furthermore, travel time, travel cost, or slope are often not included as explanatory variables to reflect the trade-offs between the alternatives. These trip and alternative specific attributes are crucial to account for different utilities provided through each alternative to the decision-maker (Ben-Akiva and Lerman, 1985). The prediction of substitution patterns using a DCM as Reck et al. (2022) offers some valuable insights. However, they do not include data on the preferred substituted mode, basing their findings on predictions.

To the author's knowledge, no DCM approach is applied to substituted modes of transport due to the e-bike in Switzerland. Therefore, this thesis contributes to the literature with a more refined approach to compare preferences across socio-demographic characteristics, including stage attributes concerning the substituted mode of the e-bike. There is currently also no study in Switzerland that examines the substitution patterns of the e-bike using SP data and uses these preferences to predict the probability of stages and individuals traveling by e-bike instead of another mode.

3 Data

3.1 The EBIS Project

The E-Biking in Switzerland Project (EBIS) aims to investigate the impact of the increasing popularity of e-biking on external costs in the mobility sector, such as CO_2 emissions, congestion, noise, and accidents. This is achieved through a randomized controlled trial (RCT) using extensive GPS-tracking and survey data to provide more accurate assessments of the climate impact of e-biking. Gaining these insights is essential for shaping transport policies and fostering exhaustive decarbonization efforts. Furthermore, the socio-demographic and mobility data gathered by EBIS during the study period between September 2022 and July 2023 are valuable resources for transport modeling and deriving mobility policy implications.

Building on the methodology of the earlier Mobility behaviour in Switzerland Study (MOBIS) study, which examined the effects of transportation pricing on travel behavior and external costs across the Swiss population using GPS data (Hintermann et al., 2024), EBIS narrows its focus to cyclists and e-bikers. The respective participants voluntarily contributed GPS data for the analysis of travel behavior, modes, and routes, enhancing the understanding of cycling behavior (Heinonen et al., 2023). Analogous to MOBIS, EBIS also explores the potential of financial incentives to promote mode choices with less negative external costs. However, in contrast to MOBIS, EBIS tries to encourage e-bike adoption among car users through an RCT. Initially, the travel behavior of all participants was observed without intervention. Later, a treatment group received a mobility budget, from which external costs incurred during travel were deducted, with the balance paid back at the study's conclusion. This method allowed for the examination of mobility pricing incentives without actual implementation. For a full overview of the methods of recruitment, sampling, and the dataset, see the methods and dataset article on the EBIS project by Heinonen et al. (2023).

The EBIS study design includes an introductory survey starting in September 2022, followed by a data collection phase through tracking to facilitate a revealed preference (RP) experiment focusing on the route and mode choices of cyclists (Phase 1). This is followed by a Randomized Controlled Trial (RCT) examining the influence of transportation pricing on mode choice (Phase 2). Additional surveys, including a retrospective survey on mode-shift and a stated preference survey concerning cycling infrastructure provisions, were done (Heinonen et al., 2023). Figure 2 shows a graphical representation of these phases.

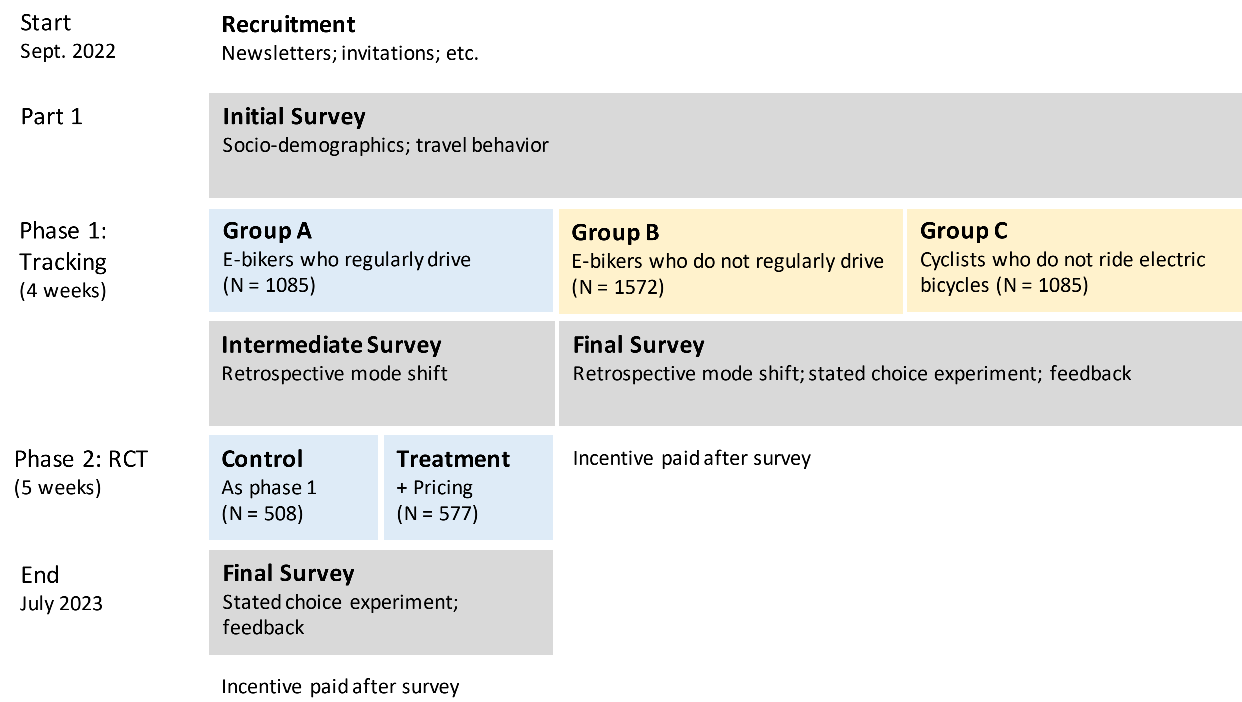


Figure 2: Overview of EBIS study design, fall 2022 recruitment wave. Based on Heinonen et al. 2023.

3.2 Data Description

The data used to investigate the research questions of this thesis is mainly from EBIS but additional external data had to be collected and added from multiple sources. An overview of the source is provided in Table 1. In the following, the different data sources and their processing in order to conduct the estimations in Chapter 5 are described. For an illustration of the process and inclusion of each of these data sources, see Figure 1 in Chapter 1. Note that only the main characteristics are represented in the following concerning the data of EBIS. The interested reader is referred to the “Conference Paper on Method and Dataset of EBIS” of Heinonen et al. (2023).

3.2.1 Introduction Survey EBIS

The introduction survey was done online via Qualtrics by all individuals who were willing to participate in the EBIS study. During the approximately 10-15 minutes of the general survey, the potential participants were asked about their socio-demographic characteristics, such as age, gender, education, citizenship, household size, and income (list not exhaustive), as well as their mobility tools available and their general travel behavior. A special focus was

Table 1: Data sources and characteristics

Data Source	**Type of data**	**Variables included**	**Method of data collection**
EBIS Introduction survey	Survey data	Socio-demographic variables, Availabilities of modes of transport (choice set generation)	Qualtrics questionnaire
EBIS Intermediate survey	Survey data, SP	Trip purposes, substituted mode choice	Qualtrics questionnaire
EBIS GPS-tracking data	GPS-Tracking data, RP	Length, duration, chosen mode of transport	GPS-tracking data from app "Catch-my-Day"
Visual Crossings	Weather data	Weather variables: Precipitation, Temperature (avg., min., max.)	API/ Data extraction from national weather services
Google Maps	Geographic Information System (GIS) data	Non-chosen alternatives: Travel time car, Distance car and pt	Own collected data, traffic data, user data (GPS)
OpenStreetMap (OSM): Geofabrik, BRouter	Vector data, crowdsourced data	Non-chosen stage attributes: Travel time walk, bike and pt, distance walk and bike, slope bike and walk	Crowdsourced contributions, aerial and satellite imagery, public datasets, external open databases (APIs)
MTMC 2021	Survey data	Socio-demographic variables of population, mobility tools availabilities, stages of one-day travel diary (and the respective attributes)	Random sampling online survey
Mobitool, based on Sacchi and Bauer (2023)	Emission life-cycle assessment of modes of transport	Emission per mode and km	Various, see Sacchi and Bauer (2023)

laid on their usage of traditional bikes and e-bikes (Heinonen et al., 2023).

3.2.2 GPS-Tracking Data

The revealed preference data of the stages and trips done by the participants was gathered through the app "Catch-my-Day" by MotionTag[2] between September 2022 and July 2023. After initialization, the app tracked automatically the travel behavior of the participants and identified the respective mode of travel. These modes of travel included car, bus, train, tram, subway, walking and cycling. However, due to similar routes and speeds, the algorithm could not reliably distinguish between e-bikes and regular bicycles (Heinonen et al., 2023, p. 6). The dataset initially consisted of 1.7 million stages conducted within Switzerland across all study groups (Heinonen et al., 2023, p. 11). This tracking data is used in this thesis for a) the stages surveyed in the intermediate survey described in the subsection below and b) to enrich the data used for estimation of the probability of substituting a specific stage with the e-bike, which will be described in Section 4.3.

3.2.3 Intermediate Survey EBIS

An essential component of EBIS for this thesis is the intermediate survey, also referred to as the retrospective survey. Up to five past stages, recognized as e-bike stages, were shown to the participants. Stages are unlinked trips, i.e., a stage is a continuous movement with one mode of transport as defined by Axhausen (2007). These stages can be aggregated to a trip, which can be a sequence of stages between origin destination pairs. For example, a trip to work may consist of walking to the car parked in a parking lot (stage 1), traveling with the car (stage 2), parking it in a parking garage, and walking to the work location (stage 3). Some of these specific stages were shown to the participants in study groups A and B after the conclusion of the GPS-tracking Phase 1. For groups B and C, this marked the final survey of the study, as these participants were not part of the RCT in Phase 2, as seen before in Figure 2. Both of these groups were also shown up to five past stages. However, as individuals in study group C do not own or regularly use an e-bike, they are not further relevant for this thesis and the following part is only concerned with groups A and B.

The specific stage is visually represented on a map, along with the travel date and time, as illustrated in Figure 3. This visualization ensures that participants can accurately recall the characteristics of that stage, minimizing recall bias. The stages shown to survey participants were selected based on the number of similar stages they completed during the study period. This ensures that the stages are representative of the entire study period. The selected stages had a minimum length of 2 km and a maximum length of 40 km. This selection was based on the imprecise GPS data below 2 km. Furthermore, the majority of the stages done with

[2] `https://motion-tag.com/`

an e-bike were also below 40 km. Then, several questions were posed concerning that stage, which are displayed in Table 2.

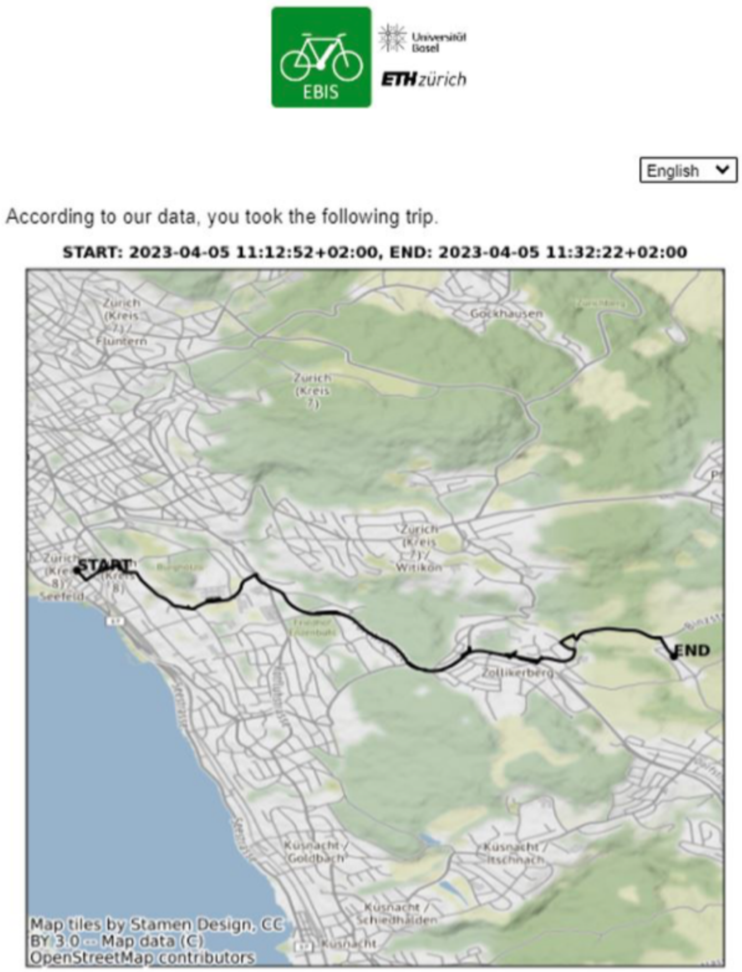


FIGURE 3: Example of retrospective trip visualization

As the app “Catch-my-day” could not perfectly distinguish the trips of the modes of transport traditional bike and e-bike (this accounts especially for electric assistance up to 25 km/h), the first question made sure that the stage was properly assigned to be an e-bike stage. There were already 1826 stages excluded out of 9895 of groups A and B due to wrongly detected e-bike trips or participants not remembering the mode of transport. The trip purpose was an additional information gathered for these specific trips. Nevertheless, the trip purpose “For recreation and exercise” was excluded from the estimation dataset as this purpose

Table 2: Intermediate survey questions: Trip purpose and substituted mode

1. Which mode of transport did you use for this trip?	**2. What was the purpose of the trip?**	**3. For the trip shown before, would you have traveled to the same destination this purpose before riding an e-bike?**	**4. For the trip shown before, which travel method would you have used before riding an e-bike?**
• Walking • Bicycle (without electric assist) • E-bike (25 km/h) • E-bike (45 km/h) • Other (please specify) • I don't remember	• Commuting (to work or school) • Shopping • Transporting children • Getting to another destination (Restaurants, Friends, Sport, Events, etc.) • For recreation or exercise • Coming home • I don't remember • Other (please specify)	• Yes, same destination • No, different destination, closer than the one shown • No, different destination, further away than the one shown • I would not have taken this trip prior to riding an e-bike • I don't know	• Walking • Bicycle (without electric assist) • car • Motorcycle • Local bus or tram • Suburban/regional train • E-bike (25 km/h) • E-bike (45 km/h) • Other (please specify)

would hardly ever be chosen to be done before by pt or car. Furthermore, the option *I don't remember* also ensures the assignment of the right trip purposes. However, as we only asked for specific stages and not a whole trip, the trip purpose assignment seemed not straightforward to some participants who were connecting several purposes. For example, a woman stated that she is commuting home from work, but during her travel, she has been going to the supermarket and taking the children home. Therefore, the trip purpose is not clearly identifiable. As trips can include several purposes, stages hardly can't. However, the determination of one stage purpose does not necessarily have to be the same as the main purpose of the whole trip. In our example, the main purpose of this woman would probably be commuting and would be assigned to this stage purpose.

The third question in Table 2 was included in order to distinguish substituted trips from newly induced trips by the e-bike. Consequently, the newly induced trips were excluded from the dataset used for the estimation as they do not have a substituted mode. The fourth question concerning a specific stage is used as the dependent variable in the DCM. The modes of transport "Local bus or tram" and "Suburban/regional train" were aggregated into the mode group pt. Furthermore, motorcycles were excluded from the dataset for the estimation

(Number of choices = 108) as it could not be aggregated to the four main groups walking, cycling, pt and car and not be detected clearly due to similar speeds as cars. An additional approximately 800 choice observations were excluded due to errors in the computation of stage attributes and weather data or missing information on crucial socio-demographic attributes. As we have several choices for the substituted mode for each participant, we obtain a panel data structure in the intermediate survey dataset (Train, 2003). This is insofar important as we have to keep some coefficients (such as person-specific attributes) constant over choice situations, which enter the utility function for each choice (Train, 2003). After cleaning the data, a total of N = 4567 choice observations of 1424 individuals could be estimated.

3.2.4 Weather Data

The historical weather data for the specific stages was fetched via API from the weather data platform "The Visual Crossing"[3]. The weather variables were requested based on the latitude and longitude of the starting point of each stage at the specific date. The variables maximum temperature, minimum temperature, precipitation probability (0 if no precipitation and 1 if there was any precipitation on that day), and precipitation coverage as a percentage of the hours per day were gathered for each recorded stage based on the nearest weather station.

3.2.5 Stage Attributes of Non-Chosen Alternatives

For the estimation of the DCM, it is necessary to know the stage attributes of the chosen mode and the non-chosen modes of transport to determine the trade-offs of a decision-maker when making his or her choice. For the stage attributes travel time, travel distance, and the slope of walk and bike, BRouter, which is based on OpenStreetMap[4], was used to estimate these variables for each stage. The stage attributes of pt were generated by incorporating the pt-timetable of the years 2022 and 2023. For cars, the API of Google Maps provided reliable data on the mentioned attributes, including estimated delays due to congestion.

3.2.6 Transport and Mobility Microcensus of Switzerland 2021 (MTMC 2021)

The Federal Office of Spatial Development (ARE) surveys every five years the travel behavior across a representative sample of Switzerland's population with a size of over 55'000 respondents, called "Transport and Mobility Microcensus of Switzerland 2021 (MTMC 2021)" (Federal Office for Spatial Development ARE, 2023). It consists of various information on ownership and availability of vehicles and subscriptions of, e.g., pt-passes or car-sharing systems,

[3] `https://www.visualcrossing.com/weather-api`
[4] `https://www.openstreetmap.org/`

daily distances traveled and time taken as well as the used means of transport and mobility purposes as well as socio-demographic characteristics. The survey of 2021 was planned for 2020, but due to the major disruption in travel behavior caused by the Covid-19 pandemic, the survey was postponed by one year (Federal Office for Spatial Development ARE, 2023). This data source was used to predict potential e-bike usage among Switzerland's citizens in order to calculate the potential for CO_2 emission savings. To make the data comparable to the EBIS dataset, the person-specific attributes of the MTMC 2021 were categorized analogously. The comparison across the shares in the person-specific attributes can be seen in the next chapter in Table 3. Furthermore, the stages of the reported travel behavior on a randomly specified qualifying date were filtered to be at least 2 km and a maximum of 40 km long. The chosen modes were aggregated into the four main groups of the intermediate EBIS-Survey and all other modes excluded. This procedure leaves the dataset with nearly 70'000 choice observations of the participants in the MTMC 2021 sample, as seen in Figure 1.

3.3 Representativeness

The unweighted EBIS sample of groups A and B overrepresents middle-aged people and underrepresents younger and older people compared to the MTMC 2021. This finding is not surprising as the MTMC 2021 aims to build a representative sample of the population of Switzerland and does not necessarily build on e-bike users. However, the subsample with households having an e-bike in the MTMC 2021, indicated in the last column of Table 3, shows that there are still differences in shares between the EBIS sample and individuals owning an e-bike in the MTMC 2021 sample. Concerning citizenship, a similar pattern is found between the EBIS sample and the MTMC 2021: Swiss citizens more often own an e-bike than non-swiss citizens. However, the shares in the EBIS sample are slightly more diverging than in the MTMC 2021 and the e-bike-owning subsample. The EBIS sample clearly overrepresents individuals with a tertiary education compared to Switzerland's (e-bike) population, which is not the case per se for employment status. We can also see that the share of female individuals participating in the EBIS study is lower than in the Swiss population. The EBIS sample also overrepresents individuals with a household income over 10'000 CHF, which could be correlated to the high education share. There are also clearly fewer people giving no answer to this question in the EBIS survey than in the MTMC 2021. In both samples, most participants live in household sizes with 1, 2, or 3 individuals. The German language is overrepresented in the EBIS sample, which is not surprising as there was no recruiting done in the Italian-speaking part of Switzerland. More individuals living in the city have an e-bike and took part in the EBIS study than the population of Switzerland. The

Table 3: Socio-demographic variables in intermediate survey data (EBIS) and MTMC 2021, in percentages

Variable	Category	EBIS Group A&B	MTMC 2021	MTMC 2021 E-bike owner
Age	Young (16-29 years)	11.17	18.35	17.61
	Middle (30-59 years)	70.37	52.38	55.13
	Old (60+ years)	18.47	29.27	27.27
Citizenship	Swiss	80.97	59.35	71.88
	Non-Swiss	19.03	40.65	28.12
Education	Tertiary	82.44	37.93	43.24
	Mandatory/Secondary	17.56	59.56	55.48
Employment Status	Employed	78.65	65.79	71.68
	Not Employed*	21.35	34.21	28.32
Gender	Female	39.47	50.98	50.63
	Male	60.32	49.02	49.37
HH Income	<= 10'000 CHF	38.69	52.97	43.56
	> 10'000 CHF	46.98	25.61	35.29
	No Answer	1.90	21.41	21.14
HH Size	1, 2 or 3	60.88	74.44	67.42
	>= 4	39.12	25.56	32.58
Language	German	86.31	68.94	78.93
	French	9.62	25.18	17.89
	English	4.07	0.00	0.00
Urbanity Level	City	74.86	48.59	50.35
	Non-City	25.14	16.15	20.36
Access Car	No	19.45	23.68	15.22
	Sometimes	12.43	16.05	18.95
	Yes	68.12	60.27	65.83
Access Bike	No	22.19	31.43	8.07
	Sometimes	1.12	9.50	7.16
	Yes	76.69	59.07	84.77

* including Retired, Unemployed, Student, and Other.

access to a car is comparable, which is also true for the access to a bike. These differences between groups A and B of the study sample and the general population of Switzerland are crucial as they might influence the generalization of the results and will be discussed in Chapter 6.

In the analysis of the shares across socio-demographic variables, we can see in the Appendix Table A1 that certain categories show very low shares compared to others, which can result in high standard errors and reduced statistical power in the estimation. Especially since the dataset of the intermediate survey contains only 1424 individuals. Therefore, some variable categories were aggregated to increase the reliability and statistical power of the models. Furthermore, aggregating categorical variables such as household income, household size, education, urbanity level, or employment status into a binary categorization helps for a more straightforward interpretation of results.

4 Methodology

This Chapter consists of three main parts. The first part delves into the theory behind the discrete choice models used in this thesis and the main aspects of the goodness-of-fit measures. The second part is dedicated to substitution on the stage level, i.e., the binary logistic regression. The third part will address these models in the empirical strategy, which will provide the foundation for the analysis in the next section.

4.1 Discrete Choice Modeling

This Chapter describes the main rationale behind the method of discrete choice. Furthermore, the Multinomial logit model (MNL) and Mixed multinomial logit model (MMNL) used in the estimation in Chapter 5.3 will be presented.

Discrete choice modeling is a method commonly used in various fields, including transportation research, to systematically analyze, understand, and predict discrete decisions made by individuals or entities (Ben-Akiva and Lerman, 1985). Discrete choice models are based on the idea that decision-makers select from a *finite* set of *mutually exclusive* options. No alternatives outside of this choice set are considered (Train, 2003, p. 15). There are four key components which we need to make assumptions about in order to develop a discrete choice model:

- **The decision-maker** n, which is an individual or entity or a group (e.g., a household) with underlying characteristics;
- **The alternatives** j, which form the choice set of each decision-maker;
- **The attributes**, which determine the characteristics of each alternative that the decision-maker considers;
- **The decision-rule**, which describes the process of getting to an actual choice.

(Bierlaire, 1998, p. 3)

In a mode choice model, the decision-maker is an individual n defined by underlying characteristics that are assumed to influence the mode choice, such as age, gender, household income, or citizenship. All alternatives available to the individual form the choice set, which varies across individuals (Bierlaire, 1998). For example, if an individual n owns a car, has access to pt, and is able to walk but does not own a bike, it is assumed that she would not consider the alternative bike as part of their choice set to travel from destination A to destination B.

To make an actual choice, the individual n considers each alternative's attributes. In a transport mode choice model, these are, for example, travel time, travel cost, the weather, or the upward and downward slope. Note that not all attributes are relevant for all alternatives. For example, an individual might consider the upward and downward slope in their choice of bike and walking as the exertion is considered as a negative trait, but it is not as relevant when traveling by car or pt. Considering all of these factors, an individual will make a choice based on an underlying decision-rule. Discrete choice models are typically derived under the assumption that the decision-maker is maximizing her utility (Train, 2003). The model estimates the likelihood of each alternative being chosen, relying on observed attributes of each alternative and the socio-demographic characteristics or attitudes of the decision-makers and unobservables and uncertain factors, such as alternative attributes or taste variations (Bierlaire, 1998). Therefore, each alternative's utility has a deterministic component captured by the model and a random component, which accounts for unmeasured influences - which is the basis of the Random Utility Model (RUM) (Manski, 1977). The random utility theory assumes that a decision-maker n gets an amount of utility U from each alternative j that he or she chooses in choice situation t as follows:

$$U_{njt} = V_{njt} + \varepsilon_{njt} \quad \forall j, t. \tag{1}$$

where

- U_{njt}: Utility function for decision-maker n and alternative j in situation t
- V_{njt}: Representative utility - Observed by the modeler, which represents attributes of the alternative j and the characteristics of the decision-maker n in a situation t
- ε_{njt}: Error term - Unobserved by the modeler, which is treated as a random variable

As a decision-maker is assumed to maximize her utility, she chooses alternative i if and only if $U_{nit} > U_{njt} \forall j \neq i$ (Bierlaire, 1998). As it is not possible to actually measure all factors influencing the total of utility U of each alternative j, V_{njt} is the representative utility, while ε_{njt} captures the aspects that affect utility but are not included as we are not able to measure them. Examples could be people experiencing less comfort or safety on bikes compared to a stage by car.

As discrete choice models are designed to analyze decisions where individuals choose one option from a discrete and finite set of alternatives, it is the suitable approach to analyze the discrete transport mode choice of the substituted mode of the e-bike. A discrete choice model, therefore, provides insights into the decision-making process of individuals, revealing why certain substituted modes of transport were preferred over others based on various factors (Ben-Akiva and Bierlaire, 1999). Furthermore, such an approach can catch and interpret categorical variables more precisely compared to a regression framework, which performs better at continuous variables (Train, 2003).

4.1.1 Multinomial Logit Model (MNL)

The Multinomial logit model (MNL) assumes ε_{njt} to be Independent and identically distributed (iid). Independent means that the unobserved factors are not correlated and do not affect the utility of another alternative. Identically means that the unobserved factors for each stage are drawn from the same probability distribution, which is in the case of the MNL a Gumbel (or Extreme Value Type I) distribution.

The probability P of an individual n choosing alternative j in situation t is defined as

$$P_{njt} = \frac{e^{V_{njt}}}{\sum_i e^{V_{njt}}} \tag{2}$$

Therefore, the choice probabilities are proportional to the exponentials of the utilities of the alternatives, which implies a very strong assumption of the MNL: That the choice probability ratio between any two alternatives is unaffected by the presence or attributes of any other alternative, known as the Independence of irrelevant alternatives (iia). Consequently, it is assumed that if an attribute of an alternative changes or a new alternative is added to the choice set, this would change the ratio of any two existing alternatives proportionally (Brownstone and Train, 1998).

Therefore, the choice probabilities are proportional to the exponentials of the utilities of the alternatives, which implies a very strong assumption of the MNL: That the choice probability ratio between any two alternatives is unaffected by the presence or attributes of any other alternative, known as the Independence of irrelevant alternatives (iia). Consequently, it is assumed that if an attribute of an alternative changes or a new alternative is added to the choice set, this would change the ratio of any two existing alternatives proportionally (Brownstone and Train, 1998).

These strong assumptions necessarily impose limitations on this model approach. First, the assumed substitution pattern is often not reflecting reality accurately and might be considered too restrictive (e.g., Red-bus/Blue-bus paradox as seen in McFadden (1972)), but it simplifies computation and interpretation of the MNL. Second, the model can represent a systematic preference variation but not a random taste variation. When there is unobserved heterogeneity of preferences, then this is not accurately reflected in the MNL. Last, in the case of panel data, where we observe the choice of one individual several times, the iid assumption of the MNL does not account for any correlation in choices across individuals (McFadden and Train, 2000).

4.1.2 Mixed Multinomial Logit Model (MMNL)

The MMNL is an extension of the MNL. It relaxes the iid assumption and accounts for heterogeneous preferences across individuals. This implies that for more realistic substitution patterns between alternatives can be accounted for. For example, we can account for travel time valuation across individuals n, or how different individuals value travel time savings in a mode choice setting: In the MNL, an increase of 1 minute of travel time of a mode is estimated to reduce the choice probability to travel with that mode for all individuals by, e.g., 10%. But, in a MMNL setting, the model might estimate a disutility of travel time across all individuals and all modes, but the magnitude of the disutility varies. For example, employed individuals will experience a greater disutility compared to retired individuals as the former have limited time resources left for traveling. The error term is able to catch this variation in taste[5].

According to Train (2003, p. 139), the choice probability of each alternative j chosen by individual n in situation t is specified as:

$$P_{njt} = \int L_{njt}(\beta) f(\beta) d\beta \tag{3}$$

while the logit Probability $L_{njt}(\beta)$ is the logit probability estimated at parameters β:

$$L_{njt}(\beta) = \frac{e^{V_{njt}(\beta)}}{\sum_{i=1}^{I} e^{V_{nit}(\beta)}} \tag{4}$$

If utility is expressed as a linear function of β, then

$$V_{njt}(\beta) = \beta' x_{njt} \tag{5}$$

with $\beta' x_{nit}$ being the scalar product, where the scalar is the linear predictor part of the model. It is the log-odds of the influence of the explanatory variables of choosing a particular option. If this linearity is specified, then the mixed logit probability turns into its standard form. As the density of $f(\beta)$ is specified to be continuous in this thesis, the choice probability P_{njt} is

[5]For a similar example, see Brownstone and Train (1998)

defined as:

$$P_{njt} = \int \left(\frac{e^{\beta' x_{njt}}}{\sum_j e^{\beta' x_{nit}}} \right) \phi(\beta \mid \mu, \sigma) \, d\beta \tag{6}$$

where the density $f(\beta)$ is specified to be normal with mean μ and covariance σ (indicated as $\phi(\beta \mid \mu, \sigma)$ in equation 6). This specification allows the coefficient β to be a random variable and, therefore, to vary across individuals, which is the main difference to the MNL. The MMNL is thus able to capture unobserved preference heterogeneity as the mixed logit probability is a weighted logit choice probability - weighted by $\phi(\beta \mid \mu, \sigma)$. $f(\beta)$ is specified to be normally distributed, except for travel cost, where we assume that every decision-maker perceives that coefficient as strictly negative and thereof follows a log-normal distribution (Train, 2003).

4.1.3 Goodness-of-fit Measures and Post-Estimation Values

The following section gives a brief overview of the goodness-of-fit measures for DCM and the different post-estimation parts.

Goodness-of-fit

To assess how well the models fit the data, the most known statistics ρ^2, also known as the McFadden R^2 or the *likelihood ratio index* as in Train (2003), and the adjusted ρ^2 measure how well the model with its estimated parameters performs compared to a model without parameters Train (2003). They are defined as:

$$\begin{aligned} \rho^2 &= 1 - \frac{LL(\hat{\theta})}{LL(0)} \\ \text{Adjusted } \rho^2 &= 1 - \frac{LL(\hat{\theta}) - K}{LL(0)}. \end{aligned} \tag{7}$$

where $LL(\hat{\theta})$ is the log-likelihood of the estimated model, and $LL(0)$ the log-likelihood of the null model, i.e., the model with zero parameters, and K is the number of parameters used in the θ-model. The ρ^2 can take values between 0 and 1, whereas $\rho^2 = 0$ would indicate that the model with parameters does not predict the data any better than a model without any parameters, and $\rho^2 = 1$ predicts the data perfectly (Train, 2003). The adjusted ρ^2 prevents the model from overfitting as it penalizes for the number of estimated parameters. In addition, there are two information criteria, Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) (also known as Schwarz (1978) Information Criterion), that are important for the evaluation of the model. They are defined as:

$$\begin{aligned} AIC &= -2LL(\hat{\theta}) + 2K \\ BIC &= -2LL(\hat{\theta}) + K\ln(N) \end{aligned} \tag{8}$$

where K is the number of estimated parameters, and N is the sample size. With a higher value of $LL(\hat{\theta})$, both criteria are correcting for an increase in the number of parameters. However, the BIC penalizes the number of estimated parameters more, especially in large sample sizes. Both criteria aim to balance model fit and model complexity while modeling the information in the data (Burnham and Anderson, 2004). In the case of BIC there is a tendency to underfitting, therefore the combination of these two criterions in evaluating the model fit is key.

Post-estimation and predictions

The elasticity expresses the sensitivity of the choice probability of an alternative j due to a change in attribute k. Mode-specific elasticities are the change in probability P_j following a change in the attribute X_j^k, while cross-elasticities are the change in probability P_j following a change in an attribute of another mode X_i^k. These elasticities can be calculated at both the individual and aggregate levels, but are only used at the aggregate level in this thesis, as follows:

$$E_{jX_{kj}} = \frac{\partial P_{kj}}{\partial X_{kj}} \frac{X_{kj}}{P_{kj}} \tag{9}$$

$$E_{jX_{ki}} = \frac{\partial P_{kj}}{\partial X_{ki}} \frac{X_{ki}}{P_{kj}} \tag{10}$$

where E_j represents the elasticity of the probability of alternative j with respect to the attribute X_{kj} (Train, 2003).

4.2 Stage Substitution: Binary Logistic Regression

Individuals for whom the e-bike is part of their choice set, i.e., owning or having access to an e-bike, do not always choose it as their mode of travel. Therefore, we can model the travel choice with more than two discrete choice possibilities as in the MNL or MMNL, but also as a binary discrete outcome, i.e., a stage is substituted by the e-bike or not. This is a typical example of a binary response, which is discussed in this section.

In a binary response setting, the dependent variable Y has two possible outcomes, i.e., is dichotomous:

- $Y = 1$, if the event of interest occurs;
- $Y = 0$, otherwise.

A binary logistic regression simultaneously determines the influence of the selected explanatory variables to predict the probability of $Y = 1$ over the probability of $Y = 0$ (Stock and Watson, 2019, Ch. 11). Performing a logistic transformation of the odds is the dependent variable, defined by the explanatory variables as:

$$log(P(Y_{nt} = 1 \mid X_{nt}, Zn)) = ln(\frac{P_{nt}}{1 - P_{nt}}) = \alpha_n + \beta_{nt} * X_{nt} + \gamma_n * Z_n \tag{11}$$

where

- P_{nt} = the probability of the event $Y_{nt} = 1$ for decision-maker n at choice situation t.
- α_n = the individual-specific constant. It captures unobserved heterogeneity and allows each decision-maker to have a unique baseline preference of the binary choice, which is assumed to be time-invariant.
- X_{nt} = a vector of time-varying variables for decision-maker n at choice situation t, with β as the corresponding vector to be estimated. In a mode choice model, these are the stage-specific attributes such as travel time, travel cost, slope, or situational variables (e.g., weather, trip purpose).
- Z_n = a vector of variables for decision-maker n, with γ_n to be estimated. These include, for example, socio-demographic variables and are assumed to be time-invariant.

The underlying assumption concerning the actual choice is the same as in a MNL or MMNL: The decision-maker n is assumed to make its binary choice to maximize her utility based on the stage attributes. She therefore gets an amount of utility U from making the choice to do a stage with the e-bike and to do it with another mode, which can be specified as:

$$U_1 = \beta' X_1 + \epsilon_1 \quad \text{(utility from choosing the e-bike)} \tag{12}$$

$$U_2 = \beta' X_2 + \epsilon_2 \quad \text{(utility from choosing another mode)} \tag{13}$$

where U_1 and U_2 represent the utilities that the decision-maker n derives from each transport mode, where X_1 and X_2 are vectors of attributes of each stage and each individual. Analogous to MNL and MMNL, there is an observed and an unobserved part of the utility. The latter is specified as ϵ_1 and ϵ_2 being random error terms, assumed to be logistically distributed (Train, 2003, p. 4).

4.3 Empirical Strategy

This section applies conceptually the DCM and binary logistic regression framework to the specific research question, incorporating the operationalization of the essential variables. It defines the estimation process in Chapter 5 and how stage- and person-specific attributes are measured to examine their influence on the substituted mode of transport before using an e-bike. It also describes the binary logistic regression approach used to estimate the impact of the explanatory variables on the log-odds of using an e-bike for a given stage.

4.3.1 Discrete Choice Models

Choice Sets

The potential choice set of each individual consists of the transportation modes *walk*, *traditional bike*, *car*, and *pt*, as defined in Section 3.2. The individual choice sets were generated based on the availability of modes of transport to the individuals. For example, if individual n owns a traditional bike or had at least sometimes access to it, the traditional bike was part of the choice set of individual n. If that individual stated that she had no access to a traditional bike, then it was not part of the choice set, i.e., the availability was set to 0. The same procedure was applied to generate the availability indicator of the car, i.e., whether the car was part of the choice set of individual n. Pt was assumed to be available to all individuals when a specific stage allowed for a pt connection. Whether walking was part of the choice set was defined based on the length of the stage. For all stages up to 8 km, I assumed that walking is part of the choice set of each individual n. For all stages being above that distance was assumed that walking is not a plausible option, as the travel time would take approximately two hours with an average speed of 4km/h.

A few exceptions were made to this availability indicator due to the nature of the study of the substituted mode, i.e., the mode before owning an e-bike. As the participants were asked about their substituted mode, it is conceivable that the e-bike substituted a mode not only for a specific stage (i.e., they still had the choice to take the traditional bike because they had one at home) but also in terms of mobility tools available to the individual n (i.e., she sold her traditional bike upon purchasing the e-bike). Therefore, if an individual n stated that she did not have access to a traditional bike at the beginning of the study (i.e., would not consist of the choice set by default) but declared in the intermediate survey that they would have done a stage with the traditional bike before owning an e-bike, then the traditional bike is still part of the choice set for the substituted mode. The same procedure was applied to car. However, for participants under 18 years, the mode car was still available, as they could be driven before by another individual, e.g., their parents. The e-bike could then be a substitute for this mode of transport, even when the individual only has limited impact on the actual mode choice in the case of car. However, as only two individuals below 18 years are included in the sample, the effect is negligible.

Attributes of each stage and alternative

For each stage and alternative, the attributes *travel time*, *travel cost*, and *slope* were included. The DCM is estimated on stage-level data. As described in Section 3.2, one trip can include several stages and, thereof, can include several modes of transport. Consequently, it would not be possible to define one substituted mode for a trip - there might be several. The travel time for the chosen alternative was given by the GPS track while those for the non-chosen alternatives were computed with different data sources as seen in Chapter 3.

Travel costs are defined by the additional marginal costs for each mode per stage traveled. This is assumed to be zero for the modes of transport walk and traditional bike. The travel cost for the mode car is calculated based on the fuel type, approximate weight of the owned car, and the distance of the stage. Pt-ticket prices are also based on distance and incorporate the pt-subscription of the individuals (e.g., GA Travelcard, Half fare Travelcard, zonal subscription.)

Slope[6] is only included for the alternatives walk and traditional bike as it is assumed not to influence the utility of the alternatives car and pt. The slope accounts for steep hills, which could be crucial in the mode choice from an exertion perspective. However, the total elevation gain could add a more nuanced understanding of the elevation increase, but it was not available for this thesis.

The slope parameter included in the DCM estimation is calculated as the average of the absolute slope up and absolute slope down of each stage s as follows, exemplary for slope bike:

[6]The slope routing procedure is best shown in an example: Imagine a simple route with different altitudes at different points of the route, i.e., 400m, 407m, 398m, 410m. Then, each difference between these altitudes is calculated, resulting in $400 - 407, 407 - 398, 398 - 410 = 7m, -9m, 12m$. Therefore, the elevation increased by 7 meters in the first segment, decreased by 9 meters in the second segment and increased by 12 meters in the last segment. For each stage segment, the slope is calculated by dividing it by the respective segment distance. These slopes are then averaged across a stage: The average slope up and the average slope down each consists of the upwards and downwards segments.

$$Slope_{\text{bike},s} = \frac{\left|\text{slope down}_{\text{bike},s}\right| + \left|\text{slope up}_{\text{bike},s}\right|}{2} \tag{14}$$

This definition accounts for the two-way problem of a stage. The assumption is that a stage in the dataset only reflects one way of a whole travel day. As we can assume that, in most cases, people are returning home at some time, we need to take that into account into their substituted mode choice. Taking the average of both slope parameters accounts for the fact that an individual also considers the way back into their mode choice. For example, riding an overall positive average slope down on one stage implies that the individual needs to ride it up again, which is assessed when choosing a mode of transport. As we only have data for one stage, taking the absolute average of the slopes up and down does account for that.

Situational Attributes

In addition to the stage-specific attributes, the weather variables *rain*, *heat*, and *cold* were included in the model. The weather variables were averaged over the whole day. We assume that the mode choice does not only depend on the specific stage but also a) on expectations of the day's weather if the stage was recorded in the morning, or b) on past events if the stage was recorded in the evening. For example, when asked about a stage done in the evening (e.g., on the way home from work), this particular mode choice depends on the choice situation in the morning.

The level of heat and cold for a particular stage s on day t is defined analogously to Hintermann et al. (2024, p. 15) as:

$$Heat_{st} = \max(t_{st}^{\max} - 25, 0) \tag{15}$$

$$Cold_{st} = \max(10 - t_{st}^{\min}, 0) \tag{16}$$

This specification incorporates a nonlinear effect of temperature on travel choices, whereas the variables $t_{st}^{\max}$ and $t_{st}^{\min}$ refer to the maximum and minimum temperature on that day, recorded in degrees Celsius at the weather station which was closest to the departure location for that stage s. Trip purposes were aggregated into three categories: *Commuting* (to work or school), *Leisure*, which includes getting to another destination, such as Restaurants, Friends, Sports, Events, etc., as well as transporting children, and *Other* where shopping, errands and coming home were included. *Coming home* is assigned to *Other* as we do not exactly know the main purpose of the stage. It could be coming home from work (commute) but also coming home from a leisure activity. Attributes concerning the time of day or weekday are not included in the estimation due to the parsimony of the model.

Person-specific attributes

The person-specific attributes also indicated as socio-demographic variables, give knowledge about diverging preferences for the substituted modes of transport across socio-demographic groups, as seen in the literature. In the model included are the following characteristics:

- Gender: Female, male (Reference)
- Age: Young (16-35 years), middle (36-59 years) (Reference), old (60+ years)
- Education: Tertiary education, i.e., higher vocational education and University, and non-Tertiary education (mandatory education, general education, vocational education and training/Apprenticeship) (Reference)
- Employment Status: Employed, all other (Reference) (students, retired, not employed, other, self-employed)
- Citizenship: Swiss (Reference), non-swiss
- Household Income: More than 10'000 CHF (Reference), Equal to or less than 10'000 CHF per month, No Answer
- Household size: Less than 4 individuals living in the same household (Reference), More than 4 individuals living in the same household
- Urbanity Level of Residence: City (Reference), non-city (Rural and Suburban)
- Language: German (Reference), French, English

The categorization of the variables is based, on the one hand, on the number of participants per category. As we see in Table A1 in the Appendix, some categories consist of only a few individuals. This fact makes it necessary to aggregate them, for example, in a binary categorization, such as in the cases of education, employment status, household size, and urbanity level. On the other hand, there are variables where an aggregation is not practical. The language variable was preferred to be included in three categories as these account for cultural and geographical differences. English only includes very few participants but makes the distinction between the German and French parts of Switzerland more precise. As the sample of the intermediate survey only includes 4567 choice observations of 1424 individuals, the efficiency of estimation had to be kept in mind when including socio-demographic variables. Therefore, the civil status (e.g., being married), the occupation, and the working percentage (part-time vs. full-time) were not included as the former is especially important for mode choice models investigating the effect of a specific household structure or across gender (see, e.g., Arman et al., 2018). As the employment status is included in the model, the occupation is excluded as it represents a refinement of the working status. The same accounts for the working percentage. Meanwhile, the possession of pt-passes was not directly included in the model as a person-specific attribute. As the calculation of the cost of pt is based on the possession of a pt-pass, these two attributes would be highly correlated and bias the estimation of the model. Furthermore, we do not know whether pt-passes were substituted by owning an e-bike as this mode is one of those that get highly substituted (see Figure 6). However, the mobility tools available to individuals were considered to determine each participant's individual choice sets.

Interactions of stage-specific attributes and person-specific attributes

Travel time is assumed to be perceived differently across the modes of transport. Due to this baseline assumption on which the MNL Base model builds, the socio-demographic variables are interacted with each travel time and mode. To account for preferences across socio-demographic groups for the travel time of each mode, these types of interactions were also included in the model. A different specification of travel time preferences across socio-demographic groups is included in the Appendix (see Table A2). The alternative specification follows the logic that an increase in travel time would be assumed to be equal across all modes and is then interacted with each mode. However, as the VTT is estimated differently across mode (see, e.g, Schmid et al., 2021), the specification of travel time per mode is chosen for the main model. The household income is assumed to only have an impact on the utility concerning an increase in cost, not on different modes. Therefore, this person-specific attribute is only interacted with travel cost.

DCM building process

Initially, the first model (MNL Base) will be estimated using only the attributes cost and time of each alternative and stage to control the behavior of these substantial variables. The attribute travel time is estimated for each mode, as it is assumed that not every additional second spent with one mode is perceived in the same manner. However, every additional spent Swiss franc is assumed to influence the utility in the same way across all modes. In the next model, situational variables are included, i.e., the slope affecting the mode choice of walk and bike and the trip purposes *Commute*, *Leisure*, and *Other* as well as weather variables. These are also interacted with the preference for each mode of transport.

The socio-demographic variables were included as a third model, building on the parsimonious situational model. It includes the characteristics as described above.

To account for heterogeneous preferences among the participants of the intermediate survey, a MMNL will be estimated as a fourth model, including only those situational variables that are expected to vary across the individuals. Therefore, only travel time, travel cost, and the slope are included, not the different purposes. I assume, therefore, that it is conceivable that individuals perceive varying utilities from an increase in travel time, travel cost, and the slope, which cannot be described by averaging preferences across socio-demographic groups. A MMNL will then be estimated as a fifth model, including the socio-demographic variables significant in the multinomial logit model (MNL Socio).

Applied Tools and Software

To process the data, provide descriptive statistics, and estimate the models, the open-source software R version 4.3.2 with the corresponding RStudio interface is used (R Core Team, 2023). For the estimation of the (mixed) multinomial logit models as well as for the prediction of the substituted mode of those e-bike stages, which were not part of the intermediate survey, the R package *apollo* was used (Hess and Palma, 2019).

4.3.2 Stage Substitution

As described in Chapter 1, only some stages are replaced with the e-bike by individuals owning an e-bike. Consequently, an approach is required to determine the impact of stage-specific attributes, socio-demographic variables, and the origin chosen modes of transport on the choice of taking the e-bike for a specific stage. This section defines the approach of the prediction of the substituted modes of transport for all comparable GPS-tracked e-bike stages in EBIS (see Figure 1, N = 19'390). Furthermore, it describes the prediction on the individuals in the MTMC 2021. The substitution rates across the intermediate survey of the EBIS can be seen in Subsection 5.1.

Prediction of Substituted Modes

As we only have knowledge of the origin mode of transport for 4567 selected e-bike stages queried in the intermediate survey, the sample used for the binary logistic regression to determine the probability to switch would be skewed away from the mode choice e-bike. Therefore, the substituted mode of transport is predicted for additionally 19'390 GPS tracks of EBIS detected as e-bike stages by using the estimation of the MMNL on the intermediate survey. The predictions are, therefore, based on the preferences across these individuals for the substituted mode of transport that they stated in the intermediate survey. The mode with the highest estimated probability is assumed to be the substituted mode. With this approach, the binary logistic regression model, which will be described below, is less prone to only a few e-bike stages and their respective underlying stage-specific attributes.

Binary logistic regression

By owning or regularly using an e-bike, the respective individuals are extending their choice set as an additional mode of transport is added to their available mobility tools. However, the answers to the retrospective survey only give us preferences for the substituted mode for specific stages done with the e-bike. This implies that we do not know the determinants of when an individual n, who has an e-bike, does actually make it its chosen mode of transport for a specific stage. For this, the determinants of choosing the e-bike for a specific stage by a specific individual using the GPS-tracked stages of EBIS are estimated. These stages are comparable to those included in the intermediate survey. Mainly, this means that only the mode groups walk, traditional bike, car, and pt and only stages which are at least 2 km and maximum 40 km long (indicated in the following as "comparable stages") is be included. The dependent variable is characterized by:

- Y = 1, if a stage is done with the e-bike and;
- Y = 0, if a stage is done with another mode.

The substituted mode of transport included in the binary logistic regression is defined as follows:

- E-bike stage and selected for the intermediate survey: Substituted mode as stated in the intermediate survey (N = 4567 choice observations)
- E-bike stage and not in intermediate survey: Predicted substituted mode based on the preferences estimated in the MMNL (N = 19390 GPS-tracks)
- Non e-bike stage: Chosen mode according to detection of the "Catch-my-Day" (N = 44231 comparable GPS-tracks)

The estimation to examine the probability of doing a stage with the e-bike is done as follows:

$$\log\left(\frac{P(\text{switched}=1)}{1-P(\text{switched}=1)}\right)=\beta_0+\sum_{js}\beta_j X_{js}+\sum_{ks}\beta_k X_k s+\sum_{n}\beta_n X_n \tag{17}$$

where:

- Switched = Dependent variable
- X_{js} = Binary indicator for each origin mode j taken for stage s
- X_{ks} = Stage-specific attributes k for stage s
- X_n = Person-specific attributes for decision-maker n assumed to be constant across all stages

4.3.3 CO_2 Emission Savings in Switzerland

The estimated parameters from the binary logistic regression can be used to predict the choice probabilities for each stage and each individual whether she is taking the e-bike or another mode. Consequently, these estimation results can be transferred to the representative Transport and Mobility Microcensus of Switzerland 2021 to induce the CO_2 emission savings in Switzerland. For the calculation of the average emission of each mode, the Mobitool[7] factors are used, which are based on a life-cycle assessment of the respective modes of transport in terms of emission (Sacchi and Bauer, 2023). The factors already account for the average emission per km across the fleet of Switzerland. The calculation of the emission savings for each stage was conducted as follows:

$$\begin{aligned}\text{Emission savings} = {} & E_j \times \text{length}_s \\ & -\left(P(\text{Switched}=1)\times E_{e-bike}\times \text{length}_s + P(\text{Switched}=0)\times E_j \times \text{length}_s\right)\end{aligned}$$

where E_j is the average emission per km of the chosen alternative j and $length_s$ the distance of the respective stage s in km.

[7] https://www.mobitool.ch/de/tools/mobitool-faktoren-v3-0-25.html

5 Analysis

In this Chapter, the collected data is examined, and underlying patterns are investigated in the descriptive statistics in the following Section 5.1. The substitution patterns of the e-biking stages of the EBIS intermediate survey are shown in Section 5.2. Then, the DCM results are presented in Section 5.3. The prediction of the substituted mode of transport across stages of the EBIS tracking data is shown as well as for the MTMC 2021 sample is presented in Section 5.4. The results from the binary logistic regression are shown in Section 5.5 including the calculation of the potential emission savings due to e-bikes in the transport sector of Switzerland.

5.1 Descriptive Statistics

Stage-specific attributes

The stage-specific attributes are important variables for estimating preferences for mode choice models (Ben-Akiva and Lerman, 1985). In Figure 4, we see a clear pattern in travel time for these short distances queried in the intermediate survey across the chosen and non-chosen modes: In general, car is the fastest mode of transport, while walking is clearly the slowest. This would imply that, in general, if a decision-maker n is choosing walk over car for a specific stage, the valuation of travel time is low. The traditional bike is, in general, slightly slower than the pt, but with an increasing difference in longer rides. In general, the travel time distribution is right-skewed due to the preliminary selection of stages, i.e., no farther than 40 km. As can be seen in Table 4, the actual duration of these stages was never exceeding 2 hours and 22 minutes, while the average duration was only 19 minutes. This suggests an average speed of the e-bike of 21.8 km/h across all stages in the intermediate survey.

TABLE 4: Summary statistics of tracked duration and length of intermediate survey stages

Attribute	Average	Median	Max
Duration (min)	18.58	14.82	141.13
Length (km)	6.76	4.75	40.54

Comparing the average travel time for each mode across the choices serves as a first indicator of preferences across the modes depending on travel time. Table 5 shows that the average time for the substituted mode bike is clearly lower when chosen compared to the travel time when not chosen, which can be explained by choosing a bike for shorter distances (see

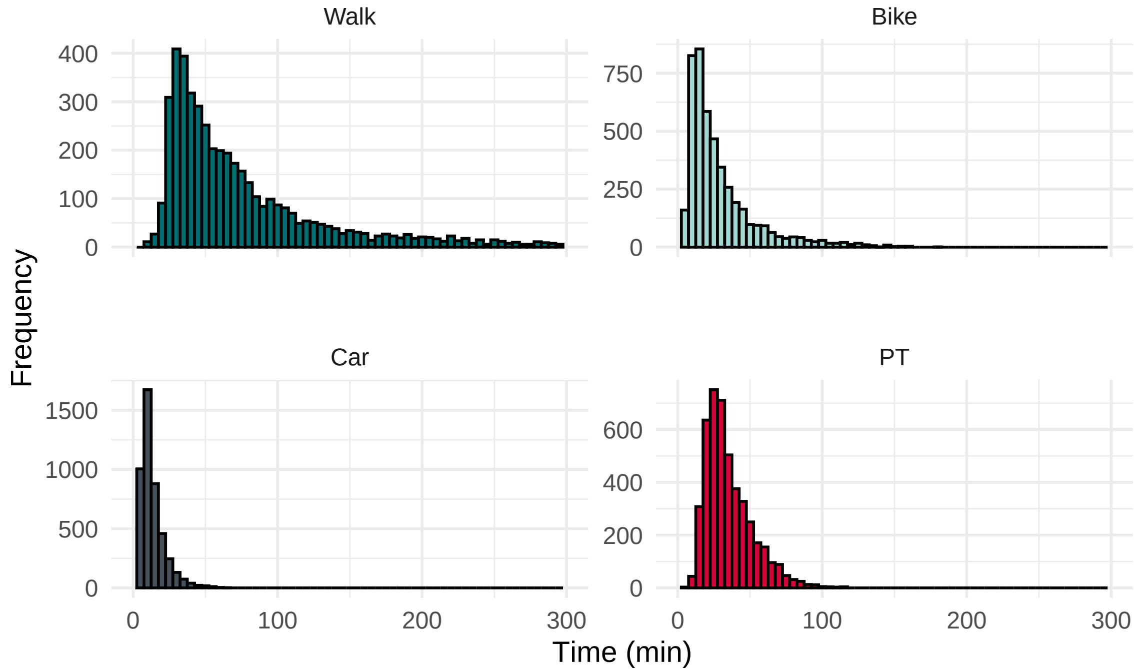


FIGURE 4: Travel time across modes in the intermediate survey including chosen and non-chosen alternatives

Appendix, Figure A2). This is also the case for walking. Still, the average time is naturally higher also in the chosen case as the travel time in walking is, on average, higher as has been seen in Figure 4. If we take a closer look at distance, people are still choosing to walk for shorter distances, as can be seen in the Appendix in Figure A2. Interestingly, there is almost no difference between the travel time for the chosen and non-chosen case for the substituted modes car and pt. However, the average time in the case of the car is clearly lower than for pt in both choice scenarios.

TABLE 5: Summary statistics: Travel time and travel cost by mode and choice

Mode	Choice	Mean Time	Mean Cost
Bike	Chosen	21.19	0.00
	Not Chosen	34.57	0.00
Car	Chosen	13.54	2.62
	Not Chosen	13.38	2.04
PT	Chosen	35.50	2.37
	Not Chosen	35.27	2.57
Walk	Chosen	35.24	0.00
	Not Chosen	80.19	0.00

The average cost has only a positive value for the modes of transport car and pt as it is assumed that an additional kilometer done with the traditional bike or by walking does not pose an additional cost to the decision-maker. The average cost of the substituted mode car is, on average, higher than when not chosen. This finding could point in the direction that, as the cost calculation is based on the travel distance, people could substitute cars rather on longer distances which can be traveled in an efficient time amount, i.e., higher speed, as the average travel time is in the chosen and non-chosen case comparably low. On the other hand, the substituted mode pt shows a slightly higher average cost for the non-chosen cases while it has a lower average time than in the chosen case. This could indicate that there could be different cost-sensitivities across these two modes. This could be explained through the fact, that for an additional travel time being correlated with travel distance, the increase in the price is noticed more directly in the case of choosing the pt. An exception is the case of pt-passes with full flat rate (GA Travelcard and zonal subscription as stages are below 40 km).

The relationship between the travel time and the travel cost across the substituted modes car and pt can better be seen in Figure 5. The tariff for pt is calculated based on the travel distance and the pt-subscription of the individuals, which is also visibly in Figure 5: The travel cost equal to zero indicates the travel cost for individuals which do own a GA Travelcard and therefore are not paying an additional cost for each stage. The Half Fare Travelcard does halve the cost for each pt stage for individuals having such a subscription. This can also be seen in the travel costs across travel time which is more dense across a travel cost of around 2.50 CHF. Summarized are the travel costs for pt rather high in the case of short stages, but an increase in travel time does not increase the cost as much as in the case of an additional travel time in a car.

The slope is included for the modes of transport, walking, and traditional bike, as it acts as a proxy for exertion. Therefore, I assume that it does not influence the utility of the alternatives car and pt. In Table 6, the summary statistics of the generated slope values for the intermediate survey stages is displayed, and the calculated absolute average slope as described in Chapter 4.3. We can see that for the mode of transport bike, when chosen, the mean and median for each slope type in percentages are closer to zero, indicating less steep climbs up and down during the stage. This does not apply to the slope types in mode walk, where the slope does not play a considerable role in substituted mode choice. While the slope percentages in walk are, in general, prone to steep climbs and descends, the mean and median for chosen modes are still further away from zero.

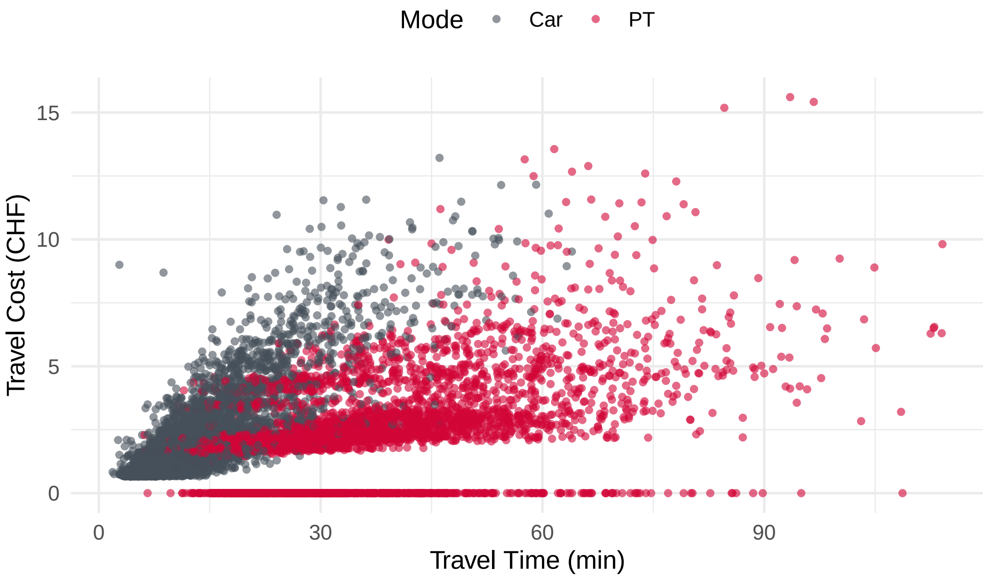


FIGURE 5: Relationship between travel time and travel cost of modes car and pt

TABLE 6: Summary statistics: Slope bike and walk by mode

Mode	Slope Type	Choice	Mean	Median	Min.	Max.
Bike	Absolute Average Slope	Chosen	4.20	4.10	0.65	13.30
		Not Chosen	4.43	4.29	0.66	13.42
	Slope Down	Chosen	-4.15	-3.98	-18.65	-0.25
		Not Chosen	-4.37	-4.17	-18.00	-0.13
	Slope Up	Chosen	4.26	4.06	0.22	20.66
		Not Chosen	4.49	4.26	0.13	20.02
Walk	Absolute Average Slope	Chosen	5.92	5.73	1.91	11.40
		Not Chosen	5.41	5.15	0.95	24.08
	Slope Down	Chosen	-5.90	-5.31	-16.79	-0.75
		Not Chosen	-5.43	-5.02	-40.69	-0.21
	Slope Up	Chosen	5.93	6.00	1.85	13.28
		Not Chosen	5.39	5.08	0.16	39.89

Trip purposes

Trip purposes can be important factors influencing the substitution of different modes through the e-bike (see, e.g., Wolf and Seebauer, 2014; MacArthur et al., 2018; Lee et al., 2015; Dill and Rose, 2012). The assignment of the purposes of the stages was not straightforward, as seen in Section 3, which was why this was aggregated into broader groups to be able to identify these influences. First of all, *Commute* was the purpose which was mostly stated for the selected stages shown to the participants in the intermediate survey with 39% of the stages. This is not surprising as the stages were also selected based on repetition, i.e., the number of similar occurrences, in the GPS-tracking data of each participant. This trip purpose was followed by the category *Other* with a share of nearly 33%, which is a collection of purposes. The trip purpose *Other* includes running errands, doing shopping, and transporting children but also *Coming home*, which cannot be assigned properly to either commuting or leisure purposes. Leisure is then the least selected purpose with approximately 28%. Note that these percentages cannot be used to provide evidence for when the e-bike is used most often, as these specific stages were particularly selected as described in Section 3.2.

Table 7: Summary statistics: Percentage of stage purposes

Stage Purpose	Percentage
Commute	39.26
Other (including errands, transporting children and coming home)	32.82
Leisure (including visiting family/friends, getting to another destination)	27.92

Socio-demographic characteristics and mode choice

Before estimating choice probabilities and preferences in the DCM for the substituted mode, exploring the choices across the socio-demographic characteristics of the individuals in the EBIS sample can provide preliminary insights and inform about potential significant person-specific attributes. Table 8 shows the chosen substituted mode across the socio-demographic variables used in the estimation of the preferences across modes.

Walking is generally the least substituted mode by frequency, with percentages of around 1%. However, speaking French shows a slightly higher percentage of people choosing walking, while individuals living in rural or suburban areas choose it the least as their substituted mode. The latter is not surprising as this might be correlated with a longer travel time, as the travel distance from destination A to destination B could be longer generally for individuals not living in a city. Bike as the substituted mode was mainly preferred by participants over 60 years and the least by individuals speaking French. Also, individuals living outside of the urban area substituted the bike less than people living in a city. This low choice frequency of the latter transfers to a high percentage of the substituted mode car. In contrast, individuals living in the city substituted fewer car stages with e-bikes. Also, individuals aged 16 to 29 are only substituting 22.7% of stages with e-bikes instead of the car. These choices could also be due to the availability of cars to the individuals, as approximately one-third has no or only sometimes access to a car in the EBIS groups A and B as previously seen in Table 3. However, the potential for car substitution also seems higher in the group of non-employed

TABLE 8: Socio-demographic variables and percentage of chosen substituted mode

		Chosen Substituted Mode			
Variable	Category	Walk (%)	Bike (%)	Car (%)	PT (%)
Age	Young (16-29 years)	1.01	35.21	22.74	41.05
	Middle (30-59 years)	1.16	36.18	27.83	34.84
	Old (60+ years)	1.02	46.44	31.17	21.37
Citizenship	Swiss	0.95	38.82	27.38	32.84
	Non-Swiss	1.81	33.75	29.80	34.65
Education	Mandatory/Secondary	1.16	40.59	30.54	27.71
	Tertiary	1.11	37.27	27.30	34.32
Employment Status	Employed	1.07	36.79	26.50	35.64
	Not Employed*	1.32	42.04	33.26	23.38
Gender	Female	1.46	36.25	25.07	37.22
	Male	0.89	38.86	29.83	30.42
HH Income	$>$ 10'000 CHF	1.07	35.63	29.25	34.05
	$\leq$ 10'000 CHF	1.16	42.57	25.92	30.34
	No Answer	1.20	36.14	27.71	34.94
HH Size	1, 2 or 3 People	1.35	38.85	27.70	32.11
	$\geq$ 4 People	0.77	36.31	28.09	34.83
Language	German	0.87	40.73	27.60	30.81
	English	1.52	27.92	23.86	46.70
	French	3.12	16.93	31.85	48.11
Trip Purpose	Commute	0.78	33.24	25.77	40.21
	Leisure	1.33	43.22	28.00	27.45
	Other	1.33	38.76	30.22	29.69
Urbanity Level	City	1.26	41.75	23.33	33.65
	Non-City: Rural/Suburban	0.64	25.30	42.32	31.74

* including Retired, Unemployed, Student, and Other.

individuals. Young individuals seem to have a clear preference for substituting pt stages with the e-bike compared to other age categories and within its age category. This pt-preference is also valid for individuals speaking French and English. Interestingly, this pt-preference is not seen in non-swiss citizens, although they might correlate with speaking English. Concerning trip purposes, pt is mainly substituted by the e-bike for commuting trips and the bike for leisure trips.

5.2 Substitution Patterns EBIS

The computation of the substitution pattern of the e-bike for the EBIS sample is straightforward, as the mode that would have been chosen before having an e-bike is known for selected, representative stages. Therefore, the mode chosen is always the e-bike, and the substituted mode is known directly through the survey. The substitution rates were therefore calculated as follows:

$$\text{subrate}_{\text{stage-level}}(\text{mode}_{\text{e-bike}}, \text{mode}_{\text{substituted}}) = \frac{\sum \text{stages}(\text{mode}_{\text{substituted}})}{\sum \text{stages}(\text{mode}_{\text{e-bike}})} \tag{18}$$

$$\text{subrate}_{\text{km-level}}(\text{mode}_{\text{e-bike}}, \text{mode}_{\text{substituted}}) = \frac{\sum \text{distance}(\text{mode}_{\text{substituted}})}{\sum \text{distance}(\text{mode}_{\text{e-bike}})} \tag{19}$$

For the stage-level, the number of stages with the substituted mode, e.g., the number of stages with the traditional bike, is divided by the total number of stages conducted with the e-bike. For the km-level, the total distance with a specific substituted mode is divided by the total distance with the e-bike. This approach follows Reck et al. (2022).

The resulting substitution rates on the stage-level and the km-level for the respective stages of the intermediate survey are shown in Figure 6. The results show that on a stage-level substitution rate, the traditional bike was replaced the most by the e-bike, followed by pt, car, and walk. Walking is also on the km-level the least substituted mode by the e-bike with a percentage of only 0.5%. The percentages of substituted km are similar to the substituted stages with substitution rates of 29.0% for the traditional bike, 35.9% for the car, and 41.4% for the pt. The difference in shares between the substituted km and the substituted stages is particularly interesting. The dispersion of these two measures in the traditional bike is an indication that the traditional bike was mainly displaced for a lot of shorter stages while the pt was displaced for fewer but longer stages in the selected data. The latter also applies to car substitution. This relationship between the substituted km and the substituted mode can further be seen in Figure 7. With an increase in distance, fewer walking stages are substituted while more stages by pt and cars got substituted by e-bikes.

Environmental implications

To measure the impact of the e-bike on the CO_2 emissions of the transport sector in Switzerland, not only the substituted mode and the substituted km are important, but also the difference in emissions between the e-bike and the substituted mode is crucial.

The substitution pattern found from the intermediate survey, as seen in Figure 6, is combined

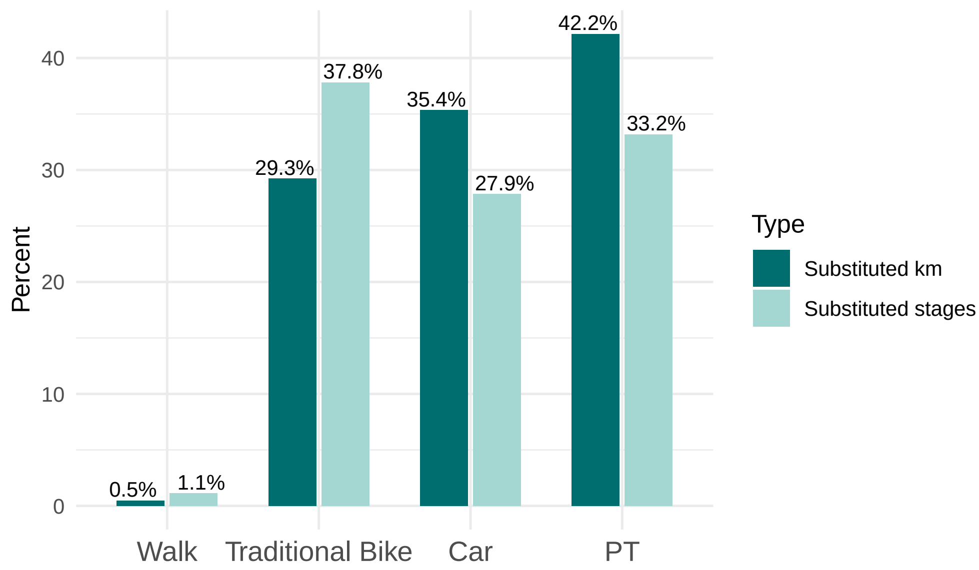


FIGURE 6: Substitution rates of e-bike stages per mode

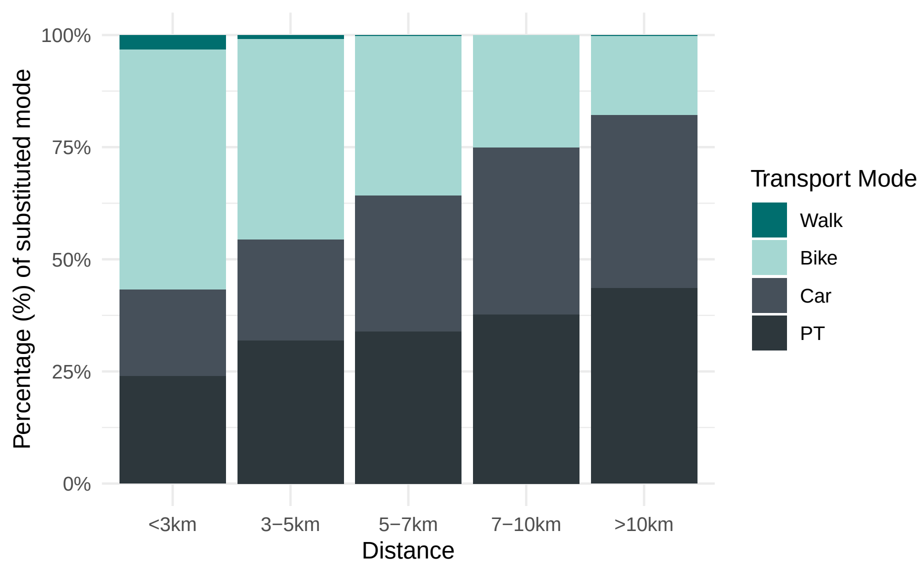


FIGURE 7: Substitution rates for e-bikes by distance brackets

with the life-cycle assessment of modes of transport in Switzerland of the MobitoolV3.0 (Sacchi and Bauer, 2023). This data on CO_2 emissions of each mode is added to the findings on substitution patterns for e-bikes to calculate the "net emissions" following Reck et al. (2022):

$$\text{Net Emissions}_{\text{e-bike}} = \text{Gross Emissions}_{\text{e-bike}} - \sum \text{Gross Emissions}_{\text{replaced mode}} \tag{20}$$

This approach incorporates the shares and, therefore, the substitution pattern calculated in

Figure 6, by which an e-bike substituted other modes over all e-bike stages of the intermediate survey. The results are shown in Table 9 for substituted km. Note that this analysis is done only for the substituted km as the analysis would be biased towards shorter stages when calculating it for the stage level, as can be seen in Figure 7.

The results show that there are negative net emissions due to the substitution of the specific stages by the e-bike, i.e., they replace, on average, more emission-intensive modes of transport such as car and pt on the km-level. On average, a CO_2 emission saving per km of 63g is achieved when taking the e-bike. This indicates that for the examined stages, an overall emission reduction could be achieved through the e-bike. Note, that stages induced by the e-bike are not included, which biases the shown results towards a higher emission saving. The stages induced by the e-bike, i.e., would not have been done before owning an e-bike, are actually adding to CO_2 emissions. In the EBIS sample these induced stages account for approximately 7% of all e-bike stages.

TABLE 9: Emission reduction per km due to e-bikes

Substituted mode	Gross emissions [g CO_2/pkm]	Substitution rate e-bike
Walk	0*	0.5%
PT (avg.)	77.2*	41.4%
Car (avg.)	138.9*	29.0%
Bike	5.6*	35.9%
Emissions of substituted modes	74.3	
Emissions of e-bike	11.3 *	
Net emissions [g CO_2/pkm]	**-63**	

* Life-Cycle Assessment Emission Calculation of Mobitool (Sacchi and Bauer, 2023).

5.3 Results of Estimation of Discrete Choice Models

The estimation results of the parsimonious models of the MNL and MMNL specifications are shown in Table 10. These parsimonious models are built upon the significant coefficients of the full models, i.e., those models that include all variables as described in Section 4.3. The initial models are in the Appendix in Table A2. The reference of substituted mode of transport is the traditional bike.

The estimated values for the ASC in the MNL Base Model show that there is an overall preference in the choice for the bike as the substituted mode compared to all other included modes of transport *walk*, *car*, and *pt*, ceteris paribus. The differences in the magnitudes of the negative values also show that there is, after the preference for the reference mode bike, a higher preference for pt followed by car than for walking as the substituted mode. This ranking aligns with the findings in Figure 6. There is a significant negative preference for increasing travel costs across all modes. The coefficient of -0.186 indicates that for each unit increase in cost, the utility of the individuals in the sample decreases on average by 0.181 units. An increase in travel time for each mode also decreases the individuals' utilities, but not for all modes in the same manner. An increase in the travel time by pt is not perceived as negative as an increase in travel time by car, bike, or walking.

Including the situational variables precipitation, hot and cold days (only included in the full situation model as seen in the Appendix in Table A2, the slope for bike and walk, and the trip purpose interacted with each mode of transport amplifies some results of the MNL Base model. First, the magnitude of the preference for the bike as a substituted mode increases, which can be seen in the coefficients for ASC of all the alternative modes walk, car, and pt. This can be explained as ASC is measuring, in this case, the market share of preference not captured by the explanatory variables. Therefore, the weather variables could still capture some negative utility from the bike compared to the MNL Base. In general, the relative differences in the preferences of the individuals across the modes are increasing compared to the MNL Base model but maintaining the same order of preference: bike > pt > car > walk. The weather variables precipitation, cold, and hot days are insignificant, as seen in the Appendix in Table A2 in model MNL A2. Consequently, they are not included in the parsimonious MNL situational model in Table 10. However, the slope for mode walk was included in the parsimonious situational model, even though it is not significant as we see a distinction to the slope of the bike, which seems to have a negative impact on the utility of the choice bike on the 1% significance level. The purpose is also an important determinant in mode choice regarding pt. There is a dislike of taking pt for the purposes *Leisure* or *Other* compared to *Commute*. While there was a slightly higher utility stemming from cars for leisure and other purposes compared to commuting in the full model seen in the Appendix in Table A2, it is not significant in the parsimonious model in Table 10. That change can occur due to interactions with other variables no longer part of the parsimonious model.

Adding household income and socio-demographic variables to the model provides a more nuanced picture of the substituted mode of the e-bike. Note that the included number of observations is slightly lower due to individuals for whom no information on the employment status was available. The included stage-specific attributes, travel cost, travel time, and the slope bike are still significantly decreasing the utility of transport mode choice. However, the coefficient for the bikes is not as high as in the MNL situational model and is only significant on the 10%-level. There is a preference for purposes "Other" done by car compared to the purpose "Commute".

Younger people are less likely to have substituted the bike than middle-aged people, and the same applies to individuals who are 60 years old or older and for the substitution of pt. Employed people had a significant preference for pt as the substituted mode compared to people not being employed (for example, students or retired) in the full model MNL A2 as seen in the Appendix in Table A2, but not when estimating the parsimonious model. Speaking French is highly important in determining the preferences for the substituted mode. These individuals prefer walking, car, and pt more than German-speaking people over the bike as a substitute mode. English-speaking individuals do not have significantly different preferences in terms of mode choice than German-speaking e-bike users. However, not being a Swiss citizen influences the choice of the substituted mode. There is a higher preference for walking, car, and pt.

An increase in walking travel time is perceived more negatively by younger people than middle-aged and older people. The same applies to the travel time with a traditional bike. Female individuals also dislike an increase in travel time more in each mode compared to male individuals. Also, if the household size is greater than 4, an increase in travel time influences the utility negatively compared to smaller household sizes for the modes bike, car, and pt. Lower-income households only seem to experience more dislike of an increase in travel cost than higher-income households in the full model, but not in the parsimonious model. Individuals who did not report their household income do not show any difference in utility than higher-income households concerning travel costs.

Table 10: Estimation results of Multinomial logit model (MNL) and Mixed multinomial logit model (MMNL)

Variable	**MNL Base**	**MNL Situational**	**MNL Socio**	**MMNL Base**	**MMNL Socio**
ASC Walk	-1.971***	-2.527***	-2.459***	-3.544**	-3.935***
	(0.656)	(0.721)	(0.663)	(1.528)	(1.215)
ASC Car	-0.808***	-1.358***	-1.45***	-2.446***	-3.264***
	(0.125)	(0.224)	(0.26)	(0.589)	(0.715)
ASC PT	-0.343**	-0.544**	-0.905***	-2.353***	-2.966***
	(0.136)	(0.215)	(0.298)	(0.566)	(0.732)
Travel Cost	-0.186***	-0.187***	-0.181***	-1.567***	-1.808***
	(0.038)	(0.037)	(0.043)	(0.456)	(0.567)
TT Walk	-0.072***	-0.071***	-0.077***	-0.182***	-0.196***
	(0.018)	(0.018)	(0.02)	(0.04)	(0.04)

Table 10: Continued

Variable	**MNL Base**	**MNL Situational**	**MNL Socio**	**MMNL Base**	**MMNL Socio**
TT Bike	-0.073*** (0.005)	-0.072*** (0.005)	-0.056*** (0.009)	-0.206*** (0.02)	-0.159*** (0.021)
TT Car	-0.088*** (0.009)	-0.081*** (0.009)	-0.056*** (0.014)	-0.289*** (0.03)	-0.227*** (0.038)
TT PT	-0.05*** (0.004)	-0.05*** (0.004)	-0.04*** (0.006)	-0.102*** (0.011)	-0.093*** (0.014)
Slope Walk		0.022 (0.052)		-0.026 (0.114)	
Slope Bike		-0.09*** (0.031)	-0.058* (0.033)	-0.129 (0.103)	-0.06 (0.108)
Purpose Leisure Car		0.14 (0.138)	0.187 (0.145)		0.54 (0.355)
Purpose Other Car		0.193 (0.134)	0.262* (0.141)		0.598* (0.328)
Purpose Leisure PT		-0.291** (0.125)	-0.227* (0.131)		-0.28 (0.316)
Purpose Other PT		-0.261** (0.12)	-0.163 (0.125)		-0.406 (0.269)
Age $\leq$ 35 Car			-1.059*** (0.354)		-1.165 (0.984)
Age $\geq$ 60 Car			-0.568* (0.299)		-1.117 (0.859)
Age $\leq$ 35 PT			-0.75** (0.349)		-0.947 (0.934)
Age $\geq$ 60 PT			-0.668** (0.315)		-1.231 (0.915)
Employed PT			0.301 (0.192)		0.853* (0.439)
French Walk			2.502*** (0.595)		5.676*** (1.436)
English Walk			1.102 (0.863)		3.099* (1.67)
French Car			1.93*** (0.408)		5.008*** (1.242)

Table 10: Continued

Variable	**MNL Base**	**MNL Situational**	**MNL Socio**	**MMNL Base**	**MMNL Socio**
English Car			0.603		1.984
			(0.547)		(1.411)
French PT			2.123***		5.449***
			(0.4)		(1.231)
English PT			0.998**		2.69**
			(0.469)		(1.299)
Non-city Walk			1.15**		2.95*
			(0.512)		(1.512)
Non-City Car			1.214***		3.189***
			(0.251)		(0.913)
Non-City PT			0.821***		2.106**
			(0.258)		(0.911)
TT Age $\leq$ 35 Walk			-0.031**		-0.04
			(0.016)		(0.031)
TT Age $\geq$ 60 Walk			-0.012		-0.012
			(0.015)		(0.037)
TT Age $\leq$ 35 Bike			-0.039***		-0.041
			(0.011)		(0.034)
TT Age $\geq$ 60 Bike			-0.019*		-0.006
			(0.01)		(0.034)
TT Female Bike			-0.029***		-0.079***
			(0.009)		(0.022)
TT Female Car			-0.046***		-0.128***
			(0.018)		(0.041)
TT Female PT			-0.011*		-0.023*
			(0.006)		(0.013)
TT HH Size $\geq$ 4 Bike			-0.016**		-0.022
			(0.008)		(0.022)
TT HH Size $\geq$ 4 Car			-0.028*		-0.035
			(0.016)		(0.044)
TT HH Size $\geq$ 4 PT			-0.012*		-0.012
			(0.006)		(0.014)
TT French Bike			0.033**		0.027
			(0.013)		(0.044)

TABLE 10: Continued

Variable	MNL Base	MNL Situational	MNL Socio	MMNL Base	MMNL Socio
TT English Bike			0.007		-0.032
			(0.015)		(0.03)
TT Non-City Bike			0.015**		-0.001
			(0.007)		(0.026)
Cost HH Income $\leq$ 10'000 CHF			-0.079		-0.247*
			(0.049)		(0.132)
Cost HH Income No Answer			0.135		0.293
			(0.137)		(0.619)
σ_{Walk}				1.694*	1.844**
				(0.996)	(0.761)
σ_{Bike}				4.587***	4.209***
				(0.566)	(0.47)
σ_{Car}				1.985***	1.373***
				(0.528)	(0.366)
σ_{PT}				0.41	0.675
				(0.954)	(0.427)
σ_{Cost}				0.59***	0.807***
				(0.162)	(0.137)
$\sigma_{TT-Walk}$				0.048***	0.045***
				(0.01)	(0.014)
$\sigma_{TT-Bike}$				0.132***	0.118***
				(0.016)	(0.016)
σ_{TT-Car}				0.277***	0.29***
				(0.028)	(0.036)
σ_{TT-PT}				0.029	0.034***
				(0.018)	(0.01)
$\sigma_{SlopeWalk}$				0.112	
				(0.115)	
$\sigma_{SlopeBike}$				0.815***	0.721***
				(0.2)	(0.086)
Estimated Parameters	8	14	42	21	52
Number of Observations	4567	4567	4437	4567	4437
LL(0)	-5472.692	-5472.692	-5317.765	-5472.692	-5317.765
LL(final)	-4031.546	-4010.359	-3749.234	-2848.529	-2695.334

Table 10: Continued

Variable	**MNL Base**	**MNL Situational**	**MNL Socio**	**MMNL Base**	**MMNL Socio**
McFadden R^2	0.2633	0.2672	0.2950	0.4795	0.4931
Adjusted R^2	0.2619	0.2646	0.2871	0.4757	0.4834
AIC	8079.093	8048.718	7582.467	5739.058	5494.667
BIC	8130.505	8138.691	7851.172	5874.017	5827.350
Number of Draws				1000	1000

Note: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Reference Mode: Traditional Bike

Compared to the MNL, the MMNL Base model shows more accentuated effects of the preferences concerning substituted mode across individuals in the sample, but, in general, the effects point in the same direction. Note that the coefficient for the travel cost is now calculated as $-e^{\mu_{cost}+\sigma_{cost}}$, ensuring that the utility is always affected negatively, as it was already estimated in the MNL models. When converting the coefficient for travel cost by $-\log(-\beta_{cost})$, then we get a slightly higher value (-0.225) than for the MNL coefficients for travel cost, indicating that there are some more cost-sensitive individuals as we now account for preference heterogeneity. This finding is emphasized by the significant σ_{cost}, which shows significant differences across individual preference distributions concerning increased travel costs. Travel time for each mode also shows a higher magnitude than in the MNL. Preferences across individuals seem largely different for travel time, except for the travel time in pt. When accounting for these heterogeneous preferences, the slope of the bike, which was on average perceived significantly negative in the MNL framework, is not significant in the MMNL.

The MMNL Socio model, which includes the socio-demographic variables that were significant in the MNL full model, as shown in Table A2 in the Appendix, are now often less significant, but without changing the direction of the effect on utility. As already discussed, the slope for the substituted mode bike does not significantly affect the utility of a substituted mode. However, these preferences are subject to significant variation as seen in $\sigma_{SlopeBike}$. For the trip purposes, only when the stage serves the category *other* do individuals assign a higher utility to the substituted mode car compared to the trip purpose *commute*. Young and old people do not experience a significantly different utility for the different modes of transport. However, being employed increases the utility for pt (including *Retired*, *Student*, and *Other*) compared to not being employed. The effects of different languages compared to German persist in the MMNL and even clearly increase the choice probability of these individuals across the substituted modes of transport walk, car, and pt. Also, individuals living in a rural or suburban area are more likely to choose walking, car, and pt than individuals living in an urban area.

The coefficients for the utility of travel time across socio-demographic groups only show an effect for being female, indicating a strong disutility for increasing travel time in the substituted modes bike, car, and pt compared to male individuals. Furthermore, people living in households with an income of less than or equal to 10'000 CHF per month do experience a greater loss of utility from an increase in travel cost compared to people living in households with an income of more than 10'000 CHF per month.

The advantage of the MMNL is that we can incorporate heterogeneous preferences across individuals in the sample for alternatives and diverging sensitivities for stage attributes (Train, 2003), such as e.g., travel time, travel cost, and the slope. This variation in the estimated β across individuals is shown as $\sigma_{parameter}$ in Table 10 and graphically in the Appendix. Except for the ASC of pt, the variations of preferences and sensitivities across individuals are significant, indicating that the MMNL specification is more extensive in explaining preferences across individuals. Still, there are different magnitudes in these varying sensitivities. The preferences for the substituted mode σ_{Bike}, as seen in the Appendix in Figure A3, has the highest variation across the individuals, followed by the substituted modes of transport walk and car. Also, an increase in travel cost is perceived significantly different across the individuals in the sample. As seen in Figure A4a in the Appendix, few individuals are highly sensitive to increased travel costs. It is reasonable to suggest that the perception of travel time may differ for different alternatives. However, there are still varying sensitivity levels to travel time increases within each mode as indicated by $\sigma_{TT-Mode}$. The highest variation in sensitivity is identifiable from an increase in travel time by car, followed by bike, walk, and pt. This ranking is in line with the order of magnitudes of estimated experienced negative utility in travel time across the respective modes of transport.

The log-likelihood *LL(final)* is decreasing with the increase of variables in the MNL, which is also the case for the AIC. This decrease suggests that the additional stage-specific attributes and person-specific attributes explain the data better, compared to the MNL Base model. However, the BIC, which penalizes additional attributes in large datasets more than in small datasets, only decreases for the parsimonious MNL Socio specification. Therefore, including trip purposes and the stages' slope does not make the model fit the data better than only including travel time and travel cost.

The MMNL specifications have much lower *LL(final)* compared to the MNL-specifications.

Also, the AIC and BIC values indicate a better balance of fit and complexity. Each model has an excellent fit with McFadden R^2 between 0.26 and 0.49. Based on these measures, the MMNL Socio suggests the best fit and is therefore considered for the following post-estimation process. Also, the prediction of the substitution patterns is based on this specification.

Post-estimation process

The performance of the prediction of the MMNL is shown in Table 11. The difference between the market share of the chosen substituted mode and the market share of the prediction of that mode is for all alternatives not significantly different from zero at all typical significance levels. Note that this does not indicate that the chosen mode is the predicted mode with the highest probability, which would not follow the logic of a probabilistic choice, as discussed by Train (2003, p. 73). The average probability of the chosen mode is 47.7% (see Appendix, Table A4).

Table 11: Prediction goodness-of-fit of e-bike stages in the intermediate survey

Statistics	Walk	Bike	Car	PT
Times chosen (data)	49.00	1670.00	1251.00	1467.00
Times chosen (prediction)	54.51	1675.48	1227.82	1479.19
Diff (prediction-data)	5.51	5.48	-23.18	12.19
t-ratio	0.77	0.19	-0.85	0.43
p-val	0.44	0.85	0.39	0.67
Share in % (data)	1.10	37.64	28.19	33.06
Share in % (prediction)	1.23	37.76	27.67	33.34

The own- and cross-elasticities of a 1% increase in the attributes travel cost and travel time are displayed in Table 12. All the signs are according to expectation: An increase in an attribute of a specific mode decreases the choice probability of taking it and increases the choice probabilities of all other modes. The highest absolute values for elasticity are discovered for travel time. An increase in the travel cost of the car does primarily benefit the modes walk and pt. Increasing the travel cost of pt decreases its choice probability by 0.13% while walking is the alternative mode most sensitive to this increase. An increase of 1% in travel time walk does highly decrease its choice probability (-3.25%), while the increase in the choice probability of all other modes is relatively low. This is different for travel time bike. An increase by 1% does primarily benefit the choice probability of pt (+0.35%), followed by walk (+0.27%) and bike (+0.24%). On the other hand, we observe that an increase in the travel time of the car least benefits bike choices. The increase in pt travel time decreases its choice probability by 0.65%. This has the greatest effect on walking, followed by cars and bikes.

Table 12: Own- and cross-elasticities in % corresponding to a 1% increase in the attributes

Attribute	Walk	Bike	Car	PT
Travel Cost Car	0.06	0.03	-0.12	0.06
Travel Cost PT	0.22	0.04	0.09	-0.13
Travel Time Walk	-3.25	0.02	0.03	0.07
Travel Time Bike	0.27	-0.49	0.24	0.35
Travel Time Car	0.34	0.07	-0.51	0.33
Travel Time PT	1.03	0.22	0.44	-0.65

The Value of travel time (VTT) for each mode is displayed in Table 13. The highest VTT is indicated by the substituted mode car by 7.52 CHF/hour, while the lowest is estimated for pt with 3.09 CHF/hour. The latter is also only half as high as the estimated VTT of walk (6.51 CHF/hour). VTT of the bike is in the middle of the travel time valuation of the car and the pt. Therefore, individuals in the EBIS sample are willing to pay between 3.09 CHF and 7.52 CHF to save one hour of travel time.

Table 13: Value of travel time (VTT) for each mode per hour

Mode	Value (CHF/h)
VTT Car	7.521***
	(2.507)
VTT PT	3.088***
	(1.116)
VTT Walk	6.512**
	(2.572)
VTT Bike	5.263***
	(1.933)

5.4 Prediction of Substituted Modes of EBIS stages

In Table 14 the prediction of the substituted modes of comparable e-bike stages of the EBIS participants is displayed. The average probability of the chosen mode is 39.0%, as can be seen in the Appendix in Table A5. When comparing the predicted and actual market shares of the intermediate EBIS sample in Table 11 with the predicted market shares of the e-bike stages not surveyed in the intermediate questionnaire in Table 14, we see that the predicted market shares for car and walk are higher in the latter than in the former. On the other hand, the predicted market shares of bike and pt are lower for these e-bike stages.

Table 14: Prediction of comparable e-bike stages

Statistics	Walk	Bike	Car	PT
Times chosen (prediction)	373.90	5907.81	5915.98	6229.31
Share in %	2.03	32.06	32.10	33.81

5.5 Stage-level Substitution and Emission Savings

The results from the coefficients, i.e., the log-odds, and the respective odds ratios that are influencing whether an individual travels a specific stage with the e-bike, are shown in Table 15. The odds ratios are calculated as $e^{Estimate}$. An odds ratio above 1 indicates that doing a stage with the e-bike is more likely to happen when a binary variable = 1 or an increase in a continuous variable occurs compared to the reference category. If the odds ratio is below 1 then doing a stage with the e-bike is less likely with an increase in this variable compared to the reference category.

Taking the car as an alternative mode does decrease the probability of doing a stage by e-bike by 54.7% compared to the traditional bike. The same effect strongly applies to the alternative pt. Walk as an alternative mode does not significantly influence the probability of doing a stage with the e-bike compared to a bike. An increase in travel distance by one kilometer does decrease the probability of doing a stage with the e-bike by 14.3%. A travel time increase of one minute does have a small positive effect on the choice probability to do a stage with the e-bike, assuming no interaction effect. However, interacting each mode with the respective travel time gives a deeper understanding of this effect: The log-odds for choosing an e-bike increase by 0.017 (odds ratio: 1.017, or increase of probability of 1.7%) for each minute increase in travel time when using a car instead of a bike. The effect is even stronger for pt with an increase of the log-odds of 0.165 (odds ratio: 1.179). However, an increase in travel time corresponds to a lower probability of choosing the e-bike when the alternative mode was walking compared to biking (log-odds: -0.116, odds ratio: 0.890). Being female does increase the odds of taking an e-bike for a specific stage. The same is true for individuals who are at least 60 years old. However, being young decreases the probability of

Table 15: Estimation results of binary logistic regression on comparable stages

	Dependent variable:			
	Switched to e-bike = 1			
	Estimate	Odds Ratio	Robust s.e.	p-value
Constant	0.288***	1.334	0.058	0.000
Car	-0.792***	0.453	0.044	0.000
PT	-3.574***	0.028	0.079	0.000
Walk	0.114	1.121	0.303	0.707
Travel Distance (km)	-0.155***	0.857	0.004	0.000
Travel Time (min)	0.042***	1.043	0.002	0.000
Travel Cost (CHF)	0.021	1.021	0.013	0.110
Gender: Female	0.049**	1.050	0.021	0.017
Young	-0.225***	0.798	0.033	0.000
Old	0.171***	1.187	0.036	0.000
Urbanity Level: City	-0.349***	0.706	0.024	0.000
Precipitation Coverage (% of 24h)	-0.001*	0.999	0.000	0.081
Temperature: Heat	0.002	1.002	0.049	0.973
Temperature: Cold	-0.004	0.996	0.003	0.104
Non-Swiss	-0.105***	0.900	0.029	0.000
HH Income $\leq$ 10'000CHF	-0.017	0.983	0.021	0.417
HH Income NA	0.073	1.076	0.064	0.251
HH Size $\geq$ 4	0.027	1.028	0.022	0.208
Language: English	0.109**	1.115	0.053	0.042
Language: French	0.252***	1.287	0.034	0.000
Tertiary Education	-0.185***	0.831	0.026	0.000
Employment Status: Employed	0.232***	1.261	0.033	0.000
Car x Travel Time	-0.025***	0.975	0.003	0.000
PT x Travel Time	0.123***	1.131	0.004	0.000
Walk x Travel Time	-0.158***	0.854	0.016	0.000
Number of observations:	68188			
Log-Likelihood:	-29630			
McFadden R^2:	0.247			
Note:				$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

choosing an e-bike compared to individuals aged 36-59. Living in a suburban or rural area is estimated to have a higher probability of choosing the e-bike compared to individuals living in a city. The weather variables have no or a negligible effect. Being a non-swiss citizen decreases the probability of doing a stage with the e-bike by 10% compared to Swiss citizens. The household income and sizes show no significant effect compared to their reference categories. Speaking French or English does increase the odds of taking the e-bike by 11.5% and 28.7%, respectively. Also, being employed has a high positive effect on the odds of taking the e-bike, while having a tertiary education does decrease the probability. The McFadden

R^2 of 0.247 indicates an accurate model fit.

The distribution of the predictions of the probability to take the e-bike for the stages by the EBIS participants groups A and B can be seen in Figure 8. In general, the estimated probabilities have relatively low values. While many predictions are 0 or close to 0, we also observe a higher frequency of values between 0.4 and 0.6. For some stages, the probability of taking the e-bike is even close to 1. Moreover, Figure A6 in the Appendix shows that the predicted probability of switching is, on average, relatively high for traditional bike stages and relatively low for walking stages. Pt shows a high variability, while car is also at the lower end.

These estimated coefficients are applied to the individuals and their reported comparable stages of the MTMC 2021. The predicted probabilities of taking the e-bike for the respective stages are displayed in Figure 9. We discover, that the probabilities, in general, are lower than in the EBIS sample with a lot of estimates close to zero and some more frequent values around 0.25. Across modes, the predicted substitution pattern is similar to the EBIS predictions, but in general with lower probabilities, as can be seen in the Appendix in Figure A7.

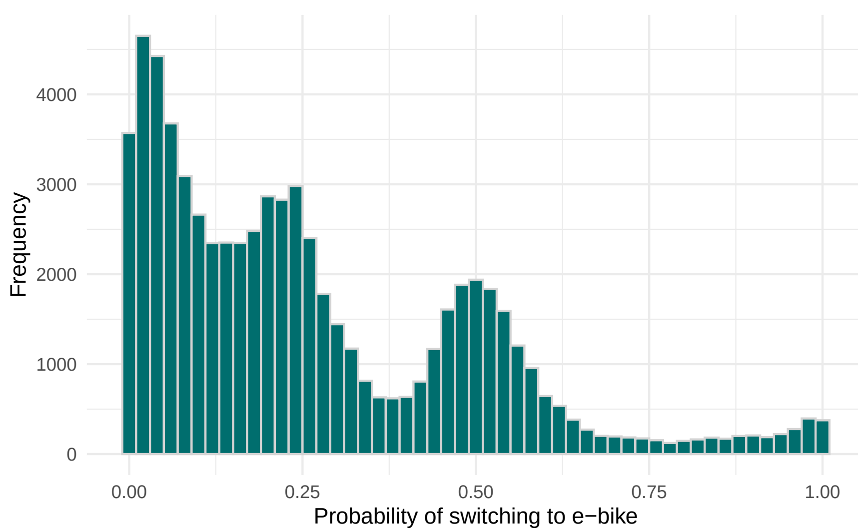


FIGURE 8: Distribution of predicted probability of substituting specific stages with an e-bike: EBIS stages and participants

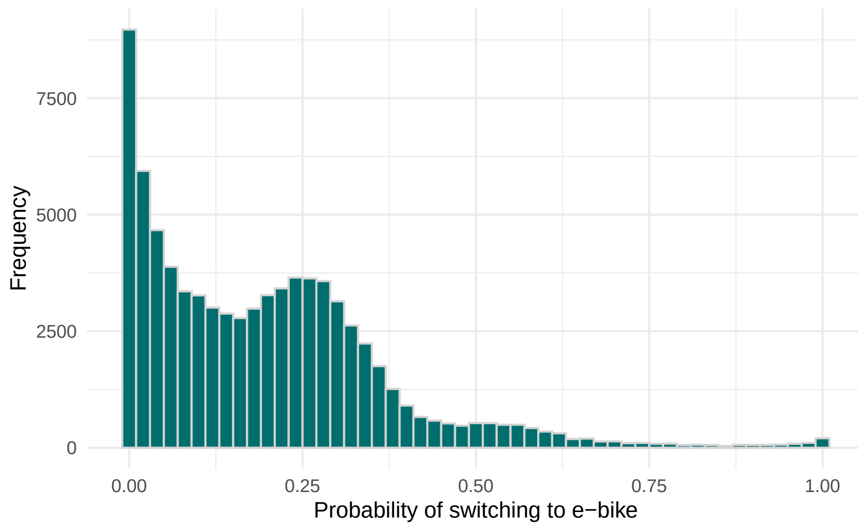


FIGURE 9: Distribution of predicted probability of substituting specific stages with an e-bike: MTMC 2021 stages and participants

Emission savings

Using the predictions of taking the e-bike for a specific stage and the underlying substituted mode of transport, we can estimate the CO_2 emission savings. In Table 16, the results are presented. The statistics concerning *Pre* are based on the emissions of the stages including the substituted mode for e-bikes, i.e., "the world pre e-bike adaption". *Post* indicates Switzerland's private transport sector with e-bikes, i.e., the current situation, where some stages are done with the e-bike, and some are not. The probability scenario indicates the potential emission savings, including the prediction of e-bike switches of the previous subsection for a specific stage by a specific individual.

TABLE 16: Emission savings for EBIS sample and MTMC 2021, in CO_2-Equivalent according to Sacchi and Bauer (2023)

Sample	Statistics, Mean	Stage	Per km	Difference g/stage	Difference g/km	Difference %
EBIS	Pre e-bikes	856.4	96.5			
	Post e-bikes	747.6	84.3	-108.9	-12.3	-12.7
	Probability scenario	683.9	77.1	-172.5	-19.4	-20.1
MTMC	Post e-bikes	1347.6	140.8			
	Probability scenario	1209.3	126.3	-138.4	-14.5	-10.3

The findings indicate that from the EBIS-participants, there already occurs a reduction of

12.7% of CO_2 emissions due to the substitution effect of the e-bike on stages between 2 km and 40 km. Including the probabilities to switch, which is calculated as outlined in Section 4.3, there could be a reduction of potentially 20.1% compared to a world without e-bikes, which corresponds to 8.5% to the current state[8]. As we do not know the substituted modes of the individuals of the MTMC 2021 sample, it is not possible to calculate the *Pre e-bikes* statistics. However, the future reduction potential, when all of these individuals would have access to an e-bike is slightly higher than for the EBIS sample with a difference of 10.3% compared to 8.5%.

[8]Reduction between the post e-bikes scenario and the probability scenario.

6 Discussion

In this thesis, substitution preferences of e-bike users were analyzed using a MNL and a MMNL. A binary logistic regression was additionally estimated to predict the substitution of other modes on the stage-level. By incorporating these findings, potential emission savings for Switzerlands' private transport sector were calculated. The first objective was to understand the factors influencing the preference for substituted modes across e-bike users and when e-bikes are used. The second objective was to use this deeper understanding of heterogeneous preferences for substituted modes to assess the impact of the e-bike on CO_2 emissions.

The results of the DCM show that among the e-bike users in the EBIS sample are differences in preferences for the substituted modes of transport due to the usage of the e-bike. The findings for older people to mainly substitute the traditional bike over the car and pt are in line with the literature outpointing that age plays a crucial role for substitution patterns (e.g., Kroesen, 2017; MacArthur et al., 2014; Rérat, 2021). However, female individuals do only differ in terms of travel time perception instead of substituted mode, which is opposed to the literature (Lee et al., 2015; Woodward et al., 2021). In general, also in Switzerland a lot of substitution occurs at the cost of traditional bikes, which is in line with the findings in the Netherlands and Denmark (de Haas et al., 2021; Haustein and Møller, 2016; Kroesen, 2017). On the other hand, the e-bike seems to be a substitution mode for car and pt as well. These shares also differ from Reck et al. (2022) in Switzerland. An explanation for this difference could occur from the recruitment of the EBIS sample as these only included e-bike users, which is different from Reck et al. (2022). This is underlined by the differences across composition in socio-demographic groups between the representative MTMC 2021 sample and the EBIS sample in Table 3.

In all models, the weather variables were not significant. Especially, in the case of active travel modes, such as e-biking, the weather normally plays a substantial role in everyday mode choices (Bucher et al., 2019; de Kruijf et al., 2021). The findings of this thesis suggest that in a retrospective survey, the weather prevalent on that day is overlooked in the response. Since typical stages were filtered for the intermediate survey and the regular substituted mode was asked for, it is possible that the "benchmark weather" is generally an average to a nice day and not an especially rainy, cold, or hot day. However, also the binary logistic regression including the GPS-tracking data, i.e., RP data, did not show large effects of weather variables, if any at all. A different operationalization of warm and cold days could be tested as well as including rain at the specific starting time of a stage or a trip, which could be a more precise approach.

The elasticites of the substituted modes concerning a change in the attributes travel time and travel costs are relatively small, but in line with other findings concerning mode choice in the literature (Litman, 2024). Walk indicates a rather high sensitivity across all changes, which is explained by the small number where it is the chosen alternative. The estimates for the VTT are comparably low. For example, in Switzerland, the value of travel time was estimated by Schmid et al. (2021) to be for motorized individual vehicles, such as cars and motorbikes, around 30.6 CHF/h, 26.7 CHF/h for walking, 18.2 CHF/h for bike, and 14.8 CHF/h for pt. However, Ben-Akiva and Lerman (1985) found that commuters may not be concerned about reducing the first 20 minutes of their commute. This attitude may be because they enjoy the time between home and work. Mokhtarian and Salomon (2001) suggests that a significant proportion of travel is undertaken for pleasure rather than as a derived travel demand. Consequently, the value of travel time may be lower on such regular short stages, which are the basis of the intermediate survey. Furthermore, as only stages are included, which are single-mode trips by default, this raises the question of whether walking is just the start or end of another mode of transport for a whole trip (e.g., walking to the bus station). As only stages above 2 km are selected for the intermediate survey, this potential threat is reduced. Still, the valuation of travel time for walk could be biased.

The binary logistic regression showed that not all individuals are equally prone to conduct a stage with the e-bike compared to other modes of transport. That female individuals are more likely to use the e-bike could be due to the fact that it lowers exertion in general and can be used for various purposes as motivated by, e.g., Lee et al. (2015), MacArthur et al. (2014) or Wolf and Seebauer (2014). Also, older individuals are more prone to use the e-bike for a specific stage, which is in line with the results of the DCM and also from other findings concerning e-bike adoption (Lee et al., 2015; MacArthur et al., 2014; Sun et al., 2020; Wolf and Seebauer, 2014). The probability of switching is rather low in general, which raises the question of whether other explanatory variables would suit better to fit the data. However, one has to keep in mind, that the prediction to switch is also relying on the prediction of the substituted mode for the e-bike stages. As the highest prediction of the substituted mode is not necessarily the actually substituted mode (as discussed in Section 5.3, there is a potential threat to these results. However, higher probabilities were estimated for the EBIS sample, indicating that the results point in the right direction as these individuals do actually also have an e-bike, which is not necessarily the case for the individuals in the MTMC 2021. This also shows that the sample composition, i.e., individuals using an e-bike, in fact, differ from those individuals (not yet) owning an e-bike.

The emission savings are in line with the findings in the literature (Bucher et al., 2019;

McQueen et al., 2020; Winslott-Hiselius and Svensson, 2017). A greater reduction potential from the current state is found for people not yet owning an e-bike (MTMC 2021 sample) compared to individuals having an e-bike (EBIS). On the one hand, one can argue that this is not surprising as the EBIS participants already own an e-bike, and therefore, an additional reduction potential is smaller. On the other, individuals in the MTMC 2021 sample could behave differently from individuals who already own an e-bike as the participants of EBIS in groups A and B. As the former did not buy an e-bike so far, there might be some underlying factors against buying or using an e-bike and, therefore, the potential in terms of emissions savings might be limited. But, as a comparison to the findings of the substitution pattern in Section 5.2 shows, we can highlight that, in general, when e-bikes are taken, a substantial reduction can be achieved. However, when accounting for the fact that not all stages are substituted, the emission reduction is put into perspective. Still, an additional decrease in stages between 2 km and 40 km of 8.5% to 10.3% shows the potential of the e-bike for a more sustainable mobility sector of Switzerland.

6.1 Policy Implications

E-bikes can help reduce the CO_2 emissions of the private transport sector in Switzerland. Therefore, encouraging the use of e-bikes through policy initiatives can further lower these negative external effects through more emission-intensive modes of transport (Scheepers et al., 2014). However, the findings of this master thesis suggest that there are varying preferences in terms of the substituted mode. Promotion of e-bikes should target individuals that have a substitution preference for car and pt. Otherwise, the potential decrease in CO_2 emissions could be negatively affected when the traditional bike and walking stages get substituted (Wolf and Seebauer, 2014). Especially, individuals not living in cities as well as French- and English-speaking individuals seem to have the potential to substitute the modes car and pt and could be addressed in targeted policy measures concerning e-bike promotion. One example could be targeted subsidies (Scheepers et al., 2014). However, whether a subsidy for e-bikes is justified from an external costs perspective cannot be fully determined from the results of this thesis and could be part of future research.

There is a clear indication of negative perception of an increase in travel time across all substituted modes. To promote e-bike usage, infrastructural elements, such as e.g. bike lanes, could decrease travel time in e-bikes and, thereof, the opportunity costs of travel time compared to other modes of transport for each stage (Wolf and Seebauer, 2014).

6.2 Limitations

Despite the contributions this thesis aims to make to the field of mode substitution through e-bikes, it also has its limitations. On the one hand, the EBIS study was conducted in Switzerland - a country characterized by high reliability and extensive rail coverage in pt. Furthermore, the pt-system is highly integrated in terms of tram, bus, and train connections. Therefore, citizens are less dependent on car usage, despite the high car ownership see, e.g., Buehler et al., 2017 compared to other countries like the United States where there is not such an extensive network of pt. These differences can highly impact mode choice. Additionally, the use of traditional bikes and walking is promoted and supported in Switzerland through networks of bike lanes and pedestrian zones, which makes non-motorized travel safer and, therefore, more attractive in mode choice compared to other countries. Furthermore, as the recruiting process also targeted only the French- and German-speaking parts of Switzerland, the preferences for transport mode choice might differ in the Italian-speaking part.

The focus of the recruitment of EBIS were e-bike owners and traditional bike owners, which comes with a selection effect on the individual level. This could be seen in the comparison to the MTMC 2021 in Table 3: The sample of the study, in general, is different from the actual population as well as the sample of the intermediate survey, which consists of groups A and B. Consequently, the application of the substitution patterns and the respective emission savings calculation onto the population of Switzerland has to be taken as an initial insight, not a conclusion. The WTP for the individuals in groups A and B in EBIS was high enough to buy an e-bike. Whether individuals with a, so far, lower WTP than the selling price will behave in the same way, we cannot say for sure.

Another shortcoming is that this thesis does not incorporate attitude variables. The selection into the study, which is connected to the selection of individuals using an e-bike, might be driven by underlying attitudes and perceptions towards mode choice, and e-bikes in particular (Biegańska et al., 2021). This is, in particular, important for policy considerations as the study of how people can be incentivized to use active travel modes by Metropia, Inc. (2023) showed. Furthermore, as SP data can provide valuable contextual insights such as asking for trip purposes and being able to validate tracking data, it also comes with some limitations. There is the possibility that the decision-making process of the choice situation may not reflect real-life behavior (Ben-Akiva et al., 1994). A typical example is when asked for their substituted mode; the participants might give idealized and socially desirable answers. They might underestimate their substituted mode of traditional bikes, as they would then admit to being lazier. Alternatively, they could not acknowledge that they used the car more often for a typical stage and would indicate that they usually cycled that stage before owning an e-bike. Furthermore, a survey is always lacking in real-life circumstances. In real-life decision-making, choices are influenced by dynamic factors, such as the weather, the mood or feeling of being healthy, or unexpected events, which can only be approximated, if at all.

6.3 Future research

The factors influencing the substituted mode were compared across different MNL and MMNL, while the latter incorporated normally distributed random coefficients. For future research, further models could incorporate error terms based on various distributions or different explanatory variables could be tested (e.g., occupation or weekday vs. weekends). Also, a DCM might not provide the best predictive model based on the available variables and their interactions. Models integrating machine learning algorithms, such as random forest or neural networks, could provide even more insights into the main drivers of preferences across individuals for substituted modes (see, e.g., Sekhar et al., 2016).

As having an e-bike is a choice per se, further research can investigate on the differences between people owning an e-bike and those who do not (yet) own an e-bike. Exploring the underlying reasons will sharpen the implications for policymakers and the potential strategies to use e-bikes as a more sustainable mode of transport. This could include doing a survey that asks more precisely about motives or attitudes towards cycling and mobility tools in general in Switzerland. Also, having an understanding of e-bike adoption mechanisms could help to examine a cost-benefit analysis, whether a subsidy for e-bikes could be an efficient tool in reducing CO_2 emissions in the private transport sector. Furthermore, the effects on other externalities in the transport sector, such as congestion, noise, or health-related external costs, could be part of future research.

7 Conclusion

This master thesis aimed to analyze substitution preferences across e-bike users using a Multinomial logit model (MNL) and a Mixed multinomial logit model (MMNL) and a Binary logistic regression analysis. The results from the discrete choice models show varying preferences across socio-demographic groups for substituted modes of transport. Older individuals are more likely to substitute traditional bikes while also being more likely to take the e-bike for a specific stage than individuals aged 36-59. While individuals not speaking German substitute walking, car, and pt stages with the e-bike, they are also more likely to take the e-bike. Female individuals experience a more significant decrease from travel time increase but are 5% more likely to take the e-bike than males. Living in the city decreases the probability of taking an e-bike, while individuals living outside of the city substitute mainly cars with the usage of the e-bike. An increase in travel time is perceived as least negative in the case of pt, while lower-income households are more cost-sensitive than higher-income households. An increase in travel distance decreases the likelihood of taking the e-bike by 14%. The overall emission reduction of these substitution effects of individuals using an e-bike is estimated to be 12.7% with a potential of reducing an additional 8.5% of CO_2 emissions. At the same time, the potential for emission savings for the whole population of Switzerland is estimated to be 10.3%. In conclusion, this master thesis undermines the promise of e-bikes as a way to decrease CO_2 emissions in the mobility sector.

Acknowledgments

I am thankful to Prof. Dr. Beat Hintermann for supervising me in this master thesis, providing expertise, and sharing knowledge in the regular discussions. Further, I want to thank the Ph.D. candidates, Jakob Roth and Laura Schwab, for providing me with the necessary resources and data and for their guidance while conducting my master thesis. Lastly, I am grateful for the infinite support of Gian and of my friends and family for their constant understanding and motivation.

References

Ahmed, F., Catchpole, J., and Edirisinghe, T. (2020). Understanding Young Commuters' Mode Choice Decision to Use Private Car or Public Transport from an Extended Theory of Planned Behavior. *Transportation Research Record: Journal of the Transportation Research Board*, 2675(3):200–211.

Andersson, A. S., Adell, E., and Hiselius, L. W. (2021). What is the substitution effect of e-bikes? A randomised controlled trial. *Transportation research part D: transport and environment*, 90:102648.

Arman, M. A., Khademi, N., and de Lapparent, M. (2018). Women's mode and trip structure choices in daily activity-travel: a developing country perspective. *Transportation Planning and Technology*, 41(8):845–877.

Astegiano, P., Fermi, F., and Martino, A. (2019). Investigating the impact of e-bikes on modal share and greenhouse emissions: a system dynamic approach. *Transportation Research Procedia*, 37:163–170.

Axhausen, K. W. (2007). *Definition of movement and activity for transport modelling: Contribution to the handbooks in transport: transport modelling*. ETH Zurich.

Bates, J. J. (1987). Measuring Travel Time Values with a Discrete Choice Model: A Note. *The Economic Journal*, 97(386):493.

Ben-Akiva, M. and Bierlaire, M. (1999). Discrete choice methods and their applications to short term travel decisions. In *Handbook of transportation science*, pages 5–33. Springer.

Ben-Akiva, M. E., Bradley, M., Morikawa, T., Benjamin, J., Novak, T., Oppewal, H., and Rao, V. (1994). Combining Revealed and Stated Preferences Data. *Marketing Letters*, 5(4):335–349.

Ben-Akiva, M. E. and Lerman, S. R. (1985). *Discrete choice analysis: theory and application to travel demand*. Number 9 in MIT Press series in transportation studies. MIT Press.

Biegańska, J., Grzelak-Kostulska, E., and Kwiatkowski, M. (2021). A Typology of Attitudes towards the E-Bike against the Background of the Traditional Bicycle and the Car. *Energies*, 14(24):8430.

Bieliński, T., Kwapisz, A., and Ważna, A. (2021). Electric bike-sharing services mode substitution for driving, public transit, and cycling. *Transportation Research Part D: Transport and Environment*, 96:102883.

Bierlaire, M. (1998). Discrete choice models. In *Operations research and decision aid methodologies in traffic and transportation management*, pages 203–227. Springer.

Bigazzi, A. and Wong, K. (2020). Electric bicycle mode substitution for driving, public transit, conventional cycling, and walking. *Transportation Research Part D: Transport and Environment*, 85:102412.

Brownstone, D. (2001). Discrete choice modeling for transportation.

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

Bucher, D., Buffat, R., Froemelt, A., and Raubal, M. (2019). Energy and greenhouse gas emission reduction potentials resulting from different commuter electric bicycle adoption scenarios in Switzerland. *Renewable and Sustainable Energy Reviews*, 114:109298.

Buehler, R., Pucher, J., Gerike, R., and Götschi, T. (2017). Reducing car dependence in the heart of Europe: lessons from Germany, Austria, and Switzerland. *Transport reviews*, 37(1):4–28.

Burnham, K. P. and Anderson, D. R. (2004). Multimodel Inference: Understanding AIC and BIC in Model Selection. *Sociological Methods amp; Research*, 33(2):261–304.

Cairns, S., Behrendt, F., Raffo, D., Beaumont, C., and Kiefer, C. (2017). Electrically-assisted bikes: Potential impacts on travel behaviour. *Transportation Research Part A: Policy and Practice*, 103:327–342.

Casier, C. and Witlox, F. (2022). An Analysis of Trip Preferences among E-bike Users in Commuting: Evidence from an Online Choice-based Conjoint Experiment. *European Journal of Transport and Infrastructure Research*, 22(1):17–41.

de Haas, M., Kroesen, M., Chorus, C., Hoogendoorn-Lanser, S., and Hoogendoorn, S. (2021). E-bike user groups and substitution effects: evidence from longitudinal travel data in the Netherlands. *Transportation*, 49(3):815–840.

de Kruijf, J., van der Waerden, P., Feng, T., Böcker, L., van Lierop, D., Ettema, D., and Dijst, M. (2021). Integrated weather effects on e-cycling in daily commuting: A longitudinal evaluation of weather effects on e-cycling in the Netherlands. *Transportation research part A: policy and practice*, 148:305–315.

DeSerpa, A. C. (1971). A theory of the economics of time. *The economic journal*, 81(324):828–846.

Dill, J. and Rose, G. (2012). Electric Bikes and Transportation Policy: Insights from Early Adopters. *Transportation Research Record: Journal of the Transportation Research Board*, 2314(1):1–6.

Fearnley, N. (2022). Factors affecting e-scooter mode substitution. *Findings*.

Federal Office for Spatial Development ARE (2023). Mobility and Transport Microcensus. `https://www.are.admin.ch/are/en/home/mobility/data/mtmc.html`. Accessed: 2024-03-01.

Fitch, D., Gao, Z., Noble, L., and Mac, T. (2022). Examining the Effects of a Bike and E-Bike Lending Program on Commuting Behavior. *Mineta Transportation Institute Publications*.

Fukushige, T., Fitch, D. T., and Handy, S. (2021). Factors influencing dock-less E-bike-share mode substitution: Evidence from Sacramento, California. *Transportation Research Part D: Transport and Environment*, 99:102990.

Fukushige, T., Fitch, D. T., and Handy, S. (2023). Estimating Vehicle-miles traveled reduced from Dock-less E-bike-share: Evidence from Sacramento, california. *Transportation Research Part D: Transport and Environment*, 117:103671.

Fyhri, A. and Fearnley, N. (2015). Effects of e-bikes on bicycle use and mode share. *Transportation Research Part D: Transport and Environment*, 36:45–52.

Gao, K., Li, A., Liu, Y., Gil, J., and Bie, Y. (2023). Unraveling the mode substitution of dockless bike-sharing systems and its determinants: A trip level data-driven interpretation. *Sustainable Cities and Society*, 98:104820.

Gao, Q., Molloy, J., and Axhausen, K. W. (2021). Trip purpose imputation using GPS trajectories with machine learning. *ISPRS International Journal of Geo-Information*, 10(11).

Goel, R., Goodman, A., Aldred, R., Nakamura, R., Tatah, L., Garcia, L. M. T., Zapata-Diomedi, B., de Sa, T. H., Tiwari, G., de Nazelle, A., Tainio, M., Buehler, R., Götschi, T., and Woodcock, J. (2021). Cycling behaviour in 17 countries across 6 continents: levels of cycling, who cycles, for what purpose, and how far? *Transport Reviews*, 42(1):58–81.

Goodman, A., Rojas, I. F., Woodcock, J., Aldred, R., Berkoff, N., Morgan, M., Abbas, A., and Lovelace, R. (2019). Scenarios of cycling to school in England, and associated health and carbon impacts: Application of the 'Propensity to Cycle Tool'. *Journal of Transport & Health*, 12:263–278.

Ha, J., Lee, S., and Ko, J. (2020). Unraveling the impact of travel time, cost, and transit burdens on commute mode choice for different income and age groups. *Transportation Research Part A: Policy and Practice*, 141:147–166.

Haustein, S. and Møller, M. (2016). Age and attitude: Changes in cycling patterns of different e-bike user segments. *International Journal of Sustainable Transportation*, 10(9):836–846.

Heinonen, S., Meyer de Freitas, L., Meister, A., Schwab, L., Roth, J., Hintermann, B., and Axhausen, K. W. (2023). The E-Biking in Switzerland (EBIS) study: Methods and dataset. Technical report, ETH Zurich.

Henning, E., Ferreira Shubert, T., and Ceccato Maciel, A. (2020). Modelling of University Student Transport Mode Choice in Joinville: A Binary Logistic Model for Active Modes. *Journal of Sustainable Development of Energy, Water and Environment Systems*, 8(4):678–691.

Hess, S. and Palma, D. (2019). Apollo: A flexible, powerful and customisable freeware package for choice model estimation and application. *Journal of Choice Modelling*, 32:100170.

Hintermann, B., Schoeman, B., Molloy, J., Götschi, T., Castro, A., Tchervenkov, C., Tomic, U., and Axhausen, K. W. (2024). Pigovian transport pricing in practice. WWZ.

Hu, Y., Ettema, D., and Sobhani, A. (2021). To e-bike or not to e-bike? A study of the impact of the built environment on commute mode choice in a small Chinese city. *Journal of Transport and Land Use*, 14(1):479–497.

Jahre, A. B., Bere, E., Nordengen, S., Solbraa, A., Andersen, L. B., Riiser, A., and Bjørnarå, H. B. (2019). Public employees in South-Western Norway using an e-bike or a regular bike for commuting – A cross-sectional comparison on sociodemographic factors, commuting frequency and commuting distance. *Preventive Medicine Reports*, 14.

Jones, T., Harms, L., and Heinen, E. (2016). Motives, perceptions and experiences of electric bicycle owners and implications for health, wellbeing and mobility. *Journal of Transport Geography*, 53:41–49.

Kroesen, M. (2017). To what extent do e-bikes substitute travel by other modes? Evidence from the Netherlands. *Transportation Research Part D: Transport and Environment*.

Lee, A., Molin, E., Maat, K., and Sierzchula, W. (2015). Electric Bicycle Use and Mode Choice in the Netherlands. *Transportation Research Record: Journal of the Transportation Research Board*, 2520(1):1–7.

Litman, T. (2024). *Understanding transport demands and elasticities.* Victoria Transport Policy Institute Victoria, BC, Canada.

MacArthur, J., Cherry, C. R., Harpool, M., and Scheppke, D. (2018). A north american survey of electric bicycle owners. Technical report, National Institute for Transportation and Communities (NITC).

MacArthur, J., Dill, J., and Person, M. (2014). Electric Bikes in North America: Results of an Online Survey. *Transportation Research Record: Journal of the Transportation Research Board*, 2468(1):123–130.

Manski, C. F. (1977). The structure of random utility models. *Theory and Decision*, 8:229–254.

McFadden, D. (1972). Conditional logit analysis of qualitative choice behavior.

McFadden, D. and Train, K. (2000). Mixed MNL models for discrete response. *Journal of applied Econometrics*, 15(5):447–470.

McQueen, M., MacArthur, J., and Cherry, C. (2020). The e-bike potential: Estimating regional e-bike impacts on greenhouse gas emissions. *Transportation Research Part D: Transport and Environment*, 87:102482.

Meister, A., Felder, M., Schmid, B., and Axhausen, K. W. (2023). Route choice modeling for cyclists on urban networks. *Transportation Research Part A: Policy and Practice*, 173:103723.

Menghini, G., Carrasco, N., Schüssler, N., and Axhausen, K. W. (2010). Route choice of cyclists in Zurich. *Transportation research part A: policy and practice*, 44(9):754–765.

Metropia, Inc. (2023). Incentivizing Active and Shared Travel Pilot Program. Prepared in collaboration with Transportation Analytics and the Behaviouralist.

Miletić, G.-M., Gašparović, S., and Carić, T. (2017). Analysis of Socio-spatial Differentiation in Transport Mode Choice Preferences. *Promet-Traffic&Transportation*, 29(2):233–242.

Mohiuddin, H., Jamal, S., and Rahman Bhuiya, M. M. (2022). To bike or not to bike: Exploring cycling for commuting and non-commuting in Bangladesh. *Transportation Research Interdisciplinary Perspectives*, 14:100614.

Mokhtarian, P. L. and Salomon, I. (2001). How derived is the demand for travel? some conceptual and measurement considerations. *Transportation research part A: Policy and practice*, 35(8):695–719.

Morikawa, T. (1989). *Incorporating stated preference data in travel demand analysis.* PhD thesis, Massachusetts Institute of Technology.

Moser, C., Blumer, Y., and Hille, S. L. (2018). E-bike trials' potential to promote sustained changes in car owners' mobility habits. *Environmental Research Letters*, 13:044025.

Philips, I., Anable, J., and Chatterton, T. (2020). E-bike carbon savings—How much and where.

Philips, I., Anable, J., and Chatterton, T. (2022). E-bikes and their capability to reduce car CO_2 emissions. *Transport Policy*, 116:11–23.

Piatkowski, D. P., Krizek, K. J., and Handy, S. L. (2015). Accounting for the short term substitution effects of walking and cycling in sustainable transportation. *Travel Behaviour and Society*, 2(1):32–41.

Pigou, A. C. (1924). *The Economics of Welfare.* Macmillan, London, 1 edition.

Plazier, P. A., Weitkamp, G., and van den Berg, A. E. (2017). "Cycling was never so easy!" An analysis of e-bike commuters' motives, travel behaviour and experiences using GPS-tracking and interviews. *Journal of Transport Geography*, 65:25–34.

Popovich, N., Gordon, E., Shao, Z., Xing, Y., Wang, Y., and Handy, S. (2014). Experiences of electric bicycle users in the Sacramento, California area. *Travel Behaviour and Society*, 1(2):37–44.

Puan, O. C., Hassan, Y. A. H., Mashros, N., Idham, M. K., Hassan, N. A., Warid, M. N. M., and Hainin, M. R. (2019). Transportation mode choice binary logit model: a case study for Johor Bahru city. *IOP Conference Series: Materials Science and Engineering*, 527(1):012066.

R Core Team (2023). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria.

Rahman, M. L. and Baker, D. (2018). Modelling induced mode switch behaviour in Bangladesh: A multinomial logistic regression approach. *Transport Policy*, 71:81–91.

Raposo, M. and Silva, C. (2022). City-Level E-Bike Sharing System Impact on Final Energy Consumption and GHG Emissions. *Energies*, 15(18):6725.

Reck, D. J. (2021). *Modelling travel behaviour with shared micro-mobility services and exploring their environmental implications.* PhD thesis, ETH Zurich.

Reck, D. J., Martin, H., and Axhausen, K. W. (2022). Mode choice, substitution patterns and environmental impacts of shared and personal micro-mobility. *Transportation Research Part D: Transport and Environment*, 102:103134.

Rérat, P. (2021). The rise of the e-bike: Towards an extension of the practice of cycling? *Mobilities*, 16(3):423–439.

Sacchi, R. and Bauer, C. (2023). Life cycle inventories for on-road vehicles.

Scheepers, C., Wendel-Vos, G., den Broeder, J., van Kempen, E., van Wesemael, P., and Schuit, A. (2014). Shifting from car to active transport: A systematic review of the effectiveness of interventions. *Transportation Research Part A: Policy and Practice*, 70:264–280.

Scheiner, J. and Holz-Rau, C. (2007). Travel mode choice: affected by objective or subjective determinants? *Transportation*, 34:487–511.

Schmid, B., Molloy, J., Peer, S., Jokubauskaite, S., Aschauer, F., Hössinger, R., Gerike, R., Jara-Diaz, S. R., and Axhausen, K. W. (2021). The value of travel time savings and the value of leisure in Zurich: Estimation, decomposition and policy implications. *Transportation Research Part A: Policy and Practice*, 150:186–215.

Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, 6(2).

Sears, J., Flynn, B. S., Aultman-Hall, L., and Dana, G. S. (2012). To Bike or Not to Bike: Seasonal Factors for Bicycle Commuting. *Transportation Research Record: Journal of the Transportation Research Board*, 2314(1):105–111.

Sekhar, C. R., Minal, and Madhu, E. (2016). Mode Choice Analysis Using Random Forrest Decision Trees. *Transportation Research Procedia*, 17:644–652.

Stock, J. and Watson, M. W. (2019). *Introduction to econometrics, global edition.* Pearson Education, London, England, 4 edition.

Sun, Q., Feng, T., Kemperman, A., and Spahn, A. (2020). Modal Shift Implications of E-bike Use in the Netherlands: Moving Towards Sustainability? *Transportation Research Part D*, 78:102202.

Ton, D. and Duives, D. (2021). Understanding long-term changes in commuter mode use of a pilot featuring free e-bike trials. *Transport Policy*, 105:134–144.

Train, K. (2003). *Discrete choice methods with simulation.* Cambridge University Press.

Truong, T. P. and Hensher, D. A. (1985). Measurement of travel time values and opportunity cost from a discrete-choice model. *The Economic Journal*, 95(378):438–451.

Van Cauwenberg, J., De Bourdeaudhuij, I., Clarys, P., de Geus, B., and Deforche, B. (2018). E-bikes among older adults: benefits, disadvantages, usage and crash characteristics. *Transportation*, 46(6):2151–2172.

Vega, A. and Reynolds-Feighan, A. (2008). Employment sub-centres and travel-to-work mode choice in the Dublin region. *Urban Studies*, 45(9):1747–1768.

Winslott-Hiselius, L. and Svensson, Å. (2017). E-bike use in Sweden – CO2 effects due to modal change and municipal promotion strategies. *Journal of Cleaner Production*, 141:818–824.

Witchayaphong, P., Pravinvongvuth, S., Kanitpong, K., Sano, K., and Horpibulsuk, S. (2020). Influential Factors Affecting Travelers' Mode Choice Behavior on Mass Transit in Bangkok, Thailand. *Sustainability*, 12(22):9522.

Wolf, A. and Seebauer, S. (2014). Technology adoption of electric bicycles: A survey among early adopters. *Transportation Research Part A: Policy and Practice*, 69:196–211.

Woodward, A., Shaw, C., and Wild, K. (2021). Gender and the E-bike: Exploring the Role of Electric Bikes in Increasing Women's Access to Cycling and Physical Activity. *Active Travel Studies*, 1(1).

Yang, Y., Wang, C., Liu, W., and Zhou, P. (2018). Understanding the determinants of travel mode choice of residents and its carbon mitigation potential. *Energy Policy*, 115:486–493.

Youssef, Z., Alshuwaikhat, H., and Reza, I. (2021). Modeling the Modal Shift towards a More Sustainable Transport by Stated Preference in Riyadh, Saudi Arabia. *Sustainability*, 13(1):337.

Yu, Y., Jiang, Y., Qiu, N., Guo, H., Han, X., and Guo, Y. (2022). Exploring built environment factors on e-bike travel behavior in urban China: A case study of Jinan. *Frontiers in public health*, 10:1013421.

Declaration of Authorship

University of Basel

Faculty of Business and Economics

Declaration of Independent Authorship

"I attest with my signature that I have written this work independently and without outside help. I also attest that the information concerning the sources used in this work is true and complete in every respect. All sources that have been quoted or paraphrased have been marked accordingly.

Additionally, I affirm that any text passages written with the help of AI-supported technology are marked as such, including a reference to the AI-supported program used.

This paper may be checked for plagiarism and use of AI-supported technology using the appropriate software. I understand that unethical conduct may lead to a grade of 1 or "fail" or expulsion from the study program.

I have read the Information Sheet on Plagiarism and Fraud dated March 3, 2022 and I am aware of the consequences of such action."

Last Name, First Name: Kipfer Noémie

Place and Date: Liestal, 07.06.2024

Signature: N. Kipfer

February 2024

A Appendix A

Descriptive Statistics of Intermediate Survey

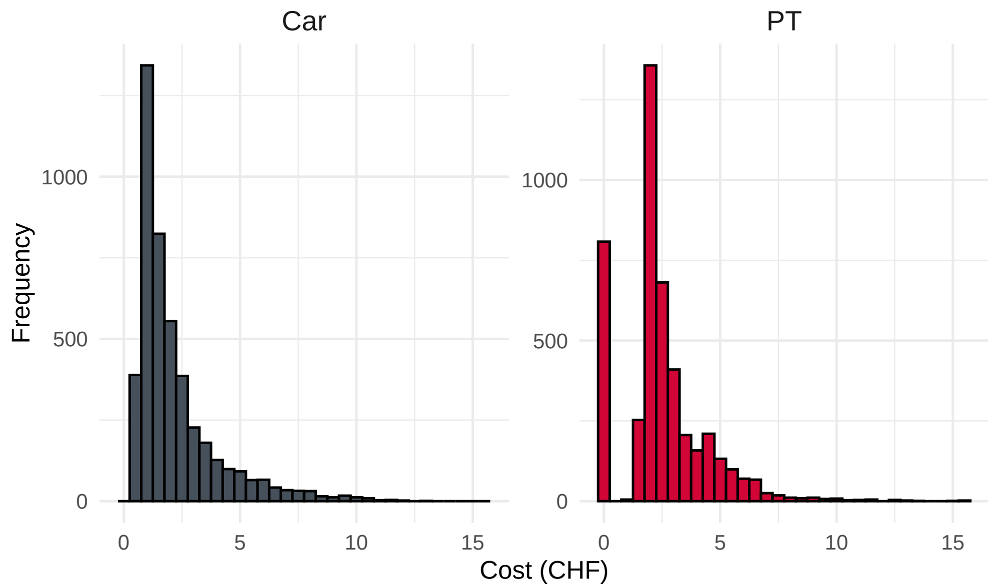


Figure A1: Travel cost in car and pt

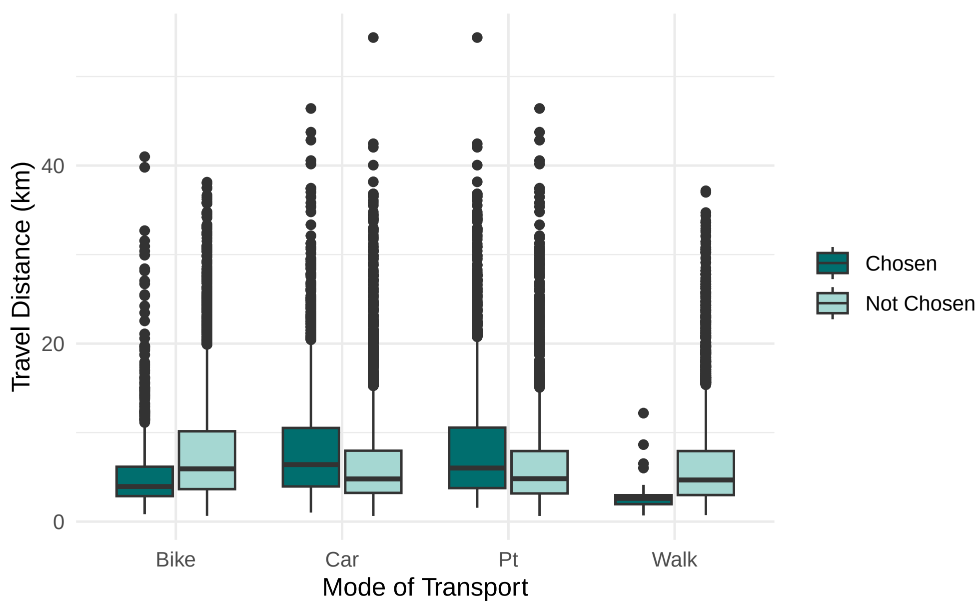


Figure A2: Boxplot of travel distance for chosen and non-chosen alternatives

Table A1: Descriptive statistics of intermediate survey data used for MNL and MMNL

Variable	Category	Percentage
Age	Young (16-29 years)	11.17
	Middle (30-59 years)	70.37
	Old (60+ years)	18.47
Citizenship	Swiss	80.97
	Non-Swiss	19.03
Education	Mandatory	0.42
	Secondary	17.13
	High	82.44
Employment Status	Employed	78.65
	Self-Employed	4.14
	Unemployed	0.35
	Student	1.05
	Retired	11.73
	Other	1.33
	NA	2.74
Gender	Female	39.47
	Male	60.32
Household Income	<= 10'000 CHF	38.69
	> 10'000 CHF	46.98
	No Answer	1.90
Language	German	86.31
	French	9.62
	English	4.07
Household Size	1	10.53
	2	34.55
	3	15.80
	4	28.23
	5 or more	10.88
Urbanity Level	City	74.86
	Suburban	18.61
	Rural	6.53
Access Car	No	19.45
	Sometimes	12.43
	Yes	68.12
Access Bike	No	22.19
	Sometimes	1.12
	Yes	76.69
		N = 1424

Additional Multinomial logit model and Mixed multinomial logit models

Table A2 shows the full models on which the parsimonous models in the main paper are built on. MNL A1 includes all situational variables, including weather variables for that specific trip. MNL A2 is the model including all socio-demographic and socio-economic variables. MNL A3 and MNL A4 are alternative specifications of the travel time interaction with the socio-demographic and socio-economic variables. Note that the models MNL A2, MNL A3, and MNL A4 are built on the significant variables of MNL A1.

Table A2: Estimation results of full MNL models and different travel time specification

Variable		**MNL A1**	**MNL A2**	**MNL A3**	**MNL A4**
ASC	ASC Walk	-2.653***	-5.364***	-2.22***	-2.617***
		(0.818)	(1.653)	(0.814)	(0.727)
	ASC Car	-1.356***	-0.88*	-0.72*	-0.882***
		(0.258)	(0.513)	(0.414)	(0.318)
	ASC PT	-0.601**	-1.744***	-1.042***	-0.814***
		(0.236)	(0.575)	(0.384)	(0.241)
	Travel Cost	-0.19***	-0.18***	-0.186***	-0.181***
		(0.038)	(0.043)	(0.043)	(0.043)
Travel Time	TT Walk	-0.066***	0.005	-0.064**	-0.068***
		(0.018)	(0.038)	(0.029)	(0.021)
	TT Bike	-0.071***	-0.081***	-0.052***	-0.061***
		(0.007)	(0.019)	(0.016)	(0.010)
	TT Car	-0.08***	-0.101***	-0.063***	-0.071***
		(0.010)	(0.034)	(0.018)	(0.012)
	TT PT	-0.048***	-0.027	-0.029*	-0.038***
		(0.006)	(0.017)	(0.016)	(0.009)
Weather	Cold weather	-0.003			
		(0.014)			
	Hot Weather	-0.148			
		(0.245)			
	Precipitation Yes	0.010			
		(0.098)			
Slope	Slope Bike	-0.091***	-0.058*	-0.058*	-0.057*
		(0.032)	(0.033)	(0.033)	(0.033)
	Slope Walk	0.012			
		(0.054)			
Purpose	Purpose Leisure Walk	0.289			
		(0.476)			

Table A2: Continued

Variable		**MNL A1**	**MNL A2**	**MNL A3**	**MNL A4**
	Purpose Other Walk	-0.061			
		(0.414)			
	Purpose Leisure Car	0.002	0.046	0.047	0.062
		(0.182)	(0.184)	(0.184)	(0.185)
	Purpose Other Car	0.307*	0.315*	0.344*	0.348**
		(0.173)	(0.176)	(0.176)	(0.177)
	Purpose Leisure PT	-0.116	-0.188	-0.145	-0.148
		(0.152)	(0.146)	(0.146)	(0.146)
	Purpose Other PT	-0.265*	-0.250*	-0.250*	-0.252*
		(0.140)	(0.134)	(0.135)	(0.134)
TT X Purpose	TT x Leisure	-0.014	-0.009	-0.010	-0.009
		(0.009)	(0.008)	(0.008)	(0.008)
	TT x Other	0.007	0.007	0.009	0.009
		(0.008)	(0.007)	(0.007)	(0.007)
Age	Age $\leq$ 35 Walk		1.493	0.185	
			(1.596)	(0.577)	
	Age $\geq$ 60 Walk		1.383	-0.365	
			(2.122)	(0.527)	
	Age $\leq$ 35 Car		-1.106**	-0.386	-0.373
			(0.456)	(0.282)	(0.276)
	Age $\geq$ 60 Car		-0.739*	-0.700**	-0.622**
			(0.398)	(0.278)	(0.249)
	Age $\leq$ 35 PT		-0.799*	0.354	0.290
			(0.461)	(0.224)	(0.208)
	Age $\geq$ 60 PT		-0.517	-0.253	-0.397**
			(0.445)	(0.268)	(0.201)
Citizenship	Non-Swiss Walk		0.095	0.76*	0.431
			(1.131)	(0.456)	(0.393)
	Non-Swiss Car		-0.396	-0.163	
			(0.326)	(0.227)	
	Non-Swiss PT		0.131	0.080	
			(0.365)	(0.219)	
Education	Tertiary Education Walk		0.914	-0.071	
			(1.439)	(0.482)	
	Tertiary Education Car		-0.124	0.008	
			(0.327)	(0.224)	
	Tertiary Education PT		0.144	0.036	
			(0.376)	(0.22)	
Employment Status	Employed Walk		2.123	-0.196	
			(1.611)	(0.478)	

Table A2: Continued

Variable		**MNL A1**	**MNL A2**	**MNL A3**	**MNL A4**
	Employed Car		-0.184	-0.558**	-0.480***
			(0.367)	(0.258)	(0.178)
	Employed PT		0.940**	0.088	
			(0.406)	(0.246)	
Gender	Female Walk		0.348	0.825**	0.737**
			(0.961)	(0.359)	(0.365)
	Female Car		-0.181	-0.233	
			(0.272)	(0.173)	
	Female PT		0.330	0.602***	0.600***
			(0.286)	(0.161)	(0.158)
HH Size	HH Size $\geq$ 4 Walk		-0.455	-0.446	
			(0.963)	(0.460)	
	HH Size $\geq$ 4 Car		-0.246	-0.332*	-0.368**
			(0.261)	(0.183)	(0.178)
	HH Size $\geq$ 4 PT		0.207	0.242	
			(0.286)	(0.167)	
Language	French Walk		3.704***	1.619***	1.775***
			(0.915)	(0.546)	(0.457)
	French Car		1.915***	1.336***	1.184***
			(0.457)	(0.295)	(0.278)
	French PT		2.071***	1.271***	1.394***
			(0.503)	(0.293)	(0.251)
	English Walk		4.489**	0.32	0.645
			(2.272)	(0.955)	(0.807)
	English Car		1.108*	0.654	0.491
			(0.66)	(0.455)	(0.401)
	English PT		0.659	0.762*	0.854**
			(0.739)	(0.433)	(0.359)
Urbanity Level	Non-city Walk		1.671*	0.556	
			(1.002)	(0.548)	
	Non-City Car		1.126***	1.103***	1.093***
			(0.280)	(0.193)	(0.190)
	Non-City PT		0.706**	0.305	0.302*
			(0.327)	(0.189)	(0.184)
TT x Socio-demographic	TT Age $\leq$ 35			-0.019*	-0.016
				(0.011)	(0.010)
	TT Age $\geq$ 60			-0.017	-0.010
				(0.012)	(0.009)
	TT Non-Swiss			-0.015	
				(0.011)	

Table A2: Continued

Variable		MNL A1	MNL A2	MNL A3	MNL A4
	TT Tertiary Education			0.007	
				(0.010)	
	TT Employed			-0.008	
				(0.012)	
	TT Female			-0.025***	-0.017**
				(0.008)	(0.007)
	TT HH Size $\geq$ 4			-0.024***	-0.019***
				(0.009)	(0.007)
	TT French			0.013	
				(0.013)	
	TT English			0.014	
				(0.021)	
	TT Non-City			0.017**	0.016**
				(0.008)	(0.007)
TT x Socio-demographic x Mode	TT Age $\leq$ 35 Walk		-0.073*		
			(0.043)		
	TT Age $\geq$ 60 Walk		-0.063		
			(0.064)		
	TT Age $\leq$ 35 Bike		-0.034**		
			(0.014)		
	TT Age $\geq$ 60 Bike		-0.015		
			(0.014)		
	TT Age $\leq$ 35 Car		0.015		
			(0.035)		
	TT Age $\geq$ 60 Car		-0.011		
			(0.030)		
	TT Age $\leq$ 35 PT		0.006		
			(0.014)		
	TT Age $\geq$ 60 PT		-0.008		
			(0.013)		
	TT Non-Swiss Walk		0.003		
			(0.031)		
	TT Non-Swiss Bike		-0.012		
			(0.012)		
	TT Non-Swiss Car		0.007		
			(0.020)		
	TT Non-Swiss PT		-0.014		
			(0.012)		
	TT Tertiary Education Walk		-0.015		
			(0.041)		

TABLE A2: Continued

Variable		**MNL A1**	**MNL A2**	**MNL A3**	**MNL A4**
	TT Tertiary Education Bike		0.014 (0.011)		
	TT Tertiary Education Car		0.035 (0.024)		
	TT Tertiary Education PT		0.010 (0.011)		
	TT Employed Walk		-0.058 (0.039)		
	TT Employed Bike		0.015 (0.014)		
	TT Employed Car		0 (0.027)		
	TT Employed PT		-0.019 (0.012)		
	TT Female Walk		-0.013 (0.027)		
	TT Female Bike		-0.031*** (0.010)		
	TT Female Car		-0.041* (0.022)		
	TT Female PT		-0.022** (0.009)		
	TT HH Size ≥4 Walk		-0.025 (0.025)		
	TT HH Size ≥4 Bike		-0.025*** (0.010)		
	TT HH Size ≥4 Car		-0.034* (0.020)		
	TT HH Size ≥4 PT		-0.024*** (0.009)		
	TT French Walk		-0.034 (0.023)		
	TT English Walk		-0.124 (0.097)		
	TT French Bike		0.035** (0.015)		
	TT English Bike		0.015 (0.023)		
	TT French Car		0.009 (0.027)		

Table A2: Continued

Variable		**MNL A1**	**MNL A2**	**MNL A3**	**MNL A4**
	TT English Car		-0.017		
			(0.044)		
	TT French PT		0.004		
			(0.015)		
	TT English PT		0.017		
			(0.024)		
	TT Non-City Walk		-0.012		
			(0.023)		
	TT Non-City Bike		0.023***		
			(0.009)		
	TT Non-City Car		0.028		
			(0.018)		
	TT Non-City PT		0.011		
			(0.008)		
Travel Cost x HH Income	Cost HH Income $\leq$ 10'000 CHF		-0.084*	-0.077	-0.087*
			(0.051)	(0.05)	(0.048)
	Cost HH Income No Answer		0.144	0.141	0.13
			(0.135)	(0.129)	(0.132)
Number of Parameters		21	57	39	87
Number of Choice Obs.		4304	4437	4437	4437
LL(0)		-5159.229	-5317.765	-5317.765	-5317.765
LL(final)		-3779.632	-3738.182	-3750.444	-3700.639
McFadden R^2		0.2674	0.2970	0.2947	0.3041
Adjusted R^2		0.2633	0.2863	0.2874	0.2877
AIC		7601.264	7590.364	7578.888	7575.278
BIC		7734.978	7955.035	7828.399	8131.881

Note: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Reference Mode: Traditional Bike

TABLE A3: Estimation results of the MMNL without trip purposes for the prediction of substituted modes across e-bike stages

Variable	**MMNL No Trip Purposes**
ASC Walk	-2.203**
	(0.946)
ASC PT	-2.812***
	(0.724)
ASC Car	-2.05***
	(0.678)
Travel Cost	-1.519***
	(0.557)
TT Walk	-0.237***
	(0.038)
TT Bike	-0.158***
	(0.028)
TT Car	-0.255***
	(0.043)
TT PT	-0.09***
	(0.014)
Slope Bike	0.011
	(0.106)
Age $\leq$ 35 Car	-1.684*
	(0.996)
Age $\geq$ 60 Car	-1.457
	(1.095)
Age $\leq$ 35 PT	-1.717*
	(0.968)
Age $\geq$ 60 PT	-1.517
	(1.15)
Employed PT	1.078**
	(0.455)
French Walk	4.961***
	(1.563)
English Walk	2.777
	(1.814)
French Car	4.222***
	(1.48)

Table A3: Continued

Variable	**MMNL No Trip Purposes**
English Car	2.498 (1.896)
French PT	4.899*** (1.44)
English PT	2.871 (1.765)
Non-city Walk	2.296** (0.932)
Non-City Car	2.454*** (0.774)
Non-City PT	1.391* (0.755)
TT Age $\leq$ 35 Walk	-0.047 (0.031)
TT Age $\geq$ 60 Walk	-0.046 (0.04)
TT Age $\leq$ 35 Bike	-0.05 (0.034)
TT Age $\geq$ 60 Bike	0.006 (0.036)
TT Female Bike	-0.064*** (0.023)
TT Female Car	-0.12** (0.049)
TT Female PT	-0.02 (0.014)
TT HH Size $\geq$4 Bike	-0.033 (0.025)
TT HH Size $\geq$4 Car	-0.017 (0.05)
TT HH Size $\geq$4 PT	-0.014 (0.017)
TT French Bike	0.012 (0.033)

Table A3: Continued

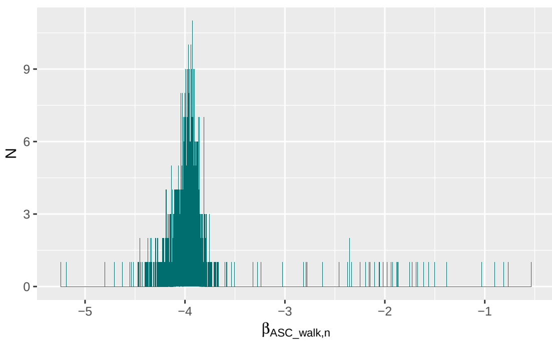
Variable	**MMNL No Trip Purposes**
TT English Bike	0.014
	(0.066)
TT Non-City Bike	-0.02
	(0.033)
Cost HH Income ≤ 10'000 CHF	-0.188
	(0.135)
Cost HH Income No Answer	0.412
	(0.834)
σ_{Walk}	0.849*
	(0.456)
σ_{PT}	-1.231*
	(0.679)
σ_{Car}	-1.206**
	(0.538)
σ_{Bike}	3.112***
	(0.824)
$\sigma_{TravelCost}$	-0.772***
	(0.226)
σ_{TTWalk}	-0.076***
	(0.013)
σ_{TTBike}	-0.137***
	(0.038)
σ_{TTCar}	-0.285***
	(0.033)
σ_{TTPT}	0.026**
	(0.013)
$\sigma_{SlopeBike}$	-0.869***
	(0.138)
Number of Estimated Parameters	48
Number of Choice Obs.	4437
LL(0)	-5317.765
LL(final)	-2715.931
McFadden R^2	0.489
Adjusted R^2	0.480

TABLE A3: Continued

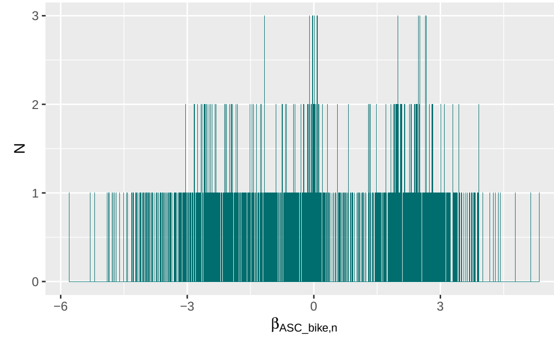
Variable	**MMNL No Trip Purposes**
AIC	5527.861
BIC	5834.953

Note: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Reference Mode: Traditional Bike

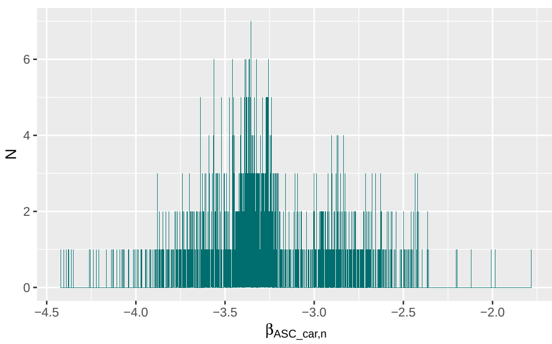
Densities of Estimated Random Coefficients in Mixed Multinomial Logit Model



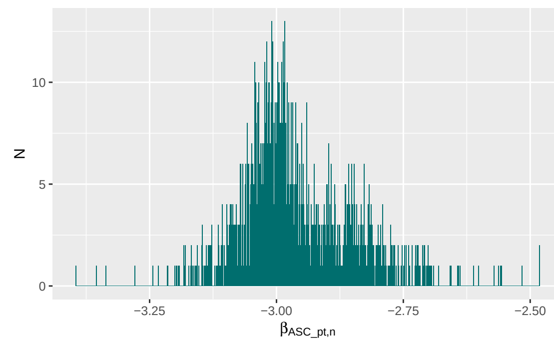
(A) Walk



(B) Bike

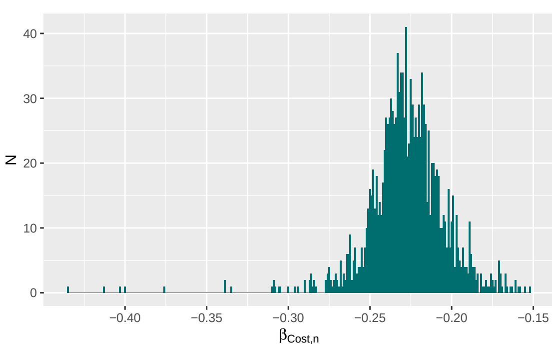


(C) Car

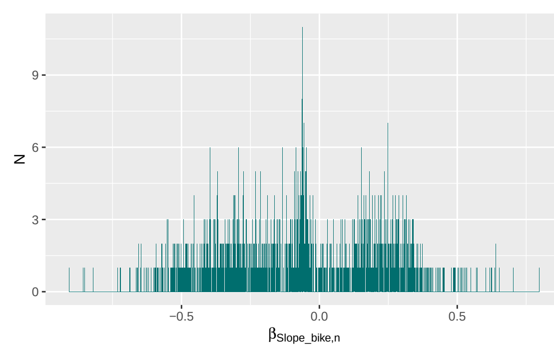


(D) pt

FIGURE A3: Estimated β_{mode}

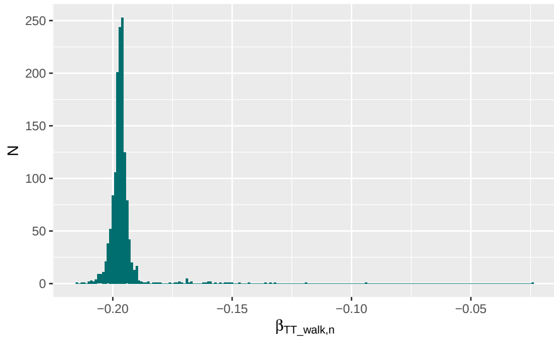


(A) Cost coefficient

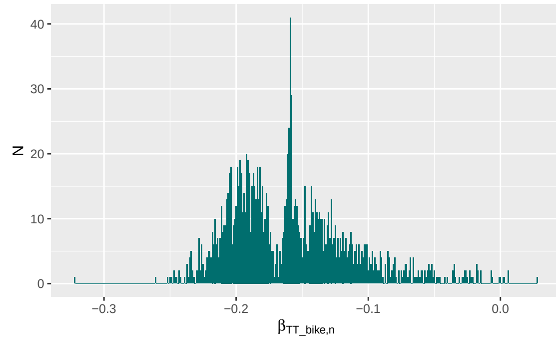


(B) Slope Bike Coefficients

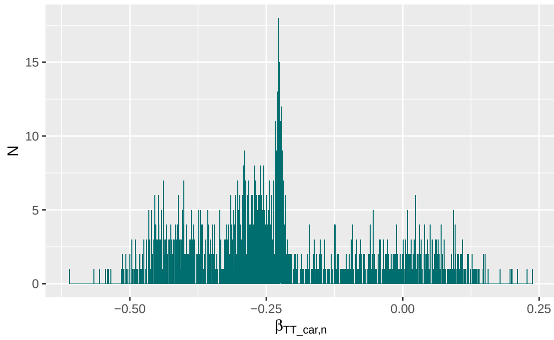
FIGURE A4: Density of cost and slope coefficients of the estimated MMNL with socio-demographic variables



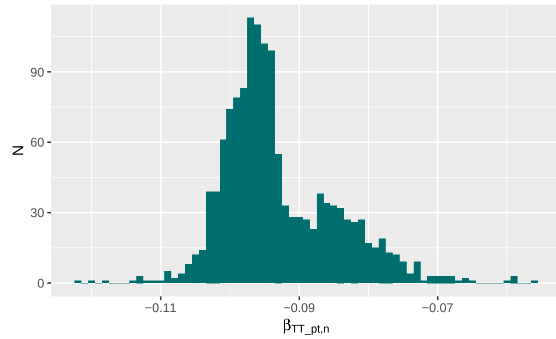
(A) TT-Walk



(B) TT-Bike



(C) TT-Car



(D) TT-PT

FIGURE A5: Estimated $\beta_{TT-Mode}$

Mixed Multinomial Logit Model Prediction Statistics

Table A4: Summary statistics of prediction of e-bike stages of the intermediate survey

	Statistic	walk	bike	car	pt	chosen
1	Min	0.00	0.00	0.00	0.01	0.00
2	Max	0.34	0.93	0.98	1.00	1.00
3	Mean	0.01	0.38	0.28	0.33	0.48
4	Median	0.00	0.41	0.27	0.29	0.48

Table A5: Summary statistics of prediction of comparable e-bike stages not in the intermediate survey

	Statistic	walk	bike	car	pt	chosen
1	Min	0.00	0.00	0.00	0.00	0.00
2	Max	1.00	1.00	1.00	1.00	1.00
3	Mean	0.02	0.32	0.32	0.34	0.39
4	Median	0.01	0.34	0.32	0.30	0.39

Binary Logistic Regression: Predicted Switching-Probabilities for each mode

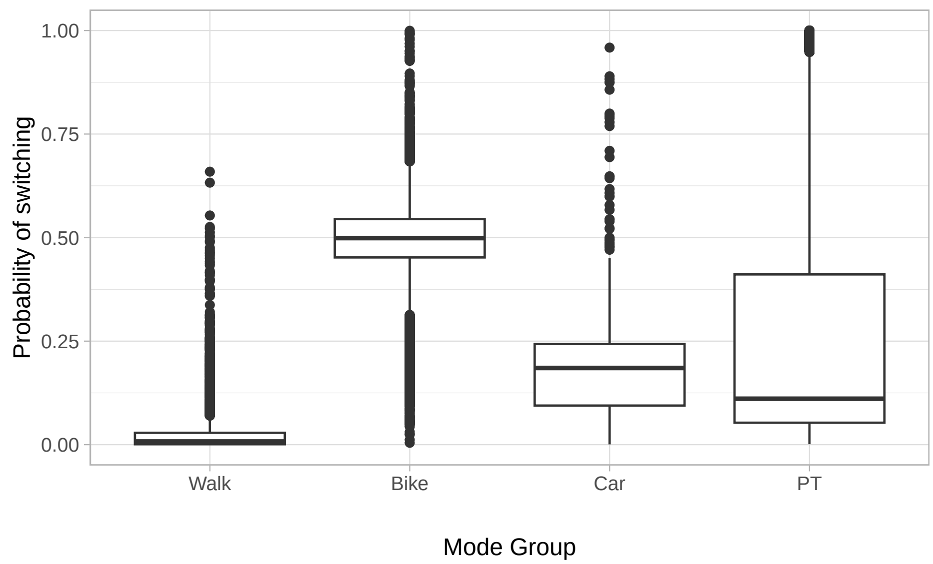


Figure A6: Switching probability across each mode, comparable EBIS stages

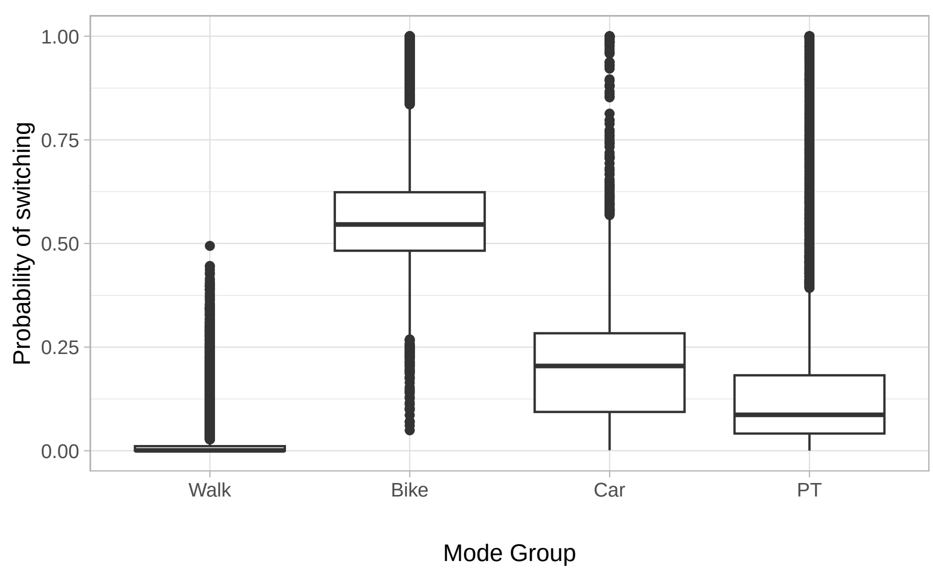


Figure A7: Switching probability across each mode, comparable MTMC stages