# Cyclists Station Choice

## Investigating railway station choice for cyclists including perceived social safety and comfort factors

A. A. Barneveld


TU Delft

Goudappel
MOBILITEIT BEWEEGT ONS

# Cyclists Station Choice

## Investigating railway station choice for cyclists including perceived social safety and comfort factors

by


A. A. Barneveld


to obtain the degree of Master of Science in
Transport, Infrastructure & logistics
at the Delft University of Technology,
to be defended publicly on May 23, 2022

| | | |
|---|---|---|
| Student number: | 4476514 | |
| Project duration: | November 26, 2021 – May 23, 2022 | |
| Thesis committee: | Prof. dr. G. P. van Wee, | TU Delft |
| | Dr. ir. N. van Oort, | TU Delft |
| | Dr. J. A. Annema, | TU Delft |
| | Drs. ir. R. Huisman, | Goudappel |

TU Delft

Goudappel
MOBILITEIT BEWEEGT ONS

# Preface

This thesis, ‘Cyclists station choice, investigating station choice for cyclists including perceived social safety and comfort factors’, is the result of my graduation project that started seven months ago. It is the last step in obtaining my Master’s degree in Transport, Infrastructure and Logistics at TU Delft.

I would like to thank Goudappel for the opportunity to perform this research at this company. I would especially like to thank Raymond Huisman, who introduced me to several colleagues and was willing to meet with me every week. Thank you, Raymond, for your involvement and willingness to help. Apart from Raymond, I am very grateful for the help of several colleagues at Goudappel, who helped me at the start of my thesis by having conversations about subject-related matters and improving my recommendations at the end of my thesis. I would like to thank my graduation committee for their guidance and feedback over the last months. Especially my daily supervisors, Jan Anne Annema and Niels van Oort who were always available when I had questions or when I needed some steering. Many thanks to Niels van Oort who introduced me to Goudappel and organised several fun and interesting meetings with thesis students involved in public transport. I would also like to thank Bert van Wee, as the chair of my committee, for his detailed comments and ideas for improving my thesis.

During the development of my models in Matlab, Sven helped me with resolving some errors. Therefore, I would especially want to express my gratitude to my boyfriend. On top of the study related comments, I also want to express my gratitude towards my (study)friends and my family for distractions and joyful events outside of my study hours. I am sure that this support helped me keep my positive attitude and the successful closure of my study years in Delft.

*A. A. Barneveld*
*Delft, May 2022*

# Summary


## Introduction

To reach the Paris Climate Agreement goals, The Netherlands should reduce its greenhouse gasses significantly in the coming years. Transport plays a significant role in emitting these gasses (Stanley et al., 2011). Besides sustainability goals, ongoing urbanisation takes place. The bicycle-train combination could play a prominent role in accommodating the growth of transport in a sustainable manner since it can compete with the mode car on speed and accessibility (Kager et al., 2016). The bicycle-train combination combines the speed of the train with the connectivity of the bicycle. Currently, over 50% of train passengers already use their bicycle as an access mode for the train (ProRail, 2019). The expanding popularity of the bicycle as a connection mode results in high parking demands at stations (van Boggelen, 2008). The largest bicycle parking facility in The Netherlands already accommodates more than 12.500 bicycles. Increasing the capacity of bicycle parking facilities at these stations requires large investments and is not always possible since these facilities have to be built in dense city centres. However, travellers could often find more than one train station within 20 minute cycle time within urban regions.

To investigate if the needed capacity of bicycle parking for travellers can be accommodated at local stations, understanding of the incentives why people cycle to which station is required. Existing research focuses primarily on dissatisfier factors, such as travel time and costs, presence of transfers and waiting time (van Mil et al., 2020). These dissatisfiers belong in the lower layers of the customer pyramid of needs (figure 1). However, van Hagen concluded that people do not have a sense of time. He states that the factor travel time is influenced by time perception time, which is influenced by other factors. Next to the factor access time, the factor travel costs is highly dependent on personal characteristics, e.g. income and age. So time perception could be different for various user groups. Therefore, the focus lies on including satisfier characteristics such as (perceived) safety and comfort, referring to the upper layers of the pyramid (figure 1), instead of only including dissatisfier characteristics like existing research does.

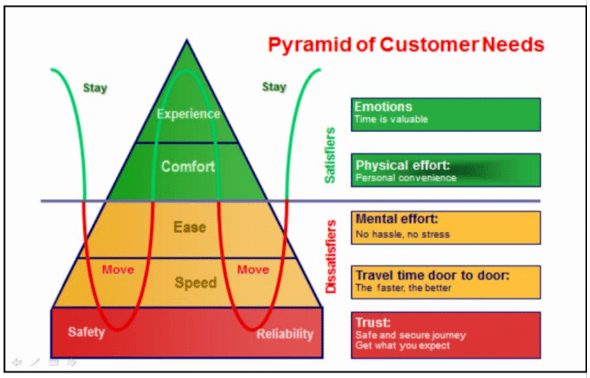



Figure 1: Pyramid of Customer Needs (Hagen and Bron, 2014)

The research question of this study is: "To what extent can satisfiers, as opposed to dissatisfiers, play a role in station choice for bicycle-train travellers regarding local railway stations?"

## Ranking Satisfiers

To answer this question, first, all possible satisfier characteristics influencing station choice by bicycle are found in a literature study. These factors are ranked by 21 respondents. The respondents are chosen such that all seven user classes of the bicycle-train combination, as found by Shelat et al. (2018), are represented by three respondents. After weighing the results to the dynamics of society, the following overall top 3 ranking is:

1. Level of shelter at the station
2. Level of lighting at the station
3. Presence of security in the bicycle station

This ranking step is solely taken to gain explorative insights into the prioritisation of factors that influence station choice. Due to the small respondents group (N=21), this step might not be statistically significant. This means the final ranking might not be fully representative for the whole population.

In this research, station choice is prioritised above investigating satisfiers, as the total station choice is not made only on the upper layers of the customer needs pyramid. For that reason, two important dissatisfiers related to station choice are included in the rest of this research: access travel time and bicycle parking costs (Ton et al., 2019; Krabbenborg et al., 2015; Weliwitiya and Eng, 2020; Young and Blainey, 2018; Givoni and Rietveld, 2007).

## Stated Choice Experiment

The top five factors do not have an equal influence on station choice. This was further investigated by using a stated choice experiment. First, a pilot survey is taken by 38 respondents to retrieve prior values. Prior values are used to create an efficient design for the final stated choice survey (ChoiceMetrics, 2018). In this final stated choice survey, 308 respondents were asked to choose between station A and station B, having different values for the parameters: level of lighting, level of shelter, presence of security, cycling access time and bicycle parking costs. All respondents varied on several socio-demographic factors, such as gender, age, income, education, social participation and household composition, train frequency and trip purpose. The distribution of socio-demographic characteristics of respondents in the final survey was very similar to the total population of bicycle-train travellers.

## Multinominal Model Results

The distribution of weights by respondents is modelled using an MNL model, with the utility function as presented in equation 1.

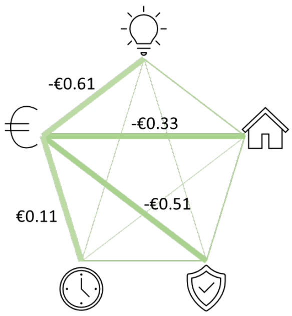
$$U_i = \beta_{Light} * Light_i + \beta_{Shelter} * Shelter_i + \beta_{Security} * Security_i + \beta_{Time} * Time_i + \beta_{Price} * Costs_i \quad (1)$$

After modelling the responses on the final survey regarding station choice using an MNL model, the following weights are found (table 1). The number of attribute levels is also shown in table 1. When an attribute has two levels, 0 and 1 are implied. A value of zero means a low level of lighting, shelter or security. A value of one indicates a high level. For access time and parking costs, three levels are used. These levels are 5, 7.5 or 10 minutes access time and €0, €0.50 or €1. The minimum and maximum columns of table 1 refer to the utility difference when the minimum and maximum attribute values are multiplied by the parameter values. This indicates the range of the effect on the total utility function for each parameter.

The weights indicate a change of utility when the parameter changes by one. Meaning when lighting is increased, the utility could increase by 0.667, and when parking costs are increased by €1, the utility of that station decreases by -1.09. Regarding the range of attributes, it is visible that the average traveller using bicycle-train is the most sensitive to changes in the access time or parking costs parameter.

Table 1: Overall Parameter Weights

| Factor | Value | # levels | minimum | maximum |
|---|---|---|---|---|
| Light | 0.667 | 2 | 0 | 0.667 |
| Shelter | 0.364 | 2 | 0 | 0.364 |
| Security | 0.554 | 2 | 0 | 0.554 |
| Access time | -0.122 | 3 | -0.610 | -1.22 |
| Parking costs | -1.09 | 3 | 0 | -1.09 |



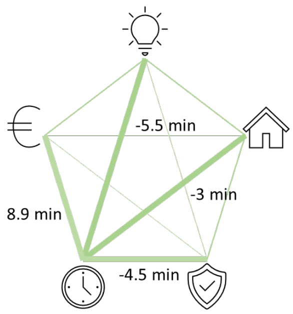




Figure 2: Costs as base (MNL) Figure 3: Time as base (MNL)

Parameter weights can only be compared with different models when a ratio is computed, for example the Value-of-Time. In order to compare the results with other models, and to show the results more clearly, all parameter ratios with the dissatisfiers costs and time are computed. The ratio values are in the unit of the base parameter: euros (figure 2) and minutes (figure 3). It can be concluded respondents are willing to pay €0.61 parking costs or to cycle 5.5 minutes extra for a well-lighted station.

## Latent Class Model Results

The parameter weights resulting from the MNL model are an average of all respondents. To allow heterogeneity between user groups, a Latent Class Model is created. A latent class choice model is applied to gain more insights into the differences between bicycle-train travellers. Two classes are found based on similar weights of respondents. Afterwards, it is analysed which socio-demographic characteristics explain class membership. The characteristics of the two user groups are presented in figure 4.

This figure shows that females, elderly and non-students are most likely to belong in group 1. The first group is sensitive to lighting, security and access travel time. Group 1 has an overall size of 35% of all travellers, which makes group 2 65%. Males and students have a probability of 83% for belonging to the second group. Group 2 is highly sensitive to access travel time and parking costs. Each class has different values for the parameter weights, shown in table 2a and 2b. In figures 5a and 5b, all parameter weights for both classes are calculated using the attribute levels.

Table 2: Final Results Latent Class

(a) Group 1 Parameters

| Parameter | Weight | Minimum | Maximum |
|---|---|---|---|
| $\beta_{Light}$ | 0.765 | 0 | 0.765 |
| $\beta_{Shelter}$ | 0.282 | 0 | 0.282 |
| $\beta_{Security}$ | 0.610 | 0 | 0.610 |
| $\beta_{Time}$ | -0.107 | -0.545 | -1.07 |
| $\beta_{Costs}$ | -0.258 | 0 | -0.258 |

(b) Group 2 Parameters

| Parameter | Weight | Minimum | Maximum |
|---|---|---|---|
| $\beta_{Light}$ | 0.588 | 0 | 0.588 |
| $\beta_{Shelter}$ | 0.738 | 0 | 0.738 |
| $\beta_{Security}$ | 0.532 | 0 | 0.532 |
| $\beta_{Time}$ | -0.272 | -0.750 | -2.72 |
| $\beta_{Costs}$ | -3.411 | 0 | -3.411 |

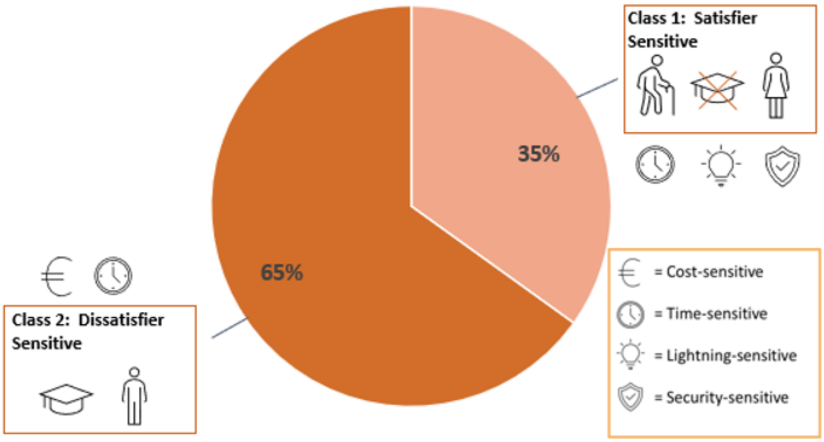
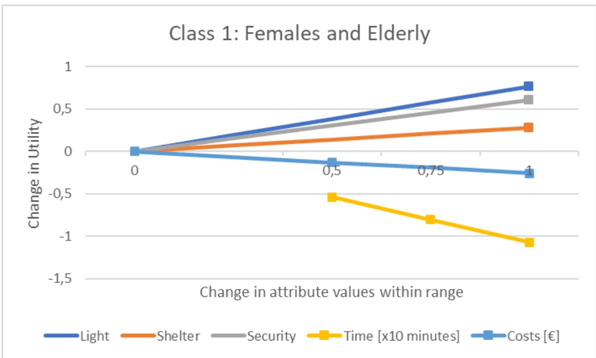
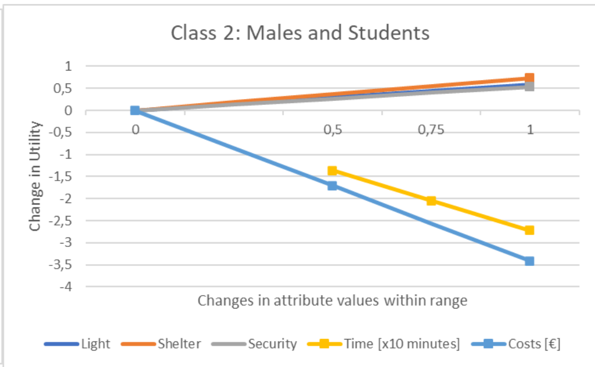



Figure 4: Summary of two classes



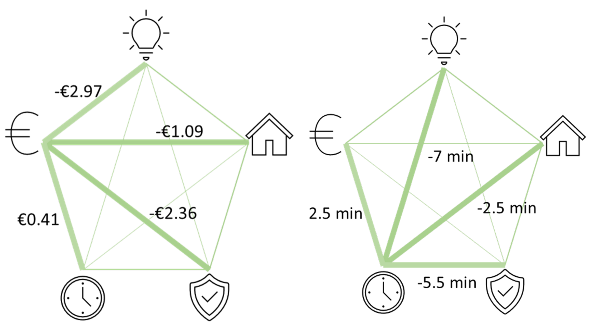

(a) Group 1: Changes in utilities for different attribute levels



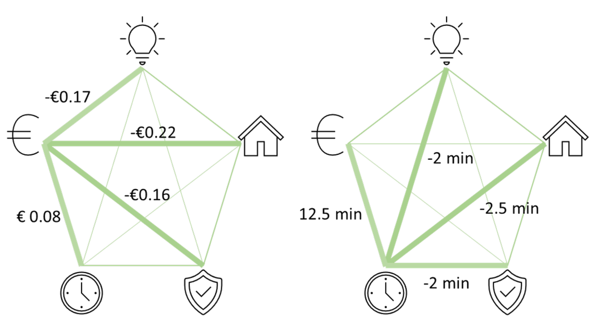

(b) Group 2: Changes in utilities for different attribute levels

Figure 5: Sensitivity to parameters of both classes

Both classes and their ratio weights for all parameters with parking costs and access time as a base, are shown in figures 6a and 6b. It is visible that the first group is willing to pay almost €3 to increase lighting from low to high, whereas the second group is only prepared to pay €0.17. Knowing the weights vary among the two user classes, it can be concluded that different measures are needed to appeal to both user classes to a station.




(a) Group 1 (Costs and Time as base parameter)




(b) Group 2 (Costs and Time as base parameter)

Figure 6: Pentagon Interrelation Costs and Time

Measures that would affect the access travel time would be most efficient for increasing the utility of a local station. Unfortunately, these measures are also the highest of costs and most technical and politically challenged.

Furthermore, group 1 highly appreciates lighting at the station and security in the bicycle parking. It could be imagined that increasing lighting at the station and security in the bicycle parking has fewer costs and challenges compared to changing the infrastructure.
For group 2, the satisfier shelter at the station has the highest positive weight of all parameters. Meaning possible extra dis-utilities of access time and parking costs could be partly compensated by placing a full shelter cubicle on the platform. Providing more shelter will also affect the first group, and providing more lighting and security is also of influence for the second group, however, in a smaller ratio.

Next to actual access travel time, which is imagined to be very difficult to adapt, also the dissatisfier parking costs is very important for the second group. Decreasing the parking price for the local station compared to the central station is not always possible, for example, when parking is free at both stations. Another optional measure could be to increase the parking price at the central station, possibly leading to more travellers choosing the local station instead. Before implementing such measures, more research concerning the acceptance of measures and actual behaviour is needed.

An example of the municipality Delft shows the effect of adapting various satisfier factors for different user groups. Imagine a traveller belonging to group 2, who lives 5 cycling minutes from station Delft and 10 minutes from Delft Campus. When adding a parking fee of €0.50 per day at station Delft and adding extra shelter at Delft Campus, the traveller has a probability of 55% for choosing station Delft and 45% for choosing Delft Campus. Concluding, under these circumstances, the probabilities for both stations are valid for 62% of bicycle-train travellers (group 2 in Delft).
When a traveller (group 1) lives 5 and 10 minutes respectively from Delft and Delft Campus, and shelter and security are added at Delft Campus. This would result a probability that 63% of 38% of train travellers will choose station Delft and 37% will choose Delft Campus. This share of 37% is compared to 19% for Delft Campus when low shelter or no security is present.

## Conclusion

This research shows that satisfiers can influence station choice. However, the effect of satisfiers on station choice depends on the socio-demographic characteristics of the traveller. For the first group (35%), the satisfiers lighting and security in the bicycle parking have a large influence on the station choice. Regarding the results and possible measures, it is seen that for the largest group (65%), the dissatisfiers have a larger influence on station choice. Therefore, the implementation of measures regarding the dissatisfiers is more effective than measures regarding the satisfiers for this group. Meaning satisfiers do influence station choice. However, the extent of influence caused by satisfiers depends on the socio-demographic characteristics of the travellers and the disutility caused by the dissatisfiers.

## Discussion and Recommendations

This research provides new insights into the role of satisfiers on station choice. However, there are several ideas to improve and expand this research. As already stated in this paper, the ranking was done on a small test group (N=21). This step was only indicative in this research. Nonetheless, it would be interesting to perform this ranking on a larger group and compare the resulting top 3. If other factors are ranked most important, these could be included in a new stated choice experiment.
Regarding dissatisfiers, access time and parking costs are researched. This research did not include other dissatisfiers, such as transfers within the train trip, and their relation to station choice. According to van Mil et al. (2020), transfers are of high importance in the station choice for travellers. Excluding transfers means that the results are applicable for train trips that can be taken from both central and local stations, or for travellers indifferent against transfers. For further research it is recommended to include other dissatisfiers combined with satisfiers in a stated choice experiment.
It is also recommended to further develop analysis using Latent Class Modelling. The differences between the resulting weights from the MNL model and the LC model in station choice are visible. Showing that two significant classes are explained by socio-demographic factors. It shows that it is

important to allow for heterogeneity in the population and not only take the average value. Allowing heterogeneity leads to results being better applicable in reality.

Practical advice for municipalities and station owners based on this research is to not solely focus on central stations when renewing or designing and keep the local stations up to travellers' standards. It could also be recommended to make a Cost-Benefit analysis for each station specific. An indication is given of how satisfiers influence travel behaviour. However, in this example of Delft, access time is stated as 5 minutes for Delft and 10 minutes for Delft Campus. Varying the access time and applying a station specific Cost-Benefit analysis allows a more accurate forecast of station choice by travellers and results in many insights into the advantages and disadvantages of increasing station lighting, shelter and security for a specific station.

# Contents

# List of Figures

# List of Tables

# 1
# Introduction

To be able to reach the climate goals of Paris 2020 and the sustainable development goals of the United Nations 2030, governments are interested in promoting sustainable travel modes. Transportation by walking and cycling are the most sustainable transport options. Since these options require more physical effort than motorised transport and are relatively slower, walking and cycling are mainly popular for smaller distances (Metz, 2014). A mode more suitable for longer distances and more sustainable than personal car ownership is public transport (Currie and Stanley, 2008).

To transport multiple passengers, public transport uses bundling of time (set of departure times) and bundling of space (amount of stops) (Sáez et al., 2012). This bundling of space results from the origin or destination of a trip hardly being at the station itself. This means access and egress modes are often needed (Brands et al., 2014). Using a bicycle as an access mode solves a fundamental weakness of public transport: the accessibility of public transport stops (Martens, 2006). An example of a multi-modal trip, like the bicycle-train mode combination, is given in figure 1.1. In this figure, two transport options are given as an example. In trip A, the trip from home to office requires 10 minutes of biking towards the station, 30 minutes of train travel, and a 5 minute walk from the station to the office. To illustrate the infinite possibilities, trip B is a different multi-modal trip, where only bicycle and public transport are used.

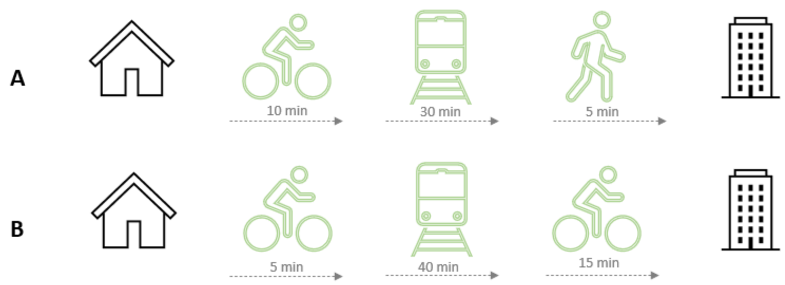



Figure 1.1: Multi-modal Trip Options

Several train stations in The Netherlands are coping with a capacity problem and crowdedness (Jacobs, 2019). This crowdedness could be translated to overfull bicycle parking (Goeverden and Correia, 2018) or crowded and unsafe events in the station or on the platform (Zhang et al., 2020) and overall capacity problems. The reason for this rising crowdedness could be the increasing number of activities, population growth and centralisation. Due to these trends, it would be unlikely that this problem would stabilise by itself, and the capacity problems could become even more prominent in the future.

According to van Mil et al. (2020), the flexibility of the bicycle combined with the speed and comfort of public transport can be a highly competitive alternative to the car. The advantages of the bicycle-train combination are in combat with the disadvantages of station crowding. With the intention of stimulating the bicycle-train combination for solving the climate problem, the capacity problems will increase. This capacity problem becomes even more complex when there is no possibility to expand, for example in dense areas.

## 1.1 Problem Statement

Municipalities and provinces in The Netherlands are aware that city centres and large central stations are reaching their maximum capacity (Jonkeren and Kager, 2021; Goeverden and Correia, 2018). However, due to the central location of these central stations, expanding is often not an easy solution. Research is needed regarding the collaboration between local railway stations and central stations to realise an equal distribution of passengers over stations. If passengers are more equally distributed (from central to local stations), the local stations could support central stations in bicycle parking capacity. To do this successfully, research is needed not only for the central stations in The Netherlands but also for the local railway stations. Using information regarding station choice, the network could be anticipated on its rising demand.
Mainly central stations are coping with capacity problems in bicycle parking facilities (Jonkeren and Kager, 2021). Unfortunately, these central stations are often limited for expanding options due to the dense areas. Ton et al. (2019) found that increasing transit capacity comes at high costs, while better integration with cycling serves as a sustainable and (cost-)efficient alternative.

To illustrate the problem, and possible solutions, of larger central stations and smaller local railway stations, an example of Rotterdam is given. In Rotterdam, many train stations are located, as shown in figure 1.2. In this illustration, the municipality of Rotterdam is indicated by a black border. All train stations in and nearby this region are pointed out, the central station being larger than the others. The many stations in this concentrated area are visible. Amongst all stations, the largest is Rotterdam Central. Due to its central location and infrastructure connections, Rotterdam Central has the most passengers and, therefore also the most capacity problems. In theory, there are a lot of stations and thus choices to be made for the passengers. An average 50% of all passengers in Rotterdam, chooses for the central station. This share of passengers has remained constant throughout all stations in Rotterdam for several years (figure 1.3).

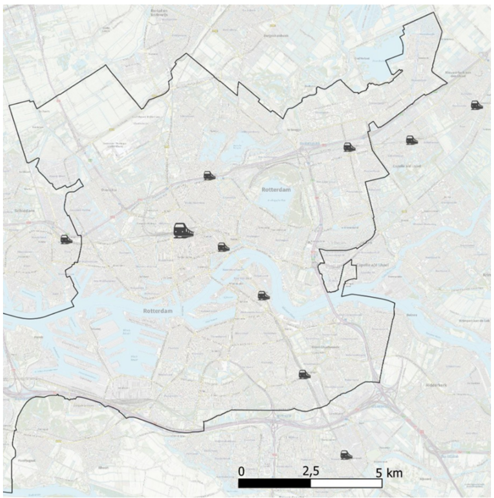



Figure 1.2: Stations in and nearby Rotterdam

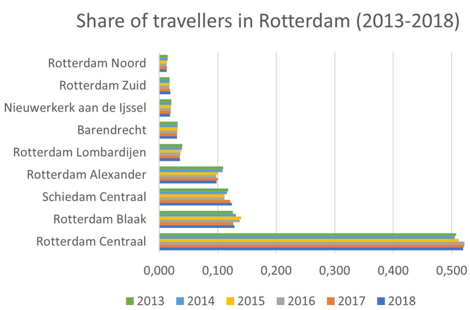



Figure 1.3: passengers Share of Stations Rotterdam based on published numbers by NS (Verkerk, 2019) checked with (NS, 2019)

Different studies (van Mil et al., 2020; Chakour and Eluru, 2014 and 2009) assign this large share by Rotterdam Central to characteristics like travel time, travel costs and the presence of transfer of the supplied public transport services. However, van Hagen et al. (2012) proved that people do not have a sense of time, only time perception. This is in line with Debrezion et al.'s findings (2009), showing that 47% of Dutch passengers do not choose the nearest station as their departure train station. For this reason, it would be useful to investigate in other, less quantifiable factors, influencing station choice.

To increase the possibilities in usage of the combination bicycle and public transport, effective customer-oriented changes have to be made. This requires more insight into the factors that influence the choices made in access and main transport modes at stations (Ton, Nijënstein & Shelat, 2019). For the reasons mentioned above, research is needed to gain more knowledge on how local railway stations can support the central stations reaching their maximum capacity, investigating the factors that influence station choice. When the factors of this mode combination are known in more detail, policy measures can be advised to serve this mode combination.

## 1.2 Research Gap

Various research is done concerning the mode choice of public transport and bicycle. This is discussed in more depth in chapter 3. However, less research is executed where bicycle and public transport are considered in one mode combination. In the niche of bicycle and train, station choice is underrepresented. Examining station choice could offer solutions in skimming passengers from overcrowded central stations to local railway stations. By doing so, these local railway stations could support the overcrowded central station regarding capacity problems. Van Mil (2020) found several factors regarding station choice and their interrelations. The factors included in his research were travel time, transit and parking costs. All dissatisfiers stem from the lower layers of the customer pyramid (figure 1.4; Hagen and Bron, 2014). Research regarding the satisfiers in the upper layers of the customer pyramid and access by bicycle including station choice, is currently limited in literature. This is unfortunate since most dissatisfiers require much effort and costs to be adapted. Another gap found in existing research is the usage of user classes. Overall, research was done taking an average traveller. The difference between several population groups was not included in station choice. It could be useful to know if there are any differences in weights of certain factors between personal characteristics. This allows the possible results to be more accurate compared to reality. The addition of different user classes and including the satisfiers (upper layers of figure 1.4) in station choice would make this research different from previous or existing research, and therefore add knowledge to the existing field. Including user characteristics allows to account for more heterogeneity in research. This is useful to enable tailored measures for the population in a specific area.

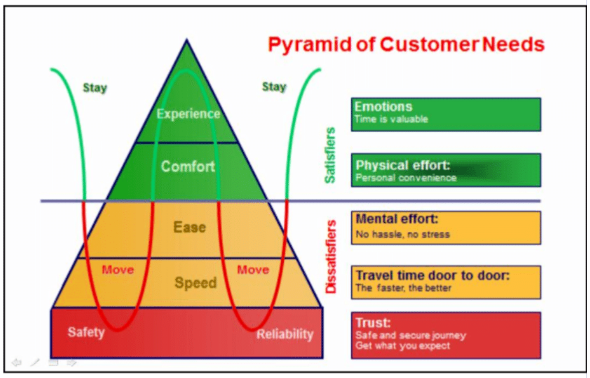



Figure 1.4: Pyramid of Customer Needs (Hagen and Bron, 2014)

Since there is more than one gap present, prioritisation is desired. The first research gap relates to insight into station choice for the bicycle-train mode combination. The second gap regarding satisfier characteristics is subordinate to the first gap, which means that the focus of this research lies on the satisfier characteristics. However, not all dissatisfier characteristics can be excluded, since passengers do not solely choose their station on satisfiers (Puello and Geurs, 2016). The third gap containing the different user classes is included with an equal priority as the first one, which is possible because they can be filled sequentially.

## 1.3 Research Objective

The overall objective of this research is to indicate the role of satisfiers in station choice. By researching in the choices made by passengers for station choice with bicycle as an access mode, more insight is gained. This insight could be used in future research to investigate the ability of local stations to assist with capacity issues of central stations. Passengers might have different travel motives or the station choice could strongly depend on personal characteristics. The outcome of this research could help any decision making on how to make local railway stations more attractive for bicycle-train mode combination users. The objective of this research could be split into two parts. One being the theoretical objective and the second being the practical objective.

### 1.3.1 Scientific Objective

As mentioned in chapter 3, previous research considering access and egress modes of public transport is present. However, within this subject, mode choice is elaborated more on than station choice (Givoni and Rietveld, 2007). This research contributes to the theoretical science trying to enclose the gap of knowledge regarding station choice. Where station choice was investigated by van Mil et al., this research focused on the dissatisfiers, meaning monetary and numerated values. This research complements the research of predecessors by adding the satisfier factors of perceived safety, experience and comfort feeling in the station choice.

### 1.3.2 Practical Objective

This research originates from the problem of crowdedness in city centres and on the existing networks. By examining the local railway stations as well, instead of only the central train stations and hubs, this research could help to solve the problems station owners, municipalities and provinces will be dealing with in the future or are already facing. By knowing the factors that influence station choice, it could be anticipated by stations that are currently less crowded. This social relevance accounts for the practical objective of this research proposal.

## 1.4 Research Questions

This research aims to fulfil the knowledge gap regarding the mode combination Bicycle and Public Transport by using station choice experimental data. This research allows to include the various weights of factors for several user classes and their characteristics. This leads to the following research question:

”To what extent can satisfiers, as opposed to dissatisfiers, play a role in station choice for bicycle-train travellers regarding local railway stations?”

Several sub-questions are required to answer this main research question. The answers to these sub-questions jointly answer the main question.

1. What are the main components of the bicycle-train trip chain regarding the choice for local railway stations and the corresponding appreciation of the traveller?
2. What are the factors related to perceived safety and comfort, influencing station choice for travellers using the mode combination bicycle - train, and their corresponding ranking?
3. Which relation could be determined between these factors, also including access travel time and parking costs, and to what extent do the weights of these factors differ between several user classes?
4. How could an application look like in reality using the weights of factors relating to station choice between user classes?

## 1.5 Scope

The primary scoping is that most monetary and time-related factors are excluded from this research. The focus of this thesis lies on the satisfiers represented in the upper part of the pyramid of customer needs (Hagen and Bron, 2014). Meaning the choice experiment doesn't include dissatisfiers other than access travel time and costs of bicycle parking. This also clarifies that the gap concerning local railway stations is prioritised above the gap concerning satisfiers, since passengers do not only make their decisions on satisfier characteristics.

This research consists of a choice experiment regarding station choice which is done in The

Netherlands due to its relative maturity of the bicycle-train mode combination (Kager et al., 2016). Therefore, data from this country is gathered. The research focus lies on the mode combination bicycle and public transport. In this research the bicycle is used as the access mode considering the availability of bicycles (Rietveld, 2000a), the egress mode is neglected. There is no distinction made between bicycle on board or bicycle and ride. This is not included since the access mode (bicycle) is the same in both options. The trip is considered bicycle-public transport once the access mode is a bicycle and the main mode is a train.

A different solution for capacity problems of central stations is to build a second central station where intercity trains stop. Although this could be an effective solution, this research only takes the current supply into account. As Lewis Mumford stated in 1955: "Adding highway lanes to deal with traffic congestion is like loosening the belt to cure obesity" (p. 92). This quote could also be implemented in the public transport network by replacing highway lanes with stations. In line with taking only current supply into account is also presence of transfers is not included in this research. van Mil et al. found transit as an important factor influencing station choice. However, including this result in a possible solution could mean changing the current supply. Due to this complexity, this dissatisfier is not included with travel time and costs in the choice experiment.

## 1.6 Thesis Outline

The first chapter of this thesis has introduced the problem and research questions. The outline is partly determined by the four subquestions. In chapter 2, the methodology used in this research will be presented and described for each subquestion. Chapter 3 includes the literature study, which provides more background information in the field of the mode combination bicycle - train. This literature study also provides the answer to the first subquestion. The second subquestion is investigated in chapter 4, including interviews. The third subquestion, where relations between factors are researched, is answered in chapter 5 describing the stated choice experiment, chapter 6 including descriptive statistics of the respondents, and chapter 7 concerning the final results. An illustrative case is included in chapter 8, and final conclusions are drawn and discussed in chapter 9.

# 2 Methodology

To answer the main research question including the four subquestions, different methods are used. In this chapter these methods are listed and explained. The main approach includes a stated choice experiment combined with latent class analyses. All methods are included in figure 2.1, which shows the methodological framework. In this framework, the different subquestions are connected with the various methods used in this research. The arrows between subquestions contain the knowledge gathered in a previous subquestion (1.4) to be used further in this research.

## 2.1 Methodological Framework

As visualised in figure 2.1, subquestion one regarding the main components of a bicycle-train trip chain and its appreciation is answered using literature research and interviews with travellers. The literature research provides the base when determining the main components. Using the interviews, these components are checked and a complete list of factors results.

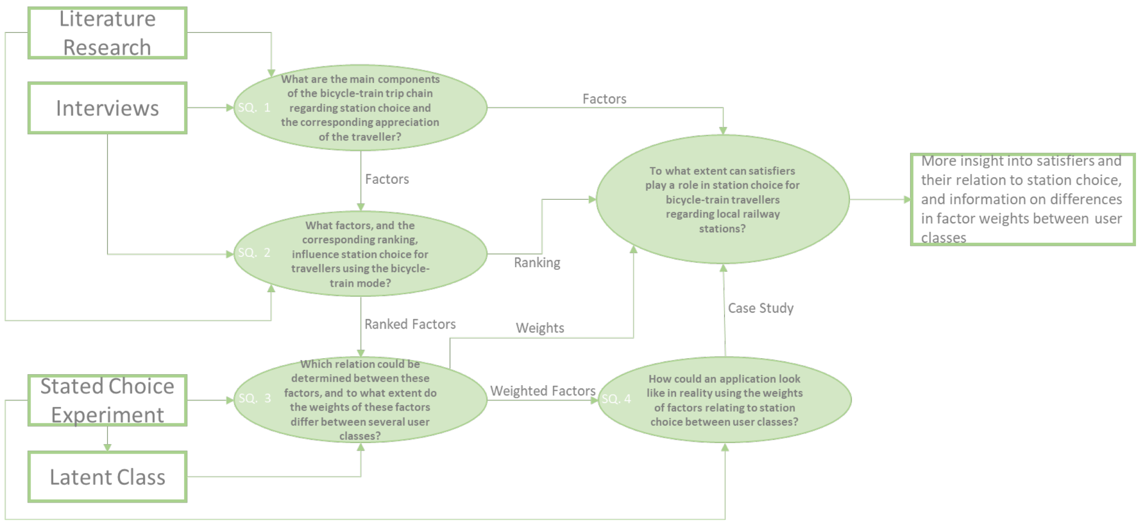



Figure 2.1: Methodology Framework

Subquestion 2 used the same methods as the first subquestion. However, different from the first subquestion, the main method used to retrieve the ranking of factors is interviewing. This is supported by using literature research. Interviewing is a crucial addition to identify all factors influencing station choice, even if these factors were not found in the existing literature. Other than verification, also the ranking of factors is performed using interviews. Respondents are asked to rank their most important factors related to station choice. A ranking list allows to make a stated choice experiment with the most important factors.

Various theories could be used when investigating choice modelling. In this thesis, the economic theory is used concerning utility maximisation. That is why, for the third subquestion, stated choice experiments are executed using Ngene. Data is analysed in SPSS and Matlab according to a Multinominal Logit Model and Latent Class Choice Model. This analysis is able to connect different weights to each factor for different user groups using the latent class technique.
To answer the fourth subquestion, the weights for different user classes are used and applied in a case study. These weights are combined with policy techniques that could have an effect on improving the factors influencing station choice for smaller stations for specific user classes.

## 2.2 Literature Research

For the first two subquestions, literature research is one of the methods used. This literature research is essential to define the different factors, to elaborate more on the bicycle - public transport combination and to investigate possible results of this research. Literature research is used to initially determine possible factors influencing station choice. In this literature research the data base of google scholar is used, including several theses and papers advised by supervisors and colleagues. The search terms included ”bicycle- train combination”, ”station choice”, ”stated choice experiments” and others. In addition, the theses of Van Mil (2015) and Krabbenborg (2015) are used, including the existing bibliography. After scanning these gathered papers, backwards and forwards snowballing technique was used.

## 2.3 Interviews

In order to decide which satisfier factors should be included in the stated choice experiment several interviews are done where interviewees were asked to rank their most important factors. Three citizens of each seven user classes predefined by Shelat et al., 2018, are interviewed to check whether or not their key factors are included. These user classes are covered in depth in section 3. In these interviews, all interviewees are asked to rank their 7 most important factors. This ranking of factors is used to determine the stated choice experiment set. It is considered that 3 interviewees per user class leading to 21 respondents might not be fully representative for the whole population. To solve this problem in significance, an obvious solution could be to retrieve more participants. However, to maintain the sequence in priority of the subquestions and to ensure the durability of approximately six months for this thesis, it is chosen to only consider this smaller amount of respondents. An important note is that this step is indicative and not meant to obtain a significant result. This step is solely taken to gain explorative insights into the ranking of factors that influence station choice and possible differences between user classes. By having at least 3 respondents per user class, a difference in characteristics of the respondents is included, concerning age, gender, education level and postal area.

## 2.4 Stated Preference Choice Experiment

After the factors are listed and ranked, five factors are included in the stated choice experiment. There are different theories that can be used when researching choice behaviour. In this research the economic utility maximisation theory is applied and combined with stated preference choice experiments.

### 2.4.1 Economic Theory

The majority of choice models in transport are based on the utility maximisation assumption within the micro-economic consumer theory (Hess et al., 2018). The principle of utility maximisation is used in discrete choice models, which means that a decisionmaker chooses the alternative which provides the highest utility (Hensher and Greene, 2003). For this reason, the Utility Maximisation Theory is used in this research, assuming that travellers act rationally and have well-defined preferences. This theory allows the participant to choose the option which provides the highest utility, according to a fixed utility function (Aleskerov et al., 2007). This utility function would determine the utility for each possible choice, being the different stations of public transport. An example of a utility function is given in formula 2.1.

$$U_s = \beta_a * A + \beta_b * B + ... + \beta_x * X \tag{2.1}$$

This should be interpreted as the utility of station s, which is equal to the sum of a factor's weight (Beta) multiplied by the corresponding units of that factor. In this study, the factors should be investigated first. This is done in subquestion 2. Afterwards, the corresponding beta's of these found factors should be determined. This is done using data from the stated choice experiment. There are mainly two types of experiments that can be used for this determination. Next to stated preference choice experiments, one could also use revealed preference choice experiments. Both types of experiments are discussed in the next section.

### 2.4.2 Revealed and Stated Preference Choice Experiments

When creating a choice experiment, stated preference and revealed preference form could be used. Stated preference studies find that the utility estimated using stated preference methods differs from the utility estimated using observed (revealed) behaviour (List and Gallet, 2001; Fifer et al., 2014; Murphy et al., 2004). Since stated preference surveys ask hypothetical questions they may not measure respondents choices accurately. This is because the discomfort of these hypothetical choices could be underestimated since they are not experienced in real life. This underestimation of discomfort would lead to hypothetical bias in the estimated utility. However, since this method uses hypothetical situations, the choice set could be under- or overestimating the weights.

Revealed preference studies use data from actual behaviour (Clinch and Kelly, 2004). Stated preference models have advantages over revealed preference models, requiring less data and avoiding multi-collinearity problems which often confound revealed preference analysis (Grisolía and Willis, 2016). However, a substantial disadvantage of comparing stated preference with revealed preference models, is that stated preference models are not estimating weights on real data. This means the choices are not made in real life and could be falsely estimated. Potential biases could therefore be included in stated choice data. Furthermore, due to the need to also create the choice set, stated preference method could be very costly and time-consuming. For the purpose of ease and ability to create a hypothetical choice set, stated preference is used in this research.

The respondents are asked to choose between a list of options with different attribute values. These respondents are gathered by spreading the survey using social media, colleagues and acquaintances. To ensure a representative group of respondents, the survey is also distributed using PanelClix. PanelClix is a Dutch online panel company which is specialised in sending out online questionnaires and gathering various respondents. Participants voluntarily applied to this panel and are rewarded a small fee once completing the questionnaire (PanelClix, 2021). In figure 2.2 an example of the choice set for the stated choice experiment is given. In this set, the factors are given with their three corresponding levels. Factor 1, 2 and 3 are the factors determined by the ranking.

**Station A / B**
**Access Travel Time: 5 / 10 / 15 min**
**Bicycle Parking Price: 0 / 2 / 5 euro's**
**Factor 1 : high / average / low**
**Factor 2: high / average / low**
**Factor 3: high / average / low**

Figure 2.2: Choice Set Example

## 2.5 Multinominal Logit and Latent Class Choice Model

The multinominal logit model (MNL) is the most used discrete choice model (Train, 2009). An MNL assumed independent and identical distribution of the random component $\epsilon_i$. To gather first insights in parameter weights, an MNL model is applied it is quite easy to understand and apply. The MNL model is applied twice, once on a pilot study, which included 38 respondents. This resulted in prior values for the weights that enabled the final choice experiment to be more efficient. Also several improvements and clarifications are applied on the pilot study to construct the final survey. For the final choice experiment 308 respondents are gathered. A known disadvantage of an MNL model is the assumption of fully rational behaviour, while in reality, individuals cope with limited knowledge and ability to know all consequences of each choice (Bhat et al., 2008). Also an MNL choice model does not include differences in the population and cannot account for their effects on parameter values.

Other models are a Nested Logit model (NL) or a Mixed MNL model (ML). An NL model can be constructed when alternatives in the choice set can be divided into subsets (nests) by certain criteria (Bai et al., 2017). An NL model is very suitable for mode choice. An ML model is an extended MNL model which compensates for random taste variation and panel effects (Shen, 2006). Compared to an MNL model an ML model is more complex to apply because it includes an additional error component (Train, 2009).

MNL choice models use population means. However in reality, people differ in this total population. Preferences and decision rules vary and thus one size, the mean, does not fit all. This could lead to misguided transport policy. Like the ML model, a Latent Class model can also capture panel data effects (Shen, 2006). In this research Latent Class choice models is used to include the differences in the population and to make transport policy more applicable for specific user groups. The latent classes are created around the heterogeneity in the parameters of the choice model. Latent Class Model assumes that a discrete number of classes is sufficient to account for preference heterogeneity (Shen, 2009; Boxall and Adamowicz, 2002). This number of classes is decided in a trade-off of latent classes and over-fitting. Following, personal characteristics can be added to the model to explain class membership (Molin and Maat, 2015). Latent Class is an elegant way to capture these types of heterogeneity that are neglected when using MNL models. Thus, according to this latent class choice model, panel data can be used to assign respondents to different classes. After the assignment, the weight for each factor is determined for all different user groups.

# 3

# Literature Research

The goal of this literature study is to provide background information to get insight into the bicycle-train mode combination, factors influencing station choice, and to introduce the different user classes of the mode combination bicycle and public transport. Literature research is useful to conclude what has already been researched and, more importantly, where there are still research gaps. During this literature research, the following search terms were used: 'bicycle', 'public transport', 'bicycle train combination', 'bicycle public transport combination', 'station choice' and 'user characteristics'. Initially, these terms already led to an extensive variety of papers. If papers were found useful, the snowballing effect was used, meaning the resources of these papers were also included in this literature research.

In this literature research, the following structure is applied. Three main gaps are discussed: 1) station choice including bicycle-train 2) the influence of satisfiers and dissatisfiers and 3) the user classes defined by Shelat et al. (2018). This literature research is completed with a conclusion and discussion where the research gap is discussed and used for section 1.2. A conceptual model (figure 3.6) is constructed to illustrate all important concepts of this literature research.

## 3.1 Bicycle-Train Mode Combination

In literature, most collected data concerns the main mode of travel on a regular basis, but tends to neglect access and egress modes (Rietveld, 2000b). In this thesis two mode choices are considered in one mode combination: the access mode choice (bicycle) and main mode choice (train). The mode choice for bicycle and public transport is influenced by several factors like travel motive, user characteristics, travel distance, car availability and climate and weather (Martens, 2004).

In bicycle usage, The Netherlands represent an extreme, as it has the highest level of bicycle use within the industrialised world. More than 27% of all trips are made by bicycle, a number that has been relatively stable over the last decades (Pucher and Dijkstra, 2000). In medium-sized cities this share is even exceeding 35% (De la Bruheze and Veraart, 1999).

This popular cycling mode is used even more when it comes to choosing an access mode to the train station. More than 50% of train travellers choose the bicycle to travel to their boarding station (ProRail, 2019). Which corresponds to Jonkeren et al. (2018) findings that 54% of train travellers choose for the bicycle-train combination. These percentages are proof that train station access by bike is very popular. An explanation for this high share is the combination of speed complemented by accessibility, as shown in figure 3.1 (Kager et al., 2016). Because of its popularity, researching the combinational mode would be beneficial. To assume a mode combination, would allow to include interrelations of factors that could be overlooked in existing research.

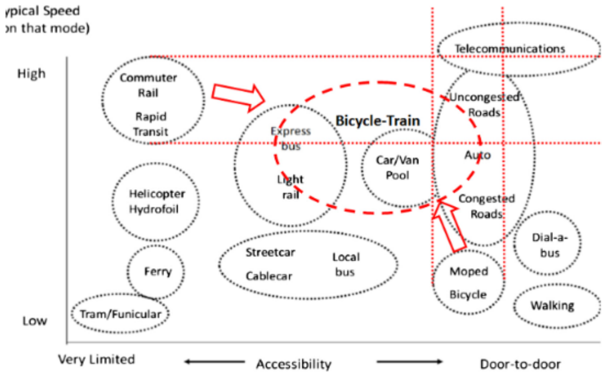



Figure 3.1: Speed and door-to-door accessibility of bicycle train intermodality (Kager et al., 2016)

## 3.2 Station Choice

Several factors concerning station choice by cyclists are listed by Krabbenborg et al. (2015), where different factors and their appreciation for cyclists to a station were investigated. Some of these factors are also used in this thesis with their corresponding reference. In her research, Krabbenborg et al. focused on the choice of the cyclists, not in combination with station choice or the remaining train trip. Weliwitiya and Eng (2020) added more factors related to the bicycle train intermodality. Station choice for cyclists was researched in 2020 by van Mil et al., where several factors were included in a stated choice experiment. Combining all the factors found in existing literature, four categories that are connected to station choice are created. These categories are 'Access Characteristics', 'Station Characteristics', 'Train trip characteristics' and 'General Factors'. Each factor in each category is paired with a utility, a positive association (+), or dis-utility, a negative association (-). All factors are thoroughly described in this section.

When listing the factors related to station choice with bicycle as an access mode the pyramid of customer needs (Hagen and Bron, 2014) is used to group the factors (figure 1.4). The upper layers of this pyramid represent the experience and comfort factors, grouping them as satisfiers. The lower layers of the pyramid include factors relating to ease, speed, safety and reliability. It is important to notice that the dissatisfier safety factor is not equal to the perceived safety of a traveller. This perceived safety is included in the satisfiers.

### 3.2.1 Access Characteristics

The trip to the train station by bike is called the access trip. There are many factors influencing this part of the station choice. The first factor is the **distance (-)** between origin and the station. An increase in the distance leads to an increase in the needed time and effort for this access trip. Due to the physical effort required, the relation between cycling distance and station choice is negative (Van Wee et al., 2006). In terms of access to railway stations, research showed the bicycle to be used most often for distances between 0.5 and 3.5 kilometres (Keijer and Rietveld, 2000; Martens, 2004). A factor strongly related to distance is **travel time (+)**. Travel time is not just influenced by travel distance, but also by **cycling infrastructure (+), accessibility to stations (+), crowdedness (+/-), right of way (+)** and **presence of obstacles (-)**. These factors are closely related because they all have an influence on travel time. Apart from their influence on travel time, most of these factors are also comfort related and important for the experience of the trip (Stinson, Bhat, et al., 2005). Implementing cycling infrastructure like boulevards may promote cycling as a station access mode (Weliwitiya and Eng, 2020; Cervero et al., 2013). Martens (2007) and Rietveld (2000b) also found that facilitating low-stress infrastructure is associated with positive numbers of access by bicycle. This same holds for the obstacles or parked cars on the bicycle lane (Garcia et al., 2015). Accessibility to stations combines the directness of the bicycle route to the station, the continuity of infrastructure and the right-of-way for cyclists (Scheltema, 2012). When a train station has higher accessibility, the route is more attractive for cyclists (Cui et al., 2014; Rietveld, 2000b). Mentioning right of way, it seems straightforward that cyclists prefer a continuous flow (Stinson, Bhat, et al., 2005) and the presence of stops leads to a lower utility for the cyclists on that route (Rietveld and Daniel, 2004).

Other comfort-related factors would be **human scale (+), topography (+/-)** and **vegetation (+)**. In 2014, Park et al. investigated the influence of the size of buildings on the walkability of that route. He concluded that streets with tall buildings or narrow streets have a negative impact on walkability. In 2015, Krabbenborg et al. extended this walkability to bike-ability and came to the same conclusion for cyclists. This means human scale of buildings on the route has a negative influence on cycling resistance. The factor **topography (+/-)** includes the possible influence on cycling resistance due to mountains and hills on the route (Semler and Hale, 2010; Weliwitiya and Eng, 2020). In The Netherlands, with its flat surface, mountains and hills will not be of a large influence. However, topography like possible bridges or railroad crossings could have a negative influence on the cycling volumes on that route. The last comfort variable in the access trip found in the literature was vegetation. According to Krenn et al. (2014) cyclists prefer green areas. Studies for New Zealand and The Netherlands show, a diverse land use mix relates to higher bicycle rates (Puello and Geurs, 2016; Mackenbach et al., 2016).

Factors that are more safety related are **vehicle volumes (+/-), speed limit (-), social safety (+), traffic safety(+)** and **street lighting (+)**. Safety could be divided into traffic safety and perceived safety. Where perceived safety is more related to social safety and traffic safety relates to the prevention of crashes or accidents. Where traffic safety is a basic need for travellers according to the pyramid of Van Hagen (figure 3.4) and perceived safety is marked in the upper layers of this pyramid. Rietveld and Daniel even stated that a lack of safety is a reason not to cycle. Due to the vulnerability of cyclists against motorised road users, traffic volumes and crowdedness are associated with lower utility for bicycle access to a station (Cui et al., 2014). However, crowdedness of other (non-motorised) road users is proven to have a positive relation to the social safety (Krabbenborg et al., 2015). This crowdedness for non-motorised road users has an optimum for the perceived safety and level of comfort. Puello and Geurs (2015) found lower cycle rates in high-density urban areas. This would mean a negative effect on station areas since these are often located in a highly dense area. Another factor related to social safety is the presence of street lighting. According to Puello and Geurs (2015) low quality of street lighting in the surrounding of railway stations has a negative influence on bicycle use as an access mode.
Related to traffic safety, the factors cycling infrastructure, right of way and presence of obstacles explained previously are applicable. Next to these factors, the speed limit applicable on motorised traffic has a negative relation with the cycling rates (Park et al., 2014). In this case study Park et al. found speed limits larger than 50 km/h results in lower cyclists.

When the destination is chosen, different routes could be used to end up at that same destination. Therefore, related to access characteristics, **route preference (+/-)** could influence station choice. Route choice for cyclists is researched in many different papers (Rietveld, 2000a; Broach et al., 2012; Liu et al., 2020). The importance of route preference is illustrated in 2012, van Hagen et al., 2012 showed that people do not have a sentence for time, only for time perception. Stating that experienced travel time should be included next to travel time. This is applicable to route choice since factors other than time influence the travel experience.

### 3.2.2 Station Characteristics

Next to access trip, another component of the trip train concerning the mode combination bicycle-train is station characteristics. Station characteristics include all aspects of the station itself that could influence station choice. For this section, it is important to define the difference between central and local railway stations used in this research. Where central stations indicate the stations that are currently dealing with capacity issues or are unable to expand due to the urbanised area, local railway stations are the stations that are coping with fewer passengers on a daily basis than the central stations nearby. This distinction between the size of the station is related to the **size of node (+)** of that station. In this factor, it is meant how many other public transport types are present at the station, other than train. There are two main train types, intercity and local trains. Where an intercity has less stops and thus a lower travel time from station to station. And a local train has more stops which could lead to a higher travel time between stations but a lower access time to the boarding station. Not all stations allow intercity trains to stop. Therefore, a preference for **train type (+/-)** is related to station choice. The centralised stations currently dealing with capacity problems are mainly intercity stations.

The factor most stated associated with the bicycle-train integration is the ability of **bicycle parking (+)** (Keijer and Rietveld, 2000; Givoni and Rietveld, 2007; Cervero et al., 2013). Where both the actual safety and perceived safety of parked bicycles can influence the choice to ride to that station (Sherwin et al., 2011). This safety is increased when **security (+)** is available at the bicycle parking location (Replogle, 1987; Givoni and Rietveld, 2007). The perceived safety increases when the bicycle parking offers a clear **overview (+)** of its area. This involves no hidden corners or corridors. Besides the positive input of secured bicycle parking, a corresponding **parking price (-)** has a negative impact on the utility of choosing that station (van Mil et al., 2020). A factor **transfer time (-)** is defined as the time spent between bicycle parking and arriving at the platform. Another factor related to station characteristics is the **station environment (+)**. This includes for example **vegetation (+)** (Krenn et al., 2014 and the **neighbourhood reputation (+)** (Ferrell and Mathur, 2012). For the neighbourhood reputation, this station environment could also influence the perceived and actual safety of bicycle parking.

Regarding station characteristics, **level of maintenance (+)** and **tidiness (+)** (Birago et al., 2017) of the station could also influence (perceived) safety and comfort. Also **level of lighting (+)** at the station is strongly related to (perceived) safety and comfort (Vos et al., 2015). Apart from tidiness, services provided at the station could increase the comfort level perceived in that station (Puello and Geurs, 2016). These **station services (+)** enhance the presence of toilets, shopping and eating facilities and a possible information desk at the station. For the travellers wanting to take their bicycle on board, facilities such as escalators or lifts provided at the station increase the perceived comfort of that station (Cervero et al., 2013). Other comfort-related station characteristics are **level of shelter (+)** and **seating (+)**. This level of shelter is provided by the waiting cubicle located at the platform. For central stations, also the station itself could provide shelter. Regarding seating there is seating at a platform and seating in the station itself. For local railway stations this could be indifferent, but central stations might have more shelter offered in the station, whereas at the platform seating, the travellers are able to see the train when it arrives.

The final factors concerning station characteristics are described as **spaciousness (+)** and **crowdedness (-)** (van Mil et al., 2020). These factors are also interrelated because the more spacious the area is, the less crowded it will be with a certain amount of travellers. Crowdedness at the station is involved in several parts of the station. When the bicycle parking is crowded, it could decrease the comfort level of that station since more time might be needed to find an empty parking space. Crowdedness at the platform decreases safety when it reaches a certain level. However, it also increases the perceived safety compared to an empty station with no or very few fellow travellers. This trade-off is comparable as described in the access trip.

### 3.2.3 Train Trip

When choosing a station, one could choose their station on access preferences but also on the train trip he/she would like to make. For example, van Mil et al. concluded in (2020) that many travellers are willing to cycle a few minutes longer to prevent a transfer. This factor is **directness (+)**. Most travellers prefer a train trip without transfer (Zhao and Ubaka, 2004). The majority of factors related to the train trip are 'dissatisfiers'. Other than directness, the **total travel time (-)** is important. The total travel time is closely related to the **train type (-/+)** stopping at the station. An intercity has an average higher speed than a local train. Compared to a local train, the intercity has fewer stops on the same route which increases the comfort level. Concluding, travel time is closely related to the train type and the train type influences the comfort level of the train trip.
Another factor related to the train trip is **reliability (+)**. This reliability enhances the probability that the train will perform as expected. Next to travel time, reliability also influences the experience of the traveller (van Mil et al., 2020). Closely linked with reliability is **frequency (+)** (Bowman and Turnquist, 1981). A high frequency will result in a lower waiting time and would increase the travel experience.

In van Mil et al.'s research, some of these factors were investigated in regarding their corresponding weights. In this study he found the following utility function and weights:

$$ALT_i = -0.19*Biketime_i + -0.13*TimetoPark_i + -0.14*traintime_i + -1.77*Price_i + -1.06*Transfer$$

Where $ALT_i$ is the total utility of the station in alternative i. $Biketime_i$ enhances the access travel time to station i by bike. The $TimetoPark_i$ includes the time needed between parking the bicycle and arriving at the platform. $Traintime_i$ describes the total time spent in train(s) during the total trip. $Price_i$ is the total monetary price paid for the trip (out of pocket costs). Finally, transfer is a penalty of 1 if a transfer is present, and 0 when the trip only includes a direct train. The units for the factors are minutes (for time), euros (for price) and number of transfers (for transfer). This means, for example, that per one minute of bicycle time the utility decreases with 0.192.

### 3.2.4 General Factors

The last category of factors influencing station choice is general factors. These general factors include several trip and user characteristics.

With trip characteristics, factors such as **weather circumstances (+/-)** and trip purpose are meant. Also, the frequency of the trip could have an influence on station choice. In The Netherlands, the shares of bike and public transport users have different values throughout various travel motives. Work and educational motives are represented the most. Hereafter comes shopping, business and others (table 3.1). This leads to the hypothesis that **trip type (+/-)** could also be important for station choice.

Slightly related to trip type is the factor **trip frequency (+/-)**. This trip frequency could influence the

| Travel motive | Train | Bus | Metro |
|---|---|---|---|
| Work (%) | 40 | 21 | 33 |
| Education (%) | 30 | 51 | 22 |
| Shopping (%) | 6 | 10 | 19 |
| Business(%) | 3 | 1 | 4 |
| Other (%) | 21 | 18 | 22 |

Table 3.1: Travel motives of bike and public transport users (van Goeverden and Egeter, 1993)

station choice in terms of the more a trip is made, the total disutility of travelling that trip is often more important for the traveller than the disutility of a singular trip.
Another general factor is **time of day (+/-)**. This includes at which time during the day the trip is scheduled for the traveller. For example at 2 PM in the morning or 1 AM at night. Furthermore, trips are often classified between peak and off-peak period trips (Bovy et al., 2006).In this research regarding the combinational mode bicycle - train, period choice is very intertwined with station choice. This could be due to the safety feeling at night but also to the opening times of stations. For example, only stations where intercity trains stop connected in the night network are in operation between 1 and 5 AM. Most other stations are not open for travellers during this time. The entrances are not always closed, but since no trains are stopping, no travellers are present, and there is no use in visiting the station for travelling (Middelkoop, 2017).

User characteristics enhance the personal characteristics of the traveller. Studies include socioeconomic attributes relating to **age (+/-), gender (+/-), income (+)** and **car ownership (-)** (Chakour and Eluru, 2014; Fox, 2005; Fan et al., 1993 ;Debrezion et al., 2009). Age is related to mobility and physical effort, where a less mobile person is less likely to choose a further station than the nearest station. Also, value of time might differ per age category. Even though no gender split is found in access to railway stations, it is possible that gender distribution matters in station choice (Krabbenborg et al., 2015). Also, in 2004, Krygsman did find that men accept a longer access, main and egress travel time than women do.
In 1997 Adcock concluded that the willingness of a passenger to use a station other than their nearest increases as the number of cars per household increases. This suggests that car ownership may

influence station choice (Young and Blainey, 2018).

People choose to travel when the utility of the destination minus the negative utility of travelling is larger than the utility of not going (Sweet, 1997). The utility of a station is the outcome of the utility function of that corresponding trip. To change a boarding train station is linked with behavioural theory. This is enhanced in the NOA-model shown in figure 3.2 (Vlek, 2000). This model shows changing behaviour requires modifications to relevant needs, opportunities and/or abilities, but also an understanding of the interaction between the needs and opportunities and the interaction between opportunities and abilities. By stating the factors influencing station choice, the 'Needs' of travellers are indicated. The results of this research are possible measures that increase a factor. When a measure is implemented it could be categorised under 'Abilities'. Opportunities are not included in this research, but should be investigated before implementing any measures to forecast changes in choice behaviour.

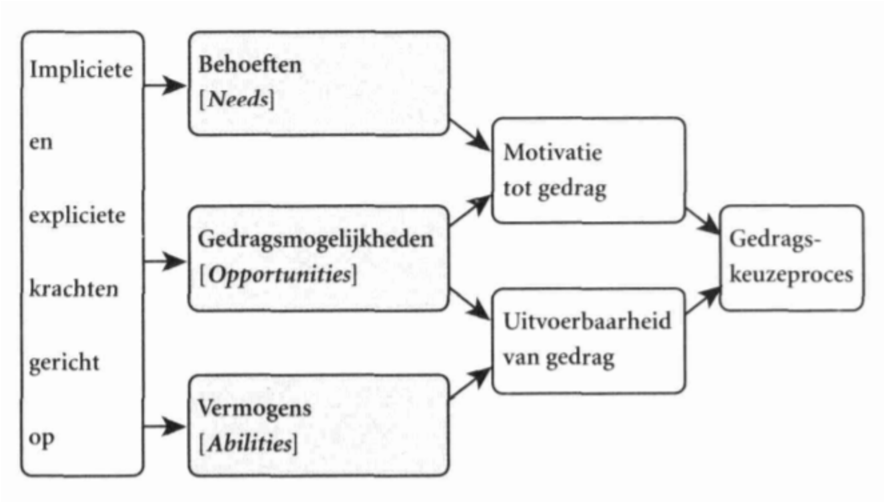



Figure 3.2: NOA-Model (Vlek, 2000)

## 3.3 User Classes

Shelat et al. (2018) found 7 different user classes of travellers using the mode combination bicycle and public transport. These classes and their corresponding labels and percentages are shown in figure 3.3.

Profile of the 7 cluster model.

| Cluster number | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| **Cluster size** | 26.50% | 22.84% | 15.10% | 14.06% | 9.64% | 9.47% | 2.39% |
| **Label** | Middle-aged full-time professionals | University students living with parents | School children | Young, low income professionals | Middle-aged part-time professionals | University students living alone | Pensioners |

Figure 3.3: User classes of train - bicycle (Shelat et al., 2018)

These seven classes are made using latent class cluster analyses. The first and biggest class is labelled as Middle-aged, full-time professionals. 74% of this group is middle-aged, meaning between 35 and 64 years old. This group mainly consists of working men (70%) and has a high overall car availability. University students living with their parents are the majority in the second group (22.8%). This group has no or limited access to a car and are daily transit users. The departure of trips is spread widely throughout the day. Most trips of this group are between a semi-rural to a strongly urbanised region. The third group is represented by school children (15.1%). This group has no car availability and the participants live with their parents. The income of the household in this group is on the lower side and this group is mainly located in rural areas. Young professionals with a low income make 14% of the travellers using bicycle-public transport combination.

The young professionals (14%) are in a younger age category than the full time professionals. Compared to the full-time, the young professionals earn less and have an equal distribution of gender and 80% never has access to a car. Next to the middle-aged full-time professionals, another group is defined as middle-aged part-time professionals. This group is responsible for 9,6% of all bicycle-train travellers. This group is very similar to the largest group. However, 85% is female and the car availability is much lower. 75% does not have regular access to a car. University students living alone (meaning not living with their parents) account for 9.5%. A large majority uses this mode for leisure purposes. 90% of the university students not living with their parents never has access to a car. The last user class defined by Shelat et al., 2018 is labelled as pensioners and makes for 2.4% of the total bicycle-public transport users. 40% of the pensioners have access to a car. The vast majority of this group are not regular transit users. Once every week or month a pensioner in this group uses this particular mode combination.

## 3.4 Literature Research Results

### 3.4.1 Discussion literature research

Jonkeren et al. (2021) addressed the needed research for the popular mode combination bicycle-train. A lot of research exists in the choice modelling of whether to make a trip by transit, bike or car (Müller et al., 2008; Hasnine and Habib, 2018). More research also accounts for the combination of multiple modes, such as bicycle and transit (van Oort et al., 2017; van Mil et al., 2020). However, most of this research focuses on mode choice instead of station choice (Givoni and Rietveld, 2007).

Research regarding station choice was performed in 2020 by Weliwitiya and Eng, and in 2015 by Krabbenborg et al. Regarding the mode combination bicycle-train van Mil et al. found several factors that might influence the station choice (2020). This is one of the few types of research that exists in this context. However, this research could be expanded on several ways.

First, it could add value to include both central and local railway stations. It is concluded that the main focus in existing research regarding mode and station choice is on the central train stations. Even more specific, the bicycle parking at these central stations. Local railway stations are hardly researched in-depth in this field of interest. Therefore, an application where these local railway stations are included could be insightful to research if passengers could be skimmed from central to local railway stations. Secondly, the research could include more factors related to perceived safety and comfort. In most research, only the lower layers of the pyramid of customer needs (figure 3.4) are included. It would add value to the current state of art to include more factors besides the dissatisfier factors used by van Mil et al. The third way to expand current research is to include the type of travellers using public transport in research related to station choice. Personal characteristics could lead to a difference in weights of factors between user groups. Combining the research of Ton et al. (2020) with station choice could contribute to more efficient measures for inhabitants in a certain area.

### 3.4.2 List of Factors and Conceptual Model

Using existing literature, several factors are identified that are related to station choice. An overview of these factors is presented in table 3.2. This list of factors provides the answer to the first research question: ”What are the main components of the bicycle-train trip chain regarding station choice and the corresponding appreciation of the traveller?”. For each factor, it is shown what type of relation it has to the probability of choosing that station. For example, distance has a negative relation, meaning the longer the distance, the less probability for the respondent to choose the station. A positive relation is found between toilets and station choice, meaning when toilets are provided at the station, the probability of choosing that station is increased. In table 3.2 also the type of factor is stated. A factor is labelled as a dissatisfier when it relates to the lower layers of the pyramid of customer needs (Hagen and Bron, 2014), and labelled as a satisfier when it relates to the upper layers of the pyramid.

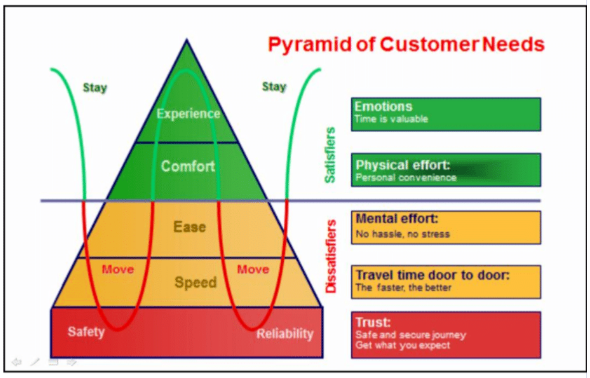



Figure 3.4: Pyramid of Customer Needs (Hagen and Bron, 2014)

For some factors there is an overlap. For example route choice, this means it relates to both satisfier and dissatisfier layers of the pyramid. These factors are labelled "both". Most general factors are not labelled in satisfiers and dissatisfiers. The personal characteristics are used in a further stage of this research to find the user groups in the bicycle-train mode combination.

To illustrate the factors and concepts influencing station choice, a conceptual model is created. This model is based on the factors noted in table 3.2 and the literature research. Besides stating the factors in the table, this conceptual model also provides the relations between those factors and concepts. The conceptual model begins with the wish for transportation by travellers, framed in grey. This introduces the eventual station choice, which is the main subject and eventually the result of the model, framed in black. Station choice is influenced by many factors. Four concepts are enlightened in this figure: comfort and dissatisfiers, safety factors and the capacity problem. An overall overview of the conceptual model is shown in figure 3.5. A detailed version is provided in figure 3.6.

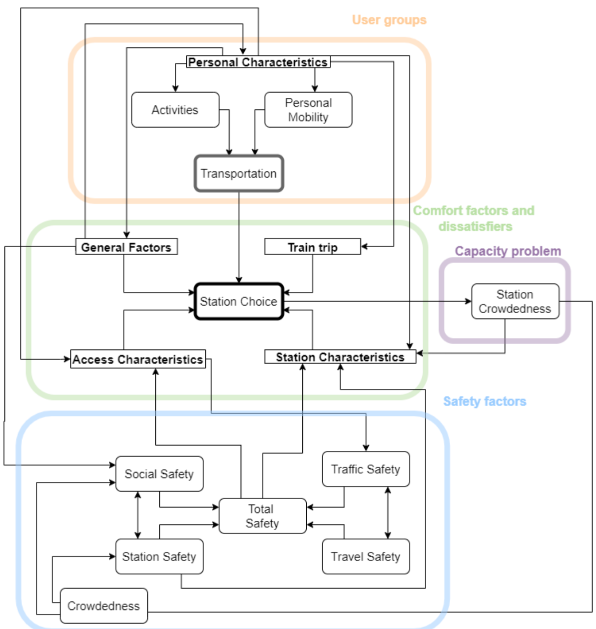



Figure 3.5: Overview and Structure Conceptual Model

Within the green frame, the comfort factors and dissatisfiers and their relation to station choice are shown. Station choice is influenced by four categories: access characteristics, station characteristics, train trip and general factors. All the factors within these four categories are described in the literature research. Some adaptations are made, such as expanding station services more specifically in presence of toilets, shopping and eating facilities. In the four boxes leading to station choice only the factors related to comfort and dissatisfiers are mentioned explicitly. Safety factors are separated in this conceptual model and are elaborated on in the blue alignment.

To clearly show the difference between dissatisfiers, comfort and safety factors, the safety factors are separated and shown in blue. The safety factors are separated from the comfort and dissatisfiers to show the extra complexity and various types of safety. One might argue that safety belongs in the lower layers of the customer pyramid, but this is only applicable for travel and traffic safety. In contrast, social safety is labelled as a satisfier in this thesis because it influences the travel experience and comfort level of the traveller. The total safety is based on four safety concepts. Traffic, travel, station and social safety. Travel safety is a concept collecting the actual safety of travelling, mostly influenced by surrounding traffic and measures as the speed limit for motorised vehicles. The concept of travel safety is more specific for the traveller itself, considering the presence of a bell and lighting on the bicycle and the experience of participating in traffic. Other contributions to the total safety are station and social safety. These concepts are elaborated on in the literature research and include the lighting, level of overview and the neighbourhood reputation. The green and blue boxes are connected by several arrows, including one between the time of day and social safety, since the time of day could influence perceived safety and the total safety to the safety within access and station characteristics.

Next to the connection between total safety and access and station characteristics, the blue and green boxes are also connected by the capacity problem, shown in the purple frame. The concept leading to the capacity problem is the station choice. The more travellers choose for the same station, the more overcrowded it can get, hence the station crowdedness related to the capacity problem. This station crowdedness is closely related to the station safety and social safety. But it also shows the feedback loop between station characteristics and station crowdedness via the bicycle parking. This simply means, the more crowded a station is for example when parking a bicycle, the less comfort is experienced, which could mean less travellers choosing that station option leading to a less crowded station.

The orange frame includes the user groups. These user groups are based on personal characteristics such as age, income social participation et cetera. These personal characteristics influence personal mobility. This mobility means to what extent the person is able to travel. This personal mobility together with the presence of activities (influenced by personal characteristics) leads to the wish for transportation. Which in its turn leads to a station choice. Personal characteristics also influence the four concepts leading to station choice. Because the weight of several factors within these concepts is dependent on personal characteristics. For example, gender influences the perceived safety on the access route, age could influence the count of comfort factors such as seating and toilets and income influences the value of time and the weight of price for the train trip or bicycle parking. And the concept general factors even includes personal characteristics.

| Factor | Relation | Type |
|---|---|---|
| **Access trip** | | |
| distance | - | dissatisfier |
| travel time | - | dissatisfier |
| cycling infrastructure | + | dissatisfier |
| accessibility to station | + | dissatisfier |
| right of way | + | satisfier |
| presence of obstacles | - | satisfier |
| human scale | + | satisfier |
| topography | +/- | satisfier |
| vegetation | + | satisfier |
| vehicle volumes | +/- | dissatisfier |
| crowdedness | +/- | both |
| speed limit | - | dissatisfier |
| social safety | + | satisfier |
| traffic safety | + | dissatisfier |
| street lighting | + | satisfier |
| route preference | +/- | both |
| **Station Characteristics** | | |
| size of node | + | both |
| train types | + | both |
| bicycle parking (presence & capacity) | + | satisfier |
| security | + | both |
| overview | + | satisfier |
| parking price | - | dissatisfier |
| transfer time | - | dissatisfier |
| station environment | + | satisfier |
| vegetation | + | satisfier |
| neighbourhood reputation | + | satisfier |
| level of maintenance | + | satisfier |
| tidiness | + | satisfier |
| level of lighting | + | satisfier |
| station services | + | satisfier |
| toilets | + | satisfier |
| shopping facilities | + | satisfier |
| eating facilities | + | satisfier |
| level of shelter | + | satisfier |
| seating at platform | + | satisfier |
| seating at station | + | satisfier |
| spaciousness | + | satisfier |
| crowdedness | - | satisfier |
| **Train trip** | | |
| directness | + | dissatisfier |
| travel time | - | dissatisfier |
| train type | +/- | both |
| reliability | + | dissatisfier |
| frequency | + | dissatisfier |
| price | - | dissatisfier |
| **General factors** | | |
| age | +/- | |
| gender | +/- | |
| income | +/- | |
| car ownership | +/- | |
| time of day | +/- | |
| trip type | +/- | |
| trip frequency | +/- | |
| weather | +/- | satisfier |

Table 3.2: Factors of Station Choice related to Pyramid of Customer Needs (Hagen and Bron, 2014)

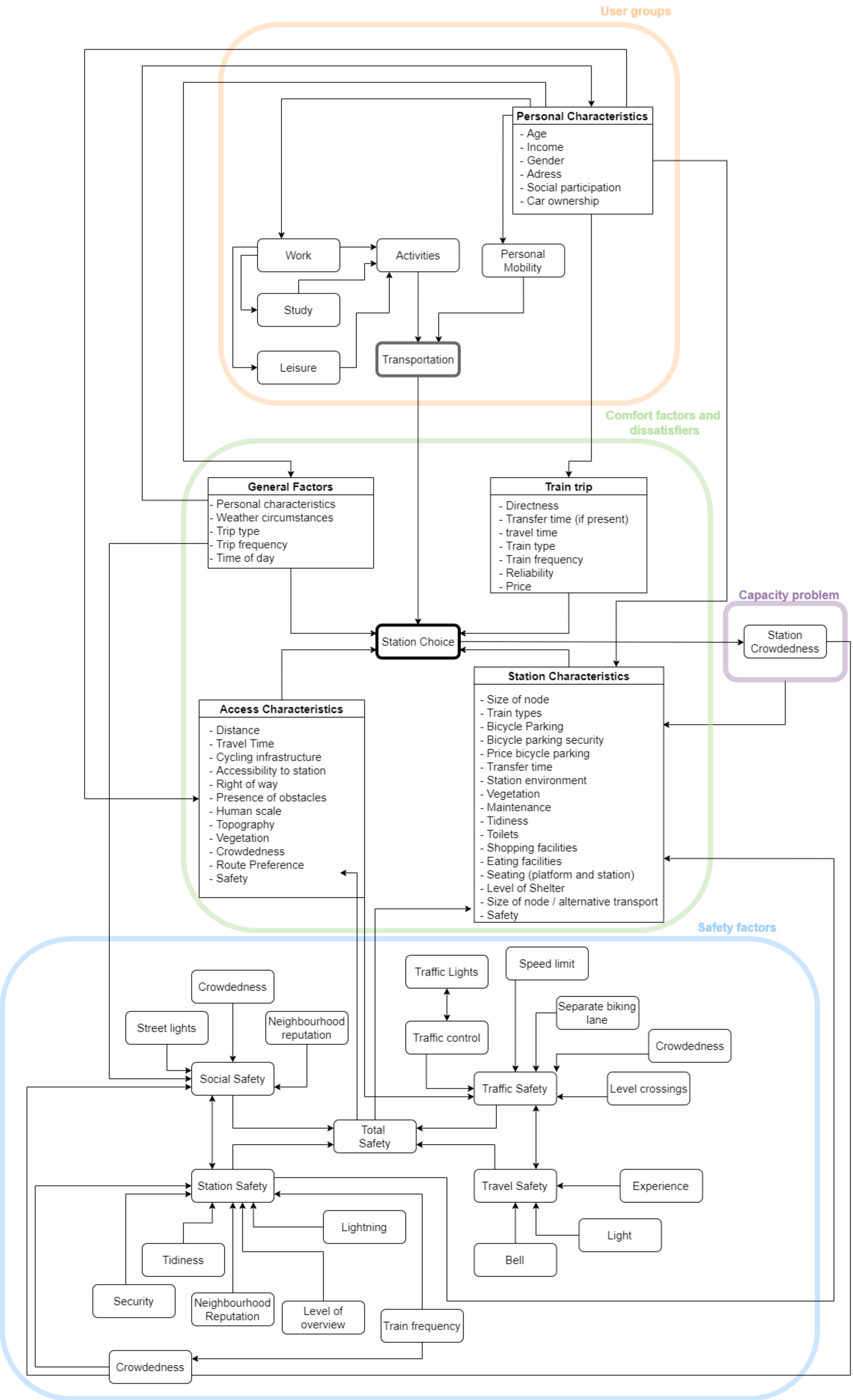



Figure 3.6: Conceptual Model

# 4 Ranking of Satisfier Factors

In section 3.2 all factors relating to station choice found in literature are listed. As stated in the research gap in chapter 1 of this project, an added value of this research in the current field is the larger focus on satisfier factors in the station choice. To determine which satisfiers should be included in the stated choice experiment first all possible factors resulting from the literature research (table 3.2) are listed. This table includes factors from the comfort and safety (green and blue) area's of the conceptual model (figure 3.6).

## 4.1 Included Satisfiers

This step of the research includes the ranking of the satisfiers. Dissatisfiers are excluded from the ranking to allow a focus on the satisfiers, since in existing research satisfiers often do not end up in the top of the ranking and are overshadowed by dissatisfiers.
Overall all satisfiers factors that are easy to adapt are included. For example, topography is left out since no mountain will be moved to influence station choice. Station services itself is also left out since its components (toilets, eating, shopping and seating facilities) are included specifically. This also holds for station environment which is explained by station services, vegetation and neighbourhood reputation, and social safety which includes street lighting, neighboorhoud reputation and other factors. Route preference is excluded from this list since it is influenced by many other factors making it a complex factor.

Components related to train trip were mostly dissatisfiers. Also train type is not included since this could confuse respondents and possibly correlates with other factors that correspond with intercity's or local trains. For example, a respondent could choose for the intercity option by incorrectly thinking this is a faster travel option. Personal characteristics are also not included in this ranking step. The personal factors are used later on to create different user groups when analysing the results of the final survey.

The final list of factors is shown in table 4.1. Due to the selection of satisfiers the categories are changed into access mode, bicycle parking and station. The factor corresponding definition is included in the table as well as the satisfier type. Using the factors from this list, the ranking of these factors can be executed by respondents. Retrieving these respondents all user classes of Shelat et al. (2018) are used and included. The ranking is explained in the next section.

## 4.2 Use of 7 User Classes

In section 3.3, research of Shelat et al. (2018) is introduced concerning the user classes amongst public transport travellers. In this research, nine indicator variables divided the total population of bicycle-train users into seven different user groups. These nine indicator variables are household composition, degree of urbanisation, household disposable income, individual gender, age group, highest education, social participation, car availability and transit use frequency. Using these indicator variables, Shelat et al. distinguished seven user groups of public transport as shown in section 3.3.

| Category | Factor | Definition | Satisfier type |
|---|---|---|---|
| Access | Intensity of lighting | Level of lighting during access trip during bicycle trip | perceived safety |
| | Warmth of lighting | Colour of lighting during access trip during bicycle tip | comfort |
| | Crowdedness | Amount of passengers by bike or walking during bicycle trip | both |
| | Neighbourhood Reputation | Neighbourhood reputation of access route | perceived safety |
| | Vegetation | Amount of trees, bushes, flowers and grass alongside route | comfort |
| | Human scale | Scale of buildings along access route | perceived safety |
| | Presence of other modes alongside access road | Amount of other public transport modes at access route | both |
| | Smoothness of bicycle lane | Presence of bumps or hitches | comfort |
| | Obstacles on bicycle lane | Presence of objects | comfort |
| | Right of way | priority of cyclists in traffic | comfort |
| Bicycle Parking | Intensity of lighting | Level of lighting in bicycle parking | perceived safety |
| | Warmth of lighting | Colour of lighting in bicycle parking | comfort |
| | Crowdedness | Amount of people in bicycle parking | both |
| | Overview | Level of overview in bicycle parking | perceived safety |
| | Bicycle Parking Security | Presence of security in bicycle parking | both |
| Station | Intensity of lighting | Level of lighting at station | perceived safety |
| | Warmth of lighting | Colour of lighting at station | comfort |
| | Crowdedness | Amount of other travellers at station | both |
| | Neighbourhood reputation | Reputation of neighboorhoud at station | perceived safety |
| | Toilets | Presence of toilets at station | comfort |
| | Seating area at platform | Presence of seats at platform | comfort |
| | Seating area in station | Presence of seats in station | comfort |
| | Shopping facilities | Presence of shops at or nearby station | comfort |
| | Eating and drinking facilities | Presence of eating and drinking facilities at or nearby station | comfort |
| | Size of node | Amount of other transport modes at station | both |
| | Vegetation | Amount of trees, bushes, flowers and grass in station environment | comfort |
| | Shelter | Level of shelter at station | comfort |
| | Camera Security | Presence of camera security in station | perceived safety |
| | Tidiness of station | Presence of dirt or smell at station | both |
| | Maintenance of station | A maintained station regarding construction | comfort |

Table 4.1: Satisfiers listed before ranking

The total profiles of this seven cluster model is included in appendix A. The resulting seven user groups and their shares of the total population of bicycle-train users are shown in table 4.2.

| | |
|---|---|
| Middle-aged full-time professionals | 26.5% |
| University students living with parents | 22.8% |
| School Children | 15.1% |
| Young low-income professionals | 14.0% |
| Middle-aged part-time professionals | 9.6% |
| University students living alone | 9.5% |
| Pensioners | 2.4% |

Table 4.2: User groups and their shares (Shelat et al., 2018)

The percentages shown in table 4.2 are not automatically representative for the total population of the Netherlands. In 2018, Jonkeren et al. concluded that the bicycle-train user is more often a younger, studying or working person. In this same research he also stated that women and men are equally represented in this group. In this research, the total population is defined as all bicycle-train users.

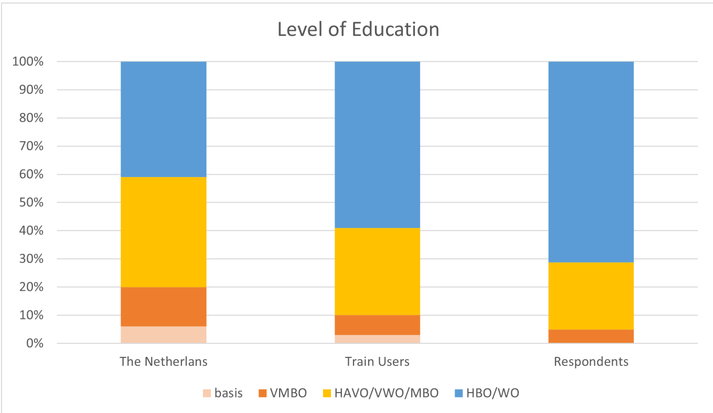



Figure 4.1: Level of education

To rank the factors retrieved in the literature review. A small interview study is performed. In this study 3 persons per user class were gathered and asked to rank their 7 most important factors from 1 (less important) to 7 (most important) out of the total list of factors (table 4.1). A clear distinction between user groups is age and social participation, included in the labels of the user groups. Next to these two characteristics, also people were selected to vary in household, income, gender, education level and level of urbanisation. It is important to mention that this ranking procedure might not lead to a significant result representative for the whole population. This is because of the small test population (N=21). To significantly substantiate the ranking for the whole population, a larger group should be interviewed. However, due to time planning and to be able to obtain the main goal of this research, it was chosen to use this ranking technique in a small population. This means the ranking is gathered in an explorative manner and is used to obtain a direction to gain knowledge about the ranking of factors.

All 21 respondents were asked to speak their decisions and arguments out loud so possible uncertainties could be solved during the interview. The travelling circumstances given to the interviewees were commuter travellers. This means to keep into mind average possible commuter circumstances. There is a possibility that it is dark, and rains. But other times it could be dry weather and daytime. If respondents would be asked to keep a rainy day in mind while ranking the factors, this could lead to an over prioritisation of shelter compared to all other factors. The reason for not giving a clear travel circumstance as raining and night-time, is keep the factors as neutral as possible.

After the ranking, two parameters are important for the factors. These parameters are the total sum and the count. The total sum is calculated by adding all points per respondent together. For example, if two respondents ranked a factor respectively two and 6 points each, the total sum of that factor is 2 + 6 = 8. The count is measured by the amount of respondents in this group that marked this specific factor in their ranking of the seven factors he/she found most important. In the example of two respondents ranking a factor and the other respondent in that user group did not include this factor in their top 7, the factor will have a count of 2 out of 3.

The results of all seven user groups are shown in Appendix C. In most user groups, similar factors ended up in the top 5. For almost all user groups, comfort and safety factors are equally represented. Only for pensioners there was a clear prioritisation of comfort factors. In addition to analysing all seven user groups separately, a weighted total is computed.

## 4.3 Weighted Total

All respondents combined represent the total group of bicycle-train users. However, when each group would be equal the respondents would not be representative for the total population. Therefore, the weighted total is computed. In this weighted total, the result of every group is multiplied by the share of that group in the total population. For example, the full-timers are multiplied by 26,5% to calculate the weighted sum and count for that factor for the largest group. To calculate the weighted total of each factor, all sums and counts for each group are added up to a total. This results in figure 4.2, where count has a maximum of 3 (if every respondent checked this factor in his top 7) and the sum has a maximum of 21 (if every respondent gave this factor 7 points).

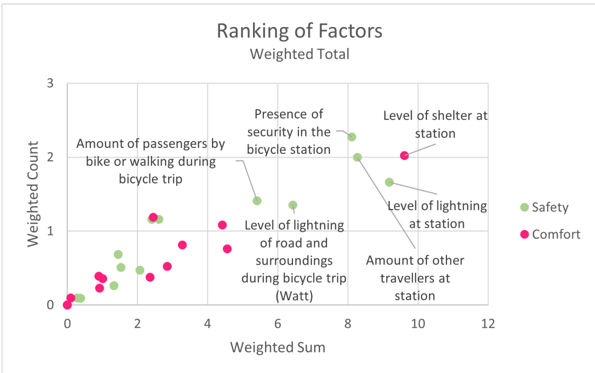



Figure 4.2: Ranking Total

The weighted total leads to the following ranking:
1. Level of shelter at the station
2. Level of lighting at the station
3. Presence of security in the bicycle station
4. Amount of other travellers at the station
5. Level of lighting at road and surroundings during the bicycle trip

When observing the data, an interesting distinction between male and female could be seen. In figure 4.3 the ranking of factors for all men is given, thus all user groups are included. In this figure the ranking factors labelled as comfort are overall more present than in any other group. In total, 11 men were interviewed, as opposed to 10 women. Therefore, the total count in the ranking of men is 11 and the possible total sum is 77.

The opposite is visible in figure 4.4, where all women interviewed are presented. In this gender, mainly factors labelled as safety are dominantly ranked. In the female ranking, the maximum count is

10 and the maximum sum is 70. An important note is that all interviewed females ranked the factor 'Amount of other travellers at station' in their top 7.

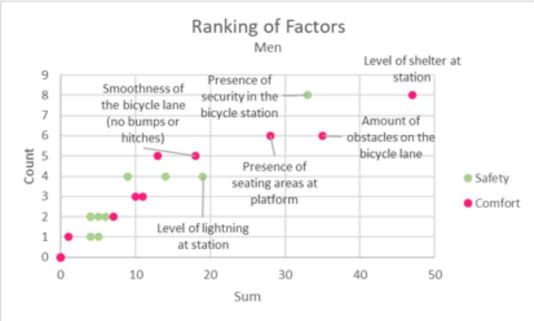



Figure 4.3: Ranking Men (N=11)

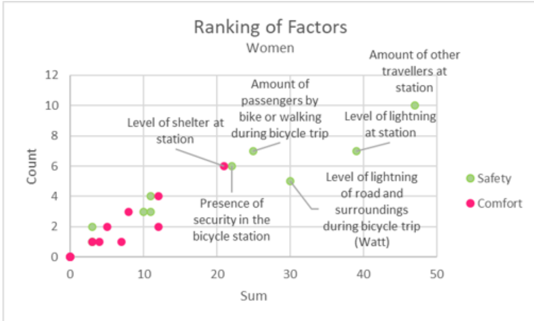



Figure 4.4: Ranking Women (N=10)

The difference in genders and their ranking is not unexpected. According to Stanko (2008) females are more aware of their (non-)safety compared to males. However, this is still very interesting to further analyse the statistical significance of this difference in priorities for each gender.

## 4.4 Conclusion

In the different user groups, most socio-demographic differences are included. However, not all user groups are equally represented in the total population. The weighted total is therefore the most useful outcome. This top 5 is used to determine the factors included in the stated choice experiment. All groups are included in this weighted total. It is useful to see that most user groups have similar factors in their ranking. For example, all groups have at least three of the same factors in their own ranking as in the weighted total.

Again, it should be noted that due to the small test population (N=21) the results might not be statistically significant. However, the factors ranked highest have a clear difference in value for the weighted count and weighted total.

Figure 4.2 presents an indicative answer on the second subquestion: "What are the factors related to perceived safety and comfort, influencing station choice for travellers using the mode combination bicycle - train, and their corresponding ranking?" Concluding the five most ranked satisfiers when weighted being:

1. Level of shelter at the station
2. Level of lighting at the station
3. Presence of security in the bicycle parking
4. Amount of other travellers at the station
5. Level of lighting at road and surroundings during the bicycle trip

# 5
# Stated Choice Experiment

Now that the satisfiers are ranked, in this chapter it will be decided which factors to include in the stated choice experiment. With the priority of research gaps in mind, not only satisfiers but also dissatisfiers will be included in the experiment. First, a pilot survey is conducted. Afterwards feedback and prior values are retrieved and implemented to construct the final survey.

## 5.1 Pilot Survey

There are two main reasons for using a pilot before sending out the final survey. The most important reason is to check the pilot on possible mistakes and to improve or clarify the survey before sending it to a large group of respondents. Another practical reason to do a pilot is to calculate the priors of the weights. These priors are the weights ($\beta$) of each factor that result from the pilot. This information can then be used to make a more efficient design of the survey. Meaning less choices are needed in the final survey for the same amount of information. The pilot choice experiment is constructed following 3 steps: model specification, generation of experimental design and the questionnaire construction (ChoiceMetrics, 2018). These steps are mainly done in this order but are also iterative.

### 5.1.1 Model Specification

The model specification includes defining the choice alternatives, containing their attributes and their attribute levels. A choice alternative is one option in the total choice set. This alternative has different attributes. Attributes are the characteristics of the choice alternatives. The attributes have different values in the alternatives that are weighted by the respondent. These different values for an attribute are called attribute levels. An overview of the terms in a stated choice experiment is given in figure 5.1

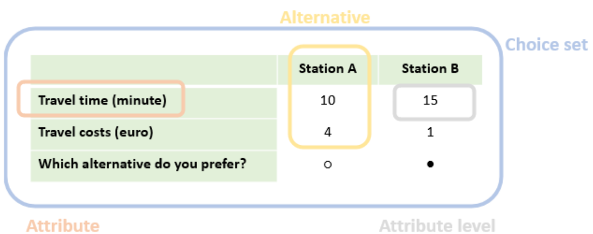



Figure 5.1: Terms stated choice experiment (adapted from Molin, 2019)

#### Alternatives

In a stated choice experiment different choice options are presented to a respondent and a choice between these options is asked. The choice options are called alternatives. There are two types of alternatives: labelled and unlabelled alternatives. Labelled alternatives are alternatives that can be labelled with a meaning. This is the case when a choice is asked between alternatives 'car' and 'train'. The labels car and train hold information and an alternative specific constant (ASC) is included in the utility function, representing all the characteristics associated with the label that are not captured within the attributes. The choice alternatives in this experiment are station A and station B, meaning that the choice alternatives are unlabelled, since A or B has no further meaning other than distinction. This is important when designing the choice alternatives and choice sets in section 5.1.2.

**Attributes**
The choice alternatives station A and station B are specified by using attributes, which are the different factors relating to station choice. A stated choice experiment works best with 2 - 6 attributes per choice alternative (Molin, 2019). The more attributes are included in each alternative, the more attention is asked of the respondent.

According to several research (Ton et al., 2019; Krabbenborg et al., 2015; Weliwitiya and Eng, 2020; Young and Blainey, 2018; Givoni and Rietveld, 2007) the factors time and costs are important choice factors relating to station choice. van Mil et al. (2020) proved that the directness of a train trip has a large influence on station choice. However, this factor is excluded in further steps of this research, due to the inability to easily change this. The intention of this thesis is not to adapt the timetable or other schedules. To focus on the current supply and demand, the factor transfer in train trip is excluded.

Regarding the importance of the factors access time and parking price, both factors are included in the stated choice experiment. This is due to the prioritisation of the research gaps. Where station choice is prioritised above satisfiers. This research's focus lies largely on including satisfiers, but the total station choice is simply not made only on the upper layers of the customer needs pyramid.

In the interviews, it was decided which satisfiers are included as attributes in the stated choice experiment. The weighted total results ranked the factors level of shelter, level of lighting, presence of security in bicycle parking, amount of other travellers at station and level of lighting during access trip. Due to the advised amount of attributes, not all five factors are used in this experiment.

After analysing the top 5, the amount of other travellers at the station and the level of lighting during the access trip are excluded. Leading back to the goal of this research being investigating if satisfiers can influence station choice, also applicable and effective measurements should be studied. Including the amount of other travellers present at the station would complicate this application. The amount of other travellers is dependent on too many different characteristics and including this factor requires a lot of extra research to effective measures. This is impossible in the time given for a thesis, however it cannot be denied that it is an important factor and therefore it is mentioned in the recommendations for further research.

The same reason holds for the factor lighting at access trip. This access trip is dependent on many other factors like route choice. Including this factor therefore does not have a direct suggestion for effective measurement, since route choice is a very complex concept.

This leads to five final attributes included in the stated choice experiment, being access travel time, bicycle parking fee, level of shelter at station, level of lighting at station and presence of security in bicycle parking.

**Model Type and Utility Function**
In the pilot stated choice experiment, a multinominal logit (MNL) model is used. In the final stated choice experiment a latent class choice model is added. This latent class choice model provides creating the different weights for all factors for different user groups. The chosen MNL model consists of two utility functions, hence two alternatives (station A and B) are present. Both alternatives have the same 5 attributes, meaning no alternative specific constants are included. For the pilot survey, the following utility function is applied.

$$U_A = \beta_{Light} * Light_A + \beta_{Shelter} * Shelter_A + \beta_{Security} * Security_A + \beta_{Time} * TT_A + \beta_{Price} * Costs_A \quad (5.1)$$

$$U_B = \beta_{Light} * Light_B + \beta_{Shelter} * Shelter_B + \beta_{Security} * Security_B + \beta_{Time} * TT_B + \beta_{Price} * Costs_B \quad (5.2)$$

Meaning the utility of station *i* is calculated by the total sum of the weight (Beta's) of every factor, multiplied by its corresponding factor value of the alternative. Where:

| | |
|---|---|
| $U_i$ | = utility of station *i* |
| $\beta_{Light}$ | = Weight for light intensity in station choice |
| $\beta_{Shelter}$ | = Weight for level of shelter in station choice |
| $\beta_{Security}$ | = Weight for presence of security at bicycle parking in station choice |
| $\beta_{Time}$ | = Weight for access travel time in station choice |
| $\beta_{Price}$ | = Weight for parking price in station choice |
| $Light_i$ | = Light intensity of station alternative *i* |
| $Shelter_i$ | = Level of shelter of station alternative *i* |
| $Security_i$ | = Presence of security at bicycle parking of station alternative *i* |
| $TT_i$ | = Access travel time by bicycle of station alternative *i* |
| $Costs_i$ | = Parking costs of station alternative *i* |

### 5.1.2 Generation of Experimental Design

An experimental design includes which choice situations are presented to the respondent in the stated choice experiment (ChoiceMetrics, 2018). Before determining which design to use, several decisions have to be made. Examples are if the design should be labelled or unlabelled, how many attribute levels are used and how many choice situations should be included.

In the model specification the utility function is written. In this utility function station A and B are used as alternatives. This means the alternatives are unlabeled since the name of alternatives only for distinction. No further meaning is included in the name of the alternatives. Also, no alternative specific constant is included in any utility function, meaning both alternatives have the same attributes. Related to unlabeled, sequential construction is applied for constructing the choice sets in a stated choice experiment with unlabeled alternatives. In sequential construction, the alternatives are constructed first before randomly placing them into choice sets.

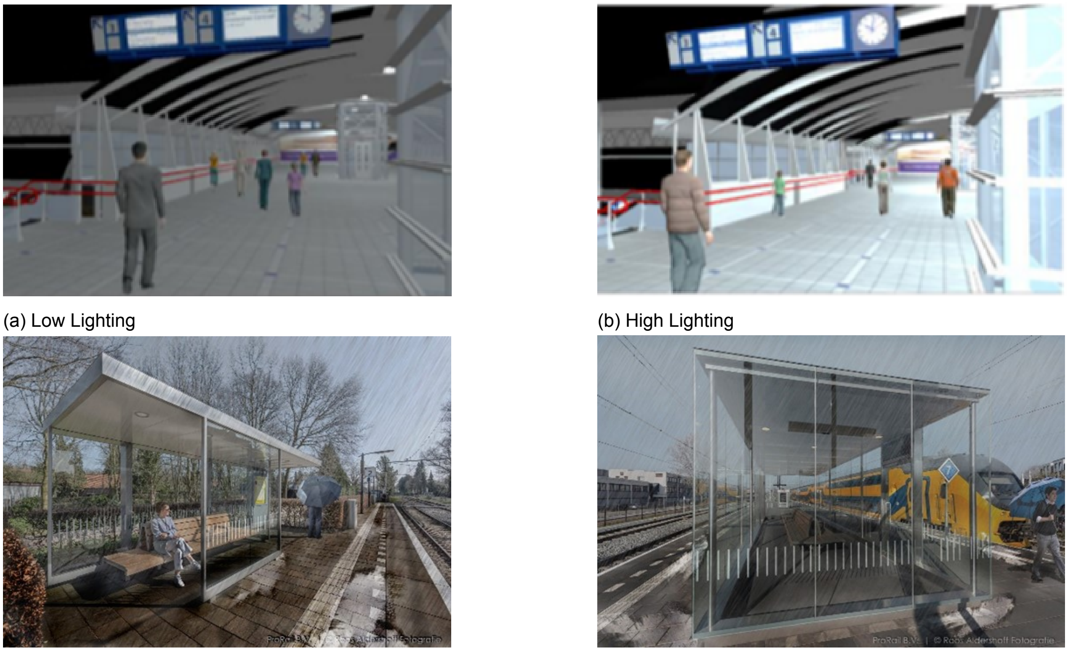
**Attribute Levels**

The actual choices of the five defined attributes is based on the values of the attributes, called the attribute levels. Typically, the amount of values is limited to 2-4 levels (Molin, 2019). Research showed the bicycle to be used most often for distances between 0.5 and 3.5 kilometres (Keijer and Rietveld, 2000; Martens, 2004). To convert travel distance to a more comprehensive factor travel time, including the attribute levels 5, 10, 15 and 20 minutes. The attributes relating to parking price for bicycles is based on the current monthly NS subscription of €0.50 per day. To not differ too much in attribute levels, also 4 levels are chosen for bicycle parking price. This is favourable in the next step, the generation of the pilot design. The pilot stated choice model is presented in figure 5.2. For the attributes shelter and lighting the levels high and low are indicated by using images in the survey, shown in figure 5.3.

**Station A / B**
**Access Travel Time (min): 5 / 10 / 15 / 20**
**Bicycle Parking Price (euro): 0 / 0.50 / 1 / 1.50**
**Shelter at Station: High / Low**
**Lighting Station: High / Low**
**Security in bicycle parking: Yes / No**

Figure 5.2: Stated Choice Pilot

The choice alternatives can be created according to various designs. The first option is a full-factorial design. This design type includes all possible choice combinations of alternatives. It is a very complete design, thus it includes many alternatives. For the model specified in figure 5.2, a



(a) Low Lighting

(b) High Lighting

(c) Low Shelter

(d) High Shelter

Figure 5.3: Images used for Shelter and Lighting in Survey

full-factorial design would include $4^2 * 2^3$= 128 alternatives. Leading to 128 choice sets in the experiment. This would consume a lot of time for respondents, even for the pilot survey.

A fractional-factorial design is a technique able in resulting a smaller number of alternatives. Fractional factorial designs are a fraction of the full factorial designs using basic plans or constructed by the software Ngene. The Ngene code used to create the fractional factorial design is written in appendix D. Within fractional factorial designs, there exist many different types of designs.

When generating the choice alternatives and choice sets in a fractional factorial design, two concepts are important: 1) orthogonality and 2) attribute level balance. The most well-known fractional factorial design type is an orthogonal design. An orthogonal design aims to minimize the correlation between the attribute levels in choices. A design is orthogonal if it satisfies attribute level balance and all parameters are independently estimable. Attribute level balance is a desirable attribute, but not as important as orthogonality. Attribute level balance includes all levels to appear an equal times within that attribute. Not having an attribute level balance affects the standard errors and creates unequal reliability in the estimates. But most importantly, an absent attribute level balance does not result in bias in the parameter estimates, which is the case when the design is not orthogonal.

To make this design both orthogonal and attribute level balance, a number of choice sets is needed which could be divided by 4 and 2 (the attribute levels). This would lead to, for example, 8, 12, 16 sets. Less choice sets are generally preferred in a stated choice experiment for respondents. However, since this pilot is also used to estimate the priors the more information gathered, the better. For this reason 12 choice sets are chosen in the pilot survey.

**Choice Set Generation**
After generating the choice alternatives the choice sets can be created. Choice sets consist of two choice alternatives. According to the sequential construction the alternatives are randomly coupled in a choice set. Each alternative being once station A and once station B. No similar pairs are included. To gain as much information as possible out of the pilot survey, also dominant choice sets are avoided. A dominant choice set holds that one alternative has a better or equal score on all attributes than the other alternative. For example in the choice set in figure 5.4. Every rational human being would prefer station B over station A. No special trade-offs are made.

The generated choice sets presented in the pilot survey are showed in appendix E. For a sequential

| | Station A | Station B |
|---|---|---|
| **Travel time (minute)** | 15 | 5 |
| **Travel costs (euro)** | 1,5 | 1 |
| **Level of lightning** | High | High |
| **Level of shelter** | Low | Low |
| **Presence of security** | Low | High |
| **Which alternative do you prefer?** | ○ | ● |

Figure 5.4: Example of dominant choice set

construction of choice sets it is important to check the correlations between attributes, these should be 0. This requirement is satisfied, which can also be seen in appendix E. Concluding that the pilot survey design is indeed orthogonal. Also when checking attribute level balance this is complied.

### 5.1.3 Questionnaire Construction

Besides the model specification and the generation of the design used for the survey, also information about the respondent is asked. Socio-demographic information helps to understand the composition of different user groups of the mode combination bicycle - train. Based on the socio-demographic questions mostly asked and the groups resulting from Ton et al. (2020) the following characteristics are asked: age, gender, education level, household composition, social participation (employment status), average days taking a train trip (before covid), the average trip goal of train trips, average access mode of train trips, annual gross household income and the first numbers of ZIP code.

After the 12 choice sets, two questions are asked that can be used for validation. First, if the respondent has two or more train stations within 20 minutes of cycling at home. And on what factor the respondent based his choices on the most (time, costs, lighting, shelter or security). The pilot survey is shown in appendix F.

### 5.1.4 Pilot Results

The pilot survey was filled in by 38 respondents. Similar to the interviewees, these respondents all varied on personal characteristics such as age, education, gender, income and household. The descriptive statistics of the respondents are presented in appendix G. Even though this variation in respondents is present, not all groups are equally represented in this pilot as they are in society. For this pilot that is not a large problem, but it should be taken into account for the final experiment.

To retrieve the weights of factors the package BisonBiogeme (Bierlaire, 2003) is used. This is a program able to estimate the values of $\beta$ from the choices made by respondents. The model code created is written in appendix H. This model is written according to an MNL model. This is valid since, for the pilot model, there is no latent class necessary. The found weights are presented in table 5.1. A weight of 0.941 for lighting means that the utility of a station increases by 0.941 when a high level of lighting is provided, compared to zero increase when low level of lighting is present. For shelter and security a higher level increases the utility by 0.306 and 0.761 compared to a lower level. For time and price the calculation is different, due to the negative relation and different attribute levels. When access time is increased by 10 minutes, the utility of that station decreases by 10*0.126=1.26. And when the parking fee is increased to €0.50 the utility will decrease by 0.5*1.16=0.58.

The resulting weights are used to create an efficient final design. Because an indication of the weights is known, a final design can be created with less choice sets and the same or more information for each set. This increases the comfort for respondents since less time is needed to complete the survey.

The risk of using priors for a more efficient design is the chance of these priors being inaccurate

| Name | Value | Std. error | t-test | p-value |
|---|---|---|---|---|
| $\beta_{Light}$ | 0.941 | 0.144 | 6.55 | 0.00 |
| $\beta_{Shelter}$ | 0.306 | 0.164 | -6.80 | 0.05 |
| $\beta_{Security}$ | 0.761 | 0.185 | 4.10 | 0.00 |
| $\beta_{Time}$ | -0.126 | 0.0185 | -6.80 | 0.00 |
| $\beta_{Price}$ | -1.16 | 0.192 | -6.04 | 0.00 |

Table 5.1: Pilot Results

(ChoiceMetrics, 2018). Where the final design of the survey is not very sensitive to the precise weight itself but foremost to the right sign of the priors (+ or -). Therefore the resulting priors are checked by their sign (positive or negative) and compared with previous research and the expected signs. The signs seem to be correct since time and costs have a negative association with the total utility, whereas light, shelter and security have a positive relation. Also, the Value-of-Time (VoT), which is a ratio of the weights for time and costs (ChoiceMetrics, 2018), is similar to the VoT found by van Mil et al. in 2020. Where for access travel time he found the value -0.19 and for travel costs -1.77, leading to a VoT of 0.11 euros per minute. The VoT of the pilot survey is also 0.11 euros per minute. The p-value of most factors is 0.00, meaning these weights are statistically significant. Only the factor shelter has a p-value of 0.05, which concludes this factor to be statistically insignificant at the 95% level, which is not unexpected with few respondents. Since this pilot is only used for improvement and no direct conclusions are drawn this prior for shelter is still the best guess of the actual weight. The lower significance could be a result of the number of respondents. Where the final result is based on a larger group of the population.

### 5.1.5 Reflection Pilot Survey

Now that the priors are found, the survey is improved according to retrieved feedback of respondents and bumps encountered. For the final survey, adaptions are made in the questions concerning personal characteristics, choice sets and the closing questions.

In the first part of the survey the following changes are made:

- Many respondents said the introduction was not always clear and concise, leading to increase the response time. This is altered by rewriting the introduction and clarifying by making important terms bold.
- Income was first asked per household. This is harder to answer by respondents since it is not always known. Therefore, this is changed halfway during the pilot to solely personal income of that respondent.
- Only in the question regarding trip frequency the time being pre-Covid was specified. This specification was not consistent throughout all socio-demographic questions. In the final survey also age, income, household, trip goal, social participation and education level should be defined as pre-Covid.
- Household characteristics did not include any outliving children. To some respondents this was not clear. In the final survey the household composition is improved by including the term 'children living at home' in the options.
- The questions concerning train trip access mode and trip goal can not be answered when the respondent stated he/she never travels by train. In the final survey these questions can be skipped when the respondent never uses the train mode.
- For respondents before the bachelor/master in education (2002) the question about education level was not complete. In the final survey also older terms of some education levels are added and it is clarified to check the most applicable option.

Regarding the choice sets in the pilot survey the following feedback is gathered:

- The comment pointed out by most respondents was the number of choices being too high. In the pilot survey there are 12 choice sets. In the final survey this number can be decreased by using information gathered in this pilot.
- To some respondents it was unclear whether or not the choices influenced the train trip, regarding transfer or total travel time. In the final survey this is now more specific included in the explanation of the choice sets.
- The pictures illustrating the level of shelter were made in the summer. This confused some respondents since the context given was an average day in autumn with chance on precipitation. This can also be an explanation of the relative lower resulting weight of this factor compared to others. In the final survey this illustration is adapted to match the context description.
- When asked what factors were mainly used to make the station choice, the factors travel time and price of the bicycle parking were indicated by most respondents. To gain more information about the station choice regarding satisfiers, it is useful to prioritise the factors related to safety and comfort. In the final survey the satisfier factors are prioritised by reducing the attribute levels and number of attribute values of the factors time and price. This is based on many comments of respondents stating they were not willing to add 10 minutes of extra access time or 2 euros of parking. In addition, the difference between the attribute level values are changed with a smaller difference. Both changes allow for more focus on trade-offs in the satisfier attribute levels (Ohler et al., 2000). Reducing the number of levels per attribute and range of attribute levels also reduces the experimental complexity for respondents (Wang and Li, 2002).

For the last part of the survey, the closing questions, were adapted:

- Some respondents reached out and noticed strange combinations of answers could be given relating to the choice sets and closing questions. These answers are verified per respondent so unrealistic combinations can be excluded.
- During the pilot it was noted that the type of bicycle of the respondent matters in the weight corresponding to security in the bicycle parking. In the final survey a question about the bicycle used in access to the station is added.
- One respondent stated that by asking a fee for using the bicycle parking, travellers using this mode combination might park their bicycle incorrectly in the station environment. This could be a valid point, however, this is outside of the scope of this research and will therefore not be adapted in the final survey. It is included in further recommendations for research.

For a more professional experience for participants, the final survey is created in Qualtrics. This is a partner of the TU Delft specialised in designing surveys for research purpose.

## 5.2 Final Survey

After improving the pilot survey regarding the feedback and gaining the prior information the final survey is constructed. The changes in attribute levels of the factors access travel time and bicycle parking price are made to prioritise the other three factors in this research. The attributes and attribute levels of the final stated choice experiment is illustrated in figure 5.5.

### 5.2.1 Final Survey Design

**Efficient Design**

In the pilot survey information is gathered about the weights of factors. This can be used to make an efficient final survey design (ChoiceMetrics, 2018). Like the orthogonal design, an efficient design is also a fractional factorial design. An efficient survey results in a survey with a lower amount of choice sets presented to each respondent and where the choices made hold more information on average. Compared to orthogonal designs, efficient designs result in smaller standard errors with the same number of respondents compared to orthogonal designs (“Constructing Efficient Stated Choice Experimental Designs”, 2009).

**Station A / B**
**Access Travel Time (min): 5 / 7.5 / 10**
**Bicycle Parking Price (euro): 0 / 0.5 / 1**
**Shelter at Station: High / Low**
**Lighting Station: High / Low**
**Security in bicycle parking: Yes / No**

Figure 5.5: Stated Choice Final Survey

There are three types efficient designs (ChoiceMetrics, 2018). These designs use the AVC (asymptotic variance-covariance) matrix of the parameters that are derived by the know priors. An A-efficient design, minimizes the sum of variances of the AVC matrix. It ignores the covariates between attributes and is not very useful if measurement units of attributes differ (Molin, 2019). D-efficient designs are based on the determinant of de AVC matrix. It is a generally preferred efficient design since it takes both variance and covariances into account. The last efficient design type is S-efficient design. This minimizes the standard error of the parameter for which it is hardest to reach statistical significance. It should be noted that the D-error of the S-efficient design is much larger. A larger D-error indicates an overall less efficient design (ChoiceMetrics, 2018). Concluding you win reliability on one parameter with an S-efficient, but you lose reliability on most other parameters. In this research all parameters should be treated equally important and an overall efficient design is preferred. Leading to the usage of a D-efficient design for constructing the final choice experiment.

Next to the advantages of the efficient design being a shorter survey and reducing the standard errors of parameters there is also a disadvantage. It should be taken into account that the efficient design is based on retrieved priors from the pilot survey (Molin, 2019). This means extra effort is needed for an efficient design. Also, a risk exists that using the wrong priors leads to incorrect results. Luckily, this risk is mainly based on the sign of the priors. Meaning if the priors are a correct positive or negative value. The access travel time and bicycle parking price being negative corresponds to research of van Mil et al. (2020). The factors shelter, lighting and security have an expected positive value. Assuming that the final survey respondents behave approximately the same as the pilot respondents, a d-efficient design is used.

After implementing the feedback from the pilot survey and determining the final attribute levels, the final choice set can be constructed.

**Final Choice Set Construction**

The pilot study consisted of 12 choice sets. A feedback point was that respondents experienced this as too long. Using an efficient design, this number of choice sets can be reduced. For orthogonal designs, attribute level balance is achieved by definition (Molin, 2019). For efficient designs attribute level may be relaxed, but is still preferred.

For choice sets with two alternatives, the number of required choice sets is determined by the number of parameters ($\beta$'s) to be estimated (equation 5.3). Which leads to a minimum of six (5/1+1) choice sets. This could be enough choice sets, however it is a bare minimum. In a later stage of this research testing for linearity might be useful (ChoiceMetrics, 2018). To do this, at least 2 parameters are needed for the attributes time and costs. Which leads to 7 (5+2) choice sets. Unfortunately, using 7 choice sets means the attribute level balance is relaxed. To achieve attribute level balance again, 12 choice sets are needed. However, the benefits of adding 5 choice sets to gain attribute level balance does not weigh up to the increased time to complete the survey for respondents. Using 10 choice sets means 60% of the attributes achieve attribute level balance. For the factors time and price one level is included in the choice sets an extra time compared to the others.

$$\#choicesets > \#parameters/(\#alternativesperchoiceset - 1) \tag{5.3}$$

Using NGENE, a final choice set is constructed with 10 choice situations. The programming code is written in appendix I. In the utility functions the priors resulting from the pilot are incorporated. Several MNL efficiency measures are shown in table 5.2. The minimal D error found was 0.338, this is a rather low D error. A lower the D error shows a more efficient design. There is no universal acceptable range for a D-error since this is dependent on how many choice situations are included. The corresponding S estimate is 23.15, meaning at least 24 respondents are needed to obtain statistically significance for all parameters, if the respondents behave equally to the pilot respondents.

| | |
|---|---|
| D error | 0.338014 |
| S estimate | 23.15 |

Table 5.2: MNL efficiency measures

An efficient design should have some degree of utility balance. No dominant alternatives are preferred in order to gain more information from all choice situations. None of expected probabilities for choosing alternative$_i$ should be exceeding 0.9 (Molin, 2019). The value for S-estimates is the required N to obtain statistical significance. In appendix J the utilities are calculated for each choice situation. No expected utility is higher than 0.9. This appendix also shows additional MNL efficiency measures. The final survey used is presented in appendix K.

# 6
# Descriptive statistics

This chapter discussed the descriptive statistics of the choice observations obtained in the final survey. In section 6.1 the data collection of the final survey is explained, followed by a description of socio-demographics of the respondents (section 6.2). In this section, the final dataset used in this research is compared to a larger dataset (ODIN-2019) for validation. The last section (6.3) shows an overview of all the choices made by respondents.

## 6.1 Data Collection

There is no need to have a sample group as large as the whole population, however more representative results are expected when more respondents are gathered. The S-estimate retrieved using Ngene stated a minimum of 24 respondents in the sample size is needed in order to obtain a statistically significant parameter estimate at a 95% confidence interval (ChoiceMetrics, 2018. However, this low amount of respondents is only sufficient if the sample group behaves the same as the pilot group. In order to collect more respondents, several methods are followed to spread the final survey.

A large part of respondents was gathered by spreading the link on social media and throughout own network. To ensure different socio-demographic characteristics also acquaintances are asked to spread. The survey was spread on different groups on social media where respondents are asked. Halfway through gathering respondents, an initial draft is made of the descriptive statistics. This gave more insight to the respondent which were underrepresented. In addition, the survey was spread via an online panel called PanelClix. Finally, some respondents were gathered by spreading the survey on the central station of Utrecht. The survey was spread on a Friday afternoon, attempting to gather both work and leisure travellers.

In total, 312 respondents filled in the final survey, of which 4 respondents completed the survey in under 90 seconds. These responses are excluded from analysing the results since it is assumed that it takes at least 90 seconds to complete the final survey. The average duration of completion was 3 minutes and 16 seconds. No further screenings are made which leads to a total of 308 respondents used for data analysis.

## 6.2 Socio Demographics

Several questions regarding personal characteristics where asked in the survey. An overview of the distributions of respondents is included in Appendix M.1. Also the geographical location of respondents is shown (Appendix M.2). A nice spread of location is seen amongst the respondents. More respondents are located in the randstad, which is not surprising since many acquintances asked to participate are habitants in the province South-Holland. Also compared to the rest of The Netherlands, the randstad is quite crowded. This could explain the representativeness of these areas.

The socio-characteristics are not only used to find different usergroups within passengers of public transport. This socio-demographic information can also be used to compare the sample of respondents to a larger data set. The dataset used for this validation is ODIN-2019. This dataset contains a large sample of passengers who are all asked to describe their travel trips of that day. In this dataset also socio-demographic characteristics are included. To validate the stated choice experiment used in this research only trips which have train as a main mode are included. When the distribution of the social-demographic characteristics do not differ significantly between sample and population, the sample is representative for these variables in the population (Molin, 2019).

Several diagrams are presented to show the differences and similarities between ODIN and the final survey. In these diagrams, the inner distribution is from ODIN 2019 and the outer distribution is from the final survey. For gender (figure 6.1), the similarities between ODIN 2019 and the final survey are almost perfect. Gender distribution is equal amongst train travellers (Jonkeren et al., 2021). Also education level and social participation do not show large differences in distributions between the two data sets.

Considering age, there are less 25-34 year old's included in the final respondents list, 26% versus 18%. Instead, there are more representatives in the 35-64 year old group, 44% versus 33%. Also a small difference is noticed considering household composition. In the respondents of the final survey there are 8% more student houses checked. This is remarkable since the social participation student does not show a large difference. The difference in household composition could be explained by the fact that also many pupils <17 are asked in the ODIN dataset, whereas this group is underrepresented in the final dataset of this survey.

One of the largest inequality between the datasets is shown in the characteristics train frequency (figure 6.1). In the ODIN dataset there are many more respondents travelling more often with the train than in the final dataset. This is due to the filter used in the ODIN dataset. To compare the respondents of this research to train travellers only the trips having train as the main mode are analysed. This filter allows for the representativeness of train trips. This could be related to the characteristic trip purpose (figure 6.1), where another difference is noted. In the final dataset work and leisure represent respectively 36% and 46%, whereas in the ODIN dataset these are switched in 46% and 32%. This change could be the underlying cause that travellers with a higher frequency often are using this mode for work or study.

Another notable dissimilarity is seen in the distribution of income (figure 6.1). In the outer circle diagram (the final dataset), slightly more higher incomes are represented. But the final dataset also includes many more incomes of below <€10.000 per year. This could be explained by the more students living independent in this dataset.
There are several small differences between the final dataset and the ODIN 2019 dataset. Since these differences are not very large or can be explained the results are suitable to proceed with, however the interpretation should be done with care. This is reflected on in the next chapters, final results and case study.

## 6.3 Choices in Data-set

The respondents answered ten choice sets in the final survey. For every choice set, station A or station B is chosen. The percentage of each choice is shown in figure 6.2. Since all choice sets include different values for the parameters, no specific results can be analysed yet. However, when comparing these final choices with the expected probabilities computed from the pilot (Appendix J), it can be noted that these are fairly similar. This means that the group of respondents of the final survey show the same behavior as the pilot group, which validates the efficient design.


Figure 6.1: Comparison ODIN (inner) and final (outer) dataset on various socio-demographics

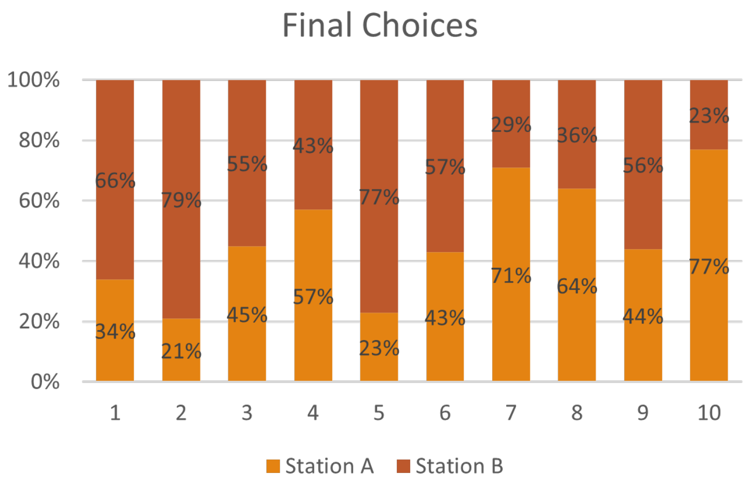



Figure 6.2: Final choices made

# 7

# Final Results

In this chapter the results stemming from the final survey are presented. Two models are applied, first, a MNL model is to compute the weights for all parameters. This is described in section 7.1. In section 7.2 a Latent Class Model is used to retrieve weights per group of respondents, which allows for heterogeneity between respondents.

## 7.1 MNL Model

First, results are retrieved using all respondents as one group, thus without making classes. This is done using the same Biogeme model as presented in appendix F, adapting the data set to the total data set. The model used to retrieve parameter weights is the MNL model. The names of the alternatives in the choice sets do not have another purpose next to distinction. This means the alternatives are unlabeled. The decision rule used in combination with the MNL model is the Random Utility Maximisation theory, which assumes people choose the alternative with the highest utility. Therefore, this model maximizes the utility per alternative and estimates the value of weights for the different parameters. The following utility function is used in the MNL model:

$$U_i = \beta_{Light} * Light_i + \beta_{Shelter} * Shelter_i + \beta_{Security} * Security_i + \beta_{Time} * Time_i + \beta_{Price} * Costs_i \quad (7.1)$$

Using an MNL model also means the results of this model can be compared to the results retrieved from the pilot, since they should be somewhat similar. This is confirmed by computing a final Value-of-Time of 0.11 euros per minute, which is the same as the VoT computed in the pilot. Some differences in the weights for light and security are noted between the pilot and final results. The ratio of costs and the parameter-weights $\beta_{Light}$ and $\beta_{Security}$ are lowered by approximately 0.2, compared to the pilot results. This means the final data proves the respondents are less sensitive to these satisfiers in station choice. All other ratios in weights for the parameters are roughly the same. The final results are presented in table 7.1.

| Name | Value | Std. error | t-test | p-value | | attribute levels | min. | max. |
|---|---|---|---|---|---|---|---|---|
| $\beta_{Light}$ | 0.667 | 0.050 | 13.22 | 0.00 | | 2 | 0 | 0.667 |
| $\beta_{Shelter}$ | 0.364 | 0.040 | 9.02 | 0.00 | | 2 | 0 | 0.364 |
| $\beta_{Security}$ | 0.554 | 0.048 | 11.52 | 0.00 | | 2 | 0 | 0.554 |
| $\beta_{Time}$ | -0.122 | 0.010 | -11.80 | 0.00 | | 3 | -0.61 | -1.22 |
| $\beta_{Price}$ | -1.09 | 0.064 | -17.00 | 0.00 | | 3 | 0 | -1.09 |

Table 7.1: Total results of MNL model and characteristics of parameters

All parameter values for the weights show the expected sign. Presence of lighting, shelter and security is positively rewarded. Whereas more access time and parking costs are perceived negatively by the traveller. Results should be interpreted in the following way. By increasing the lighting at the station by one point (from absent to present), the average utility experienced for that station increases by 0.667. For providing more shelter on the platform and security in the bicycle parking the utility of that station increases respectively by 0.364 and 0.554. It should be mentioned that the utility increase due to lighting and shelter is very dependent on the specific images used in this research. When adding one minute of cycling time or one euro of parking costs, the utility of that station decreases with 0.122 and 1.09. Meaning travellers are more sensitive for one euro extra parking price than for one minute extra access time. The resulting weights in table 7.1 have 'utility' as interconnecting unit.

The value-of-time can be computed by dividing the beta of time by the beta of price, resulting in a VoT of 0.11 euros per minute. Meaning the respondents are willing to pay 0.11 euros to lower the cycling time by one minute. A VoT of 6.72 euros per hour is lower than the average VoT of 10.50-15.91 euros per hour for cycling (Börjesson and Eliasson, 2012). This could be explained by the difference in values for the attributes time and price. The levels of these attributes vary slightly in the choice sets presented to the respondents. This was done to be able to focus on the satisfier attributes. The underestimation of the Value of Time should be taken into account while drawing conclusions.

Parameter weights between two models can only be compared in ratio, for example the VoT. In order to compare the results with other models, and to show the results more clearly, all parameter ratios with the dissatisfiers costs and time are computed. The ratio values in figure 7.1a and 7.1b are in the unit of the base parameter, respectively euros and minutes. The interpretation of results is indicated by a few examples:

Increasing the lighting at the station from Low to High is equal to decrease the parking costs by €0.61. In other words, respondents are willing to pay €0.61 euros for a well-lighted station.
Providing security at the bicycle parking is equal to decreasing the parking costs by €0.51. In other words, respondents are willing to pay €0.51 to increase bicycle parking security.

Increasing the lighting at the station from Low to High is equal to decreasing the access time by 5.5 minutes. In other words, respondents are willing to cycle 5.5 minutes extra for a station well-lighted station.
Providing security at the bicycle parking is equal to decreasing the access time by 4.5 minutes. In other words, respondents are willing to cycle 4.5 minutes extra for a station including bicycle parking security.

### 7.1.1 MNL including interaction of personal characteristics

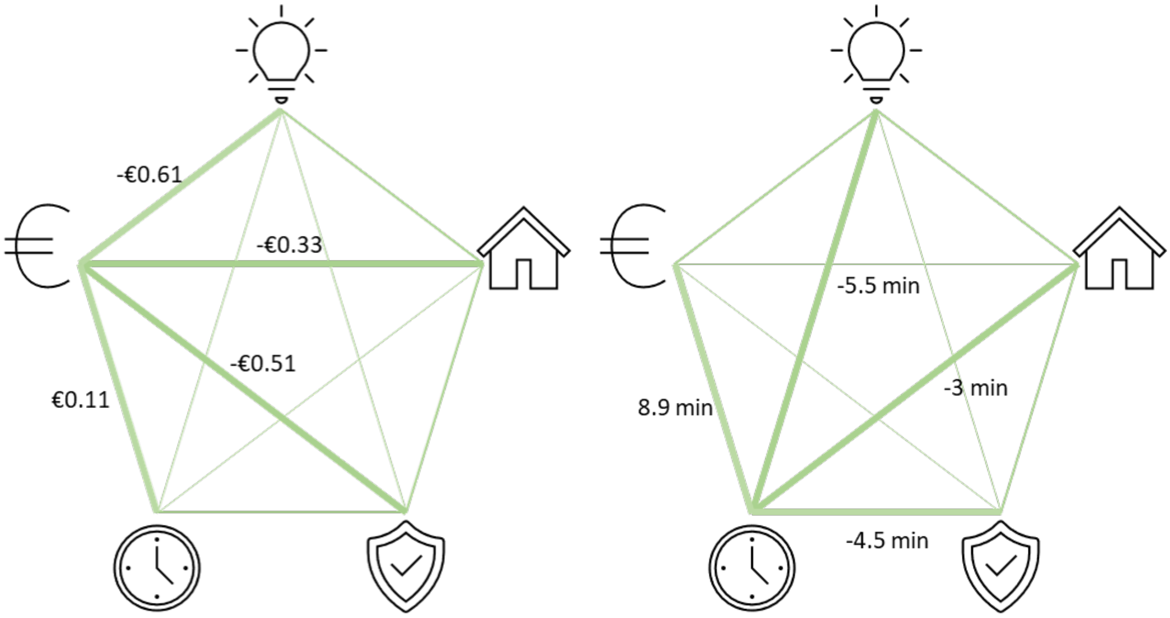
It is also possible to estimate the factor weights per personal characteristics. By computing the weights every distribution of weights per personal factor can be compared. The MNL model and its including interactions with socio-demographic variables are estimated using Biogeme. The results of these interaction effects are shown in table 7.2. The interaction model is coded per socio-demographic characteristic, for all parameters at once. An extended version including all p-values is shown in appendix L.

When computing the weights, every personal characteristics has a base case, indicated by a '*'. The value for a personal characteristic is chosen for base case when it has the most respondents in that group. Other values for the same characteristic are compared with this base case. For example, a female respondent (base case) has a weight of -0.104 for travel time, whereas a male respondent has a weight of -0.144 for this parameter. Interaction effects are found to be significant on a 95% confidence interval, which means the parameters have a p-value lower than 0.05. This significance means the difference between that value and that of the base case is statistically significant. Significant parameters are marked in bold and green. The large and notable significant interaction effects are:

- **Gender and lighting + parking costs**: Compared to male respondents, females are more sensitive for the presence of lighting at the station. This is in line with the difference in gender in ranking the satisfiers. Where level of lighting was ranked highly by many females, it was not

Table 7.2: Interaction effects on socio-demographic characteristics

| Socio-demographic Characteristics | Light 0.667 | Shelter 0.364 | Security 0.554 | Cycle time -0.122 | Parking costs -1.09 |
|---|---|---|---|---|---|
| Gender | | | | | |
| ** Female* | **0.746** | **0.366** | **0.592** | **-0.104** | **-0.868** |
| *Male* | **0.584** | 0.368 | 0.515 | **-0.144** | **-1.349** |
| Age | | | | | |
| *<17 years* | 0.342 | 0.471 | 0.783 | -0.086 | -0.734 |
| *18/24* | 0.769 | **0.597** | 0.548 | **-0.196** | **-2.13** |
| *25/35 years* | 0.784 | 0.292 | 0.616 | -0.142 | **-1.49** |
| **35/64 years* | **0.631** | **0.310** | **0.530** | **-0.091** | **-0.549** |
| *>65* | 0.680 | 0.300 | 0.597 | -0.111 | -0.104 |
| Education | | | | | |
| *None* | 1.106 | 2.752 | 1.63 | **-0.058** | -1.289 |
| *Basis* | 0.418 | 0.643 | 0.785 | -0.133 | -1.27 |
| *VMBO* | 0.444 | 0.255 | 0.427 | **-0.039** | **-0.435** |
| *HAVO/VWO* | 0.781 | 0.323 | 0.654 | -0.118 | -1.317 |
| *MBO* | 0.517 | 0.294 | 0.432 | **-0.056** | **-0.603** |
| ** HBO/WO Bachelor* | **0.787** | **0.436** | **0.557** | **-0.142** | **-1.46** |
| *WO Master* | 0.692 | 0.391 | 0.634 | **-0.205** | **-1.04** |
| *PhD* | -0.056 | 0.302 | **2.365** | **-0.118** | -2.707 |
| Household | | | | | |
| *Single* | 0.614 | 0.427 | 0.498 | -0.098 | -1.054 |
| *Two parents no kids* | 0.591 | 0.331 | 0.596 | -0.079 | **-0.655** |
| ** Two parents 1+ kids* | **0.723** | **0.220** | **0.636** | **-0.127** | **-1.02** |
| *Single 1+ kids* | 0.376 | **0.649** | **0.152** | -0.147 | -0.598 |
| *Student* | 0.956 | **0.674** | 0.514 | **-0.254** | **-2.39** |
| Work Participation | | | | | |
| *No job* | 0.853 | 0.463 | 0.391 | -0.016 | **-0.273** |
| ** Fulltime* | **0.652** | **0.327** | **0.564** | **-0.117** | **-0.976** |
| *Parttime* | 0.581 | 0.239 | 0.511 | -0.100 | **-0.605** |
| *Student* | 0.838 | **0.603** | 0.620 | **-0.211** | **-2.205** |
| *Pension* | **0.436** | **0.367** | 0.647 | -0.067 | **-0.118** |
| Days using train | | | | | |
| *4+ per week* | 0.547 | 0.381 | 0.576 | -0.145 | -1.368 |
| ** 1-3 per week* | **0.681** | **0.449** | **0.513** | **-0.151** | **-1.300** |
| *1-4 per month* | 0.707 | 0.388 | 0.488 | -0.116 | -1.145 |
| *<1 per month* | 0.930 | **0.190** | 0.800 | **-0.087** | **-0.571** |
| *Never* | 0.19 | **0.431** | 0.368 | -0.036 | -0.175 |
| Trip purpose | | | | | |
| *Work* | 0.642 | 0.297 | 0.547 | -0.143 | -0.950 |
| *School* | 0.671 | 0.423 | 0.560 | -0.112 | **-1.726** |
| ** Leisure* | **0.694** | **0.402** | **0.568** | **-0.111** | **-0.989** |
| Acces mode | | | | | |
| *Walk* | **0.924** | 0.433 | 0.702 | -0.135 | -0.905 |
| ** Bicycle* | **0.579** | **0.340** | **0.546** | **-0.117** | **-1.220** |
| *Car* | 0.561 | **0.058** | 0.499 | -0.091 | **-0.669** |
| *PT* | 0.889 | **0.682** | 0.513 | -0.165 | -1.033 |
| Yearly income | | | | | |
| *None* | 0.575 | 0.523 | 0.879 | -0.199 | **-2.320** |
| *<€10.000* | 0.830 | **0.551** | 0.558 | -0.153 | **-1.755** |
| *€10.000-€30.000* | 0.563 | 0.248 | 0.487 | **-0.056** | -0.747 |
| ** €30.000-€50.000* | **0.724** | **0.233** | **0.539** | **-0.131** | **-0.960** |
| *€50.000-€70.000* | 0.639 | 0.283 | 0.650 | -0.146 | -0.818 |
| *>€70.000* | 0.705 | 0.432 | 0.687 | -0.159 | **-0.451** |
| *Unknown* | 0.580 | **0.558** | 0.356 | -0.067 | **-0.516** |
| Two stations nearby | | | | | |
| ** Yes* | **0.636** | **0.419** | **0.530** | **-0.127** | **-1.24** |
| *No* | 0.714 | 0.299 | 0.590 | -0.111 | **-0.923** |
| Bicycle type | | | | | |
| ** Bicycle <€150* | **0.834** | **0.519** | **0.525** | **-0.169** | **-1.62** |
| *Bicycle >€150* | 0.695 | **0.134** | 0.624 | **-0.099** | **-0.771** |
| *Sport* | **0.018** | 0.329 | 0.628 | **-0.001** | -1.065 |
| *Electric* | **0.385** | 0.346 | 0.640 | **-0.066** | **-0.457** |
| *No bicycle* | 0.043 | 0.991 | 0.235 | -0.003 | -0.306 |
| *Never to station* | 0.742 | 0.297 | 0.581 | -0.136 | **-0.264** |




(a) MNL Costs as base parameter

(b) MNL Time as base parameter

Figure 7.1: Pentagon Interrelation Costs and Time as base

even in the top 5 in the ranking by males. Another interaction with gender is related to parking costs. Male respondents are more sensitive to the parking costs compared to females.

- **Work participation + shelter, travel time and parking costs**: For the characteristic 'Work participation', full-time workers are the base case. Compared to this group, students assign a higher positive value to shelter with 0.603 compared to 0.327 of full-time workers. Next to this, students experience travel time and parking costs (almost) twice as negative.
- **Days using train + shelter and parking costs**: Respondents that travel less than once per month have a lower weight for shelter and parking costs. This could be explained since the parking costs add up for each time a trip by train is taken. Thus for respondents travelling more, the parking costs become more important for their station choice.
- **Two stations + parking costs**: The respondents that do not have more than one station nearby their house are less sensitive to the parking costs of the station within 15 minutes of cycling.
- **Bicycle type + lighting, cycle time and parking costs**: For the characteristic bicycle type the base case is chosen to be a bicycle with a maximum worth of €150. Respondents using a sport bicycle or an electric bicycle are less sensitive for lighting present at the station. Also the cycle time is rated less negative. This could be explained by the comfort experienced while cycling on a more expensive bike, so the access trip is rated less negative. Next to cycle time, parking costs seem less important when the value of the bicycle used increases.
- **Education + cycle time**: Overall a trend can be seen that the higher the education, the more negative the weight for cycle time. An exception is the PhD respondents, this could be due to their bicycle type. This observation is explained below.

Other significant interaction effects can also partly be explained within socio demographic characteristics. For example, the interaction between education and security in the bicycle parking. A significant difference is shown where PhD respondents assign a weight of 2.365 to security in the bicycle parking, against the 0.557 of HBO/WO Bachelor respondents. When checking the data the PhD respondents all owned a more expensive bicycle. This could explain the significant larger value for the weight of security.

Another similar observation is between the characteristics work participation, income and trip purpose. This is mainly due to the group consisting of students. In all characteristics, the option student (or school/lower income/ and age 18-24) shows the same significant differences. This group has a high value for shelter at the station, a strong dislike for travel time compared to other respondents and a more negative weight for parking costs. This weight for parking costs is even often twice as low compared to the base case. This also shows in the socio-demographic characteristic household, when student housing is checked.

## 7.2 Latent Class Method

In the MNL model, cross sectional data is used. Cross sectional data threats observations the same as ID's. Panel data could section the observations made by the same person. 1,000 choice mode observations of 10 individuals (i.e. 10 × 100 observations) do not provide the same amount of information on the sample mean as 1,000 × 1 choice observation. Not accounting for the panel structure leads to underestimation of the standard errors thus, overestimation of the accuracy by which the parameters are being retrieved (van Cranenburgh, 2021). Panel data structure is accounted for in a latent class model, meaning the respondent ID is coupled within each choice set.

To create latent classes and to compute their factor weights, a model is constructed. The Matlab code is included in appendix N. The validation of this model is done by checking if the outcomes are reasonable and it is checked if the total results of both classes are correspond to the MNL model. Next to estimating the Latent class choice model, the MNL model is estimated using the Matlab code, thus without making any classes. This results the same parameter values and their corresponding standard errors and p-values calculated by Biogeme.

### 7.2.1 Number of Classes

The log-likelihood and rho-squared values improve by adding more parameters. However, adding parameters can lead to overfitting. To determine which number of classes fits the best, the Bayesian Information Criterion is used (Van Cranenburgh, 2018). This criterion uses a penalty term for the number of parameters in the model to solve the problem of overfitting. The lower the value of BIC, the better the model fits the data. The BIC value is computed according to equation 7.2.

$$BIC = k * ln(N) - 2 * LL \tag{7.2}$$

Where:
k = number of parameters to be estimated
N = number of observations
LL = LogLikelihood of class model

As shown in table 7.3, the model with 3 classes has the lowest BIC value and therefore has the best model fit. However, when creating the 3 classes one group is very small of size (7% of the population). Due to the sample of 308 respondents, this group of 7% would be represented by only 21 travellers, which is below the common 30 respondents for reliable results. Also no socio-demographic characteristic is significant for this small group, meaning class membership cannot be explained by personal factors. It is concluded that the results of a small group are not very representative. When constructing four classes, this same event occurs and two very small groups (8% and 10%) are defined. The found weights for three and four classes are shown in Appendix O.

| Number of classes | LL | BIC | $r^2$ | k |
|---|---|---|---|---|
| 1 (MNL) | -1956 | 3940 | 0.09 | 5 |
| 2 | -1703 | 3469 | 0.20 | 11 |
| **3** | **-1685** | **3467** | **0.22** | **17** |
| 4 | -1668 | 3468 | 0.23 | 23 |

Table 7.3: BIC values of multiple class models

When constructing two classes the results are reasonable. All values of weights are significant, which means the difference between classes is not due to coincidence and can be explained by class membership. Also the socio-demographic variables are modelled in the latent class model, to discover if differences between characteristics of travellers can explain the user groups.

### 7.2.2 Factor Weights and Class Description

When applying the two class choice model on the total data set. Two main classes (groups) with a difference in parameter values are created. The various factor weights are presented per group in table 7.4a and 7.4b. Also, the corresponding p-values are shown. When this p-value is lower than 0.05 it means the factor weight is significant for that class.

Table 7.4: Final Results Latent Class

(a) Group 1 Parameters

| Parameter | Weight | SE | t-test | p-value |
|---|---|---|---|---|
| $\beta_{Light}$ | 0.765 | 0.0752 | 10.17 | 0.00 |
| $\beta_{Shelter}$ | 0.282 | 0.0592 | 4.77 | 0.00 |
| $\beta_{Security}$ | 0.610 | 0.0699 | 8.73 | 0.00 |
| $\beta_{Time}$ | -0.107 | 0.0154 | -6.95 | 0.00 |
| $\beta_{Costs}$ | -0.258 | 0.0951 | -2.71 | 0.01 |

(b) Group 2 Parameters

| Parameter | Weight | SE | t-test | p-value |
|---|---|---|---|---|
| $\beta_{Light}$ | 0.588 | 0.1429 | 4.12 | 0.00 |
| $\beta_{Shelter}$ | 0.738 | 0.1507 | 4.89 | 0.00 |
| $\beta_{Security}$ | 0.532 | 0.1458 | 3.65 | 0.00 |
| $\beta_{Time}$ | -0.272 | 0.0422 | -6.45 | 0.00 |
| $\beta_{Costs}$ | -3.411 | 0.3083 | -11.06 | 0.00 |

The weights from table 7.4a and 7.4b are combined with the attribute levels to see the actual influence of the factors, shown in table 7.5a and 7.5b. A resulting conclusion of these tables is that for group 1, the satisfiers are approximately around the same magnitude as dissatisfiers. Group 2 however, is much very sensitive to the dissatisfiers to costs and time, as opposed to the satisfier factors. This is concluded by the range of weights of the parameters.

Table 7.5: Minimum and maximum weights of Parameters

(a) Group 1 Parameters

| Parameter | Weight | Minimum | Maximum |
|---|---|---|---|
| $\beta_{Light}$ | 0.765 | 0 | 0.765 |
| $\beta_{Shelter}$ | 0.282 | 0 | 0.282 |
| $\beta_{Security}$ | 0.610 | 0 | 0.610 |
| $\beta_{Time}$ | -0.107 | -1.07 | -0.545 |
| $\beta_{Costs}$ | -0.258 | -0.258 | 0 |

(b) Group 2 Parameters

| Parameter | Weight | Minimum | Maximum |
|---|---|---|---|
| $\beta_{Light}$ | 0.588 | 0 | 0.588 |
| $\beta_{Shelter}$ | 0.738 | 0 | 0.738 |
| $\beta_{Security}$ | 0.532 | 0 | 0.532 |
| $\beta_{Time}$ | -0.272 | -2.72 | -0.750 |
| $\beta_{Costs}$ | -3.411 | -3.411 | 0 |

Now that the factor weights per class are known, it is interesting to see if these classes can be connected to user characteristics. In table 7.6 socio-demographic factors are coupled for the second group. All the socio-demographic characteristics are modelled in the latent class model. Only the socio-demographic factors that are significant for at the second group are included in this table. Which means the user characteristics gender, age and social participation is are used to create the two classes. Since group 2 is compared to group 1, all parameter values are 0 for group 1.

In table 7.6 the values represent the probability of a higher factor value belonging to that group. The delta parameter is the class membership parameter. A positive value means more respondents belong to the second group. Initially, there was no parameter called students. However, it is discovered that social participation did not always explain class membership for every function. However, the social participation for student was always significant. That is why social participation is now translated into students and non-students. The parameter student is 0.978 for the second group. Meaning students are more likely to belong in the second group. Age, which is defined as 65+, has a negative relation for group 2. This means elderly are more likely to belong in the first group. The parameter gender is coded in such a way that a positive relation means the chances of a male belonging to that group are high. Thus males are more likely to belong in the second group.

| Parameter | Value | SE | t-test | p-value |
|---|---|---|---|---|
| $Delta-2$ | 0.623 | 0.8544 | 0.73 | 0.47 |
| $Student-2$ | 0.978 | 0.4757 | 2.04 | 0.04 |
| $Age-2$ | -0.602 | 0.2231 | -2.70 | 0.01 |
| $Gender-2$ | 0.969 | 0.3081 | 3.18 | 0.00 |

Table 7.6: Socio-demographics of the second group

The values from table 7.6 can be calculated into percentages for analysing the results. This is done using formula 7.3. Where the class membership is calculated using the exponential of the observed utility of station i divided by the exponential of the observed utilities of all alternatives summed. Resulting in table 7.7.

$$P(Y=i)=\frac{e^{V_i}}{\sum e^{V_j}} \tag{7.3}$$

| | **Group 1** | **Group 2** |
|---|---|---|
| Class membership probability | 34.9% | 65.1% |
| Male | 16.8% | 83.2% |
| Female | 58.8% | 41.2% |
| Student | 16.9% | 83.1% |
| Non-student | 58.6% | 41.4% |
| 65+ | 49.5% | 50.5% |
| 64- | 22.7% | 77.3% |

Table 7.7: Class membership probabilities related to socio-demographic characteristics

**Group 1: Satisfier Sensitive**
The first group would mainly consist of females, elderly and non-students. Even though these characteristics seem to be quite likely to belong to both groups equally, it should be considered that the first group is smaller than the second group. Leading to speculation that the first group would mainly consist of females, elderly and non-students. However, the exact distributions can only be calculated when socio-demographics of bicycle-train travellers of a station are known. Students and males have a probability of 17% to belong in group 1.
The parameter weights are computed according to the variation of attribute levels and their relation to the changes in utility. For the first group this is shown in figure 7.4a. Here it can be noted that 10 minutes of cycling time reduces the utility by -1.07. However, this utility can be increased by 0.765 and 0.610 when adding lighting in the station or security in the bicycle parking.

To be able to compare the resulting weights, the interrelation ratios for group 1 between factors with access time and parking costs as a base are shown in figure 7.2. The values in this figure are in the unit euros figure 7.2a and minutes 7.2b. For the first group found in the latent class analysis, increasing the lighting at the station from low to high is worth €0.61 euros. The ratio of lighting and access time is -7 minutes. This means that increasing the lighting at the station is worth 7 minutes of extra access time for respondents in group 1. Whereas increasing shelter at the station is worth €0.33 euros of parking costs for the first group, meaning 2.5 minutes of extra access time. The Value-of-Time for the first group is equal to €0.41 euro per minute.

**Group 2: Dissatisfier Sensitive**
The second group is the largest class of the two, with an overall probability of 65% to belong to this group. It is interesting that this probability is much higher when the respondent is a student or male, with 83%. Females and non-students are a bit less likely to belong in group 2 opposed to group 1.

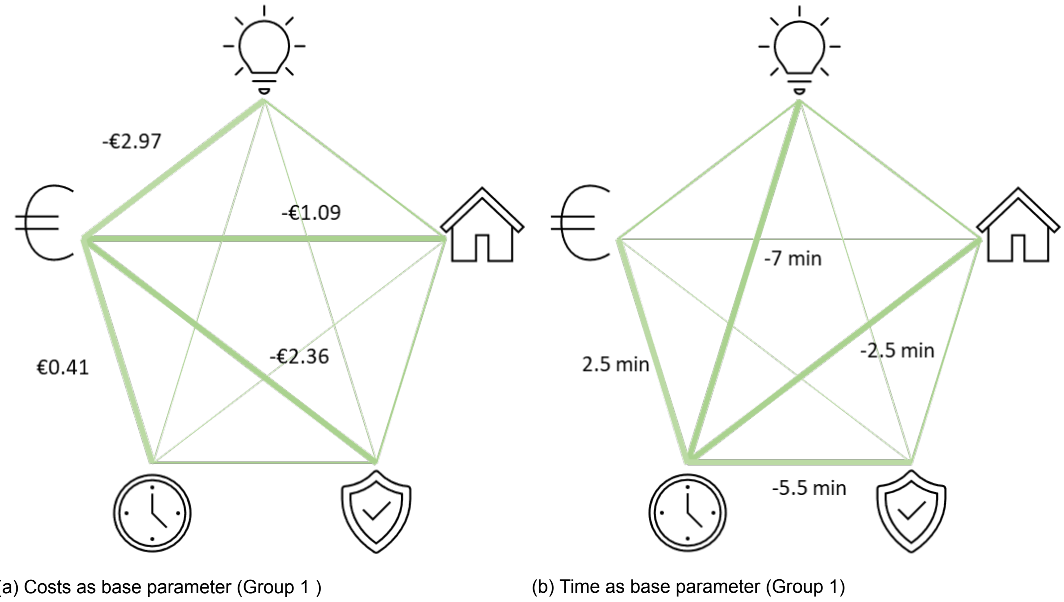



Figure 7.2: Pentagon Interrelation Costs and Time as base for Group 1

However, with 59% for group 1 and 41% for group 2, these distributions are more equal than for students and males among the classes.
The changes in utility for all attribute levels are shown in figure 7.4b. Here it is visible that 10 minutes cycling time decreases the utility of that station by 2.72. Also, parking costs decrease the total utility. A parking price of €1 decreases the total utility of that station by 3.41. The presence of lighting, shelter and security could increase the utility by 0.588, 0.738 and 0.532.

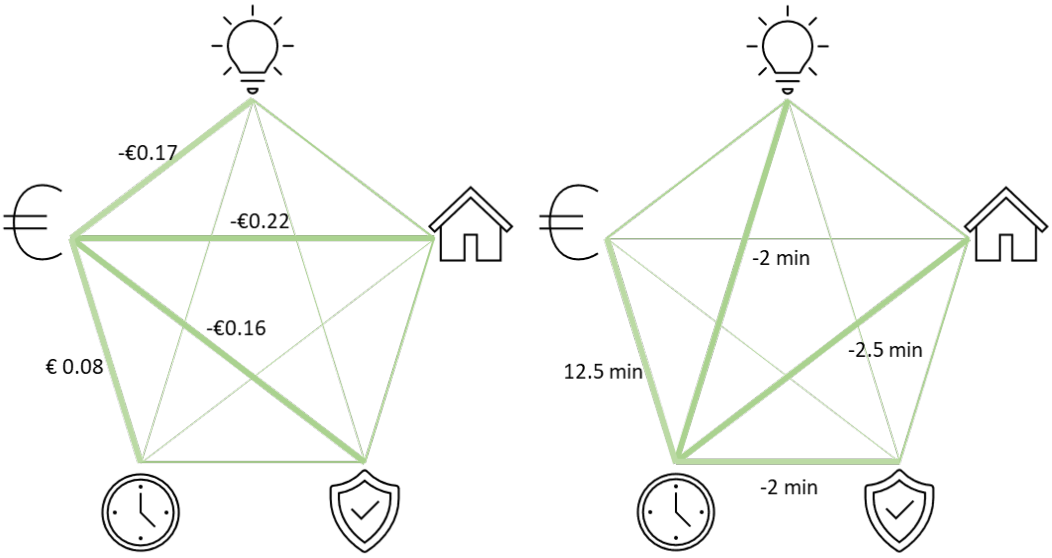
In figure 7.3 the interrelation ratios of group 2 for all parameters with the dissatisfiers are shown. Figure 7.3 shows that for the second group, increasing lighting at the station is worth €0.17 euros or equal to 2 minutes of access time. This is much less than group 1. Increasing the shelter at the station is worth €0.22 euros or 2.5 minutes of access time. Which is approximately the same as group 1. The Value-of-Time of the second group is €0.08 euro per minute, which is in line with the costs-sensitivity of this group. Leading to the conclusion that the second group is less sensitive for presence of light and security, and more sensitive to shelter, cycling time and costs, compared to the first group. This means, in the second group, the effect of the satisfiers on the total utility is much smaller than the effect of the dissatisfiers.

The percentages in table 7.7 are class membership probabilities. Thus it is not excluded that a male belongs in the first group, however, the probability is much lower. The probability of a male student to belong in the second group is higher than that of a female student.

## 7.3 Conclusion Final Survey

The stated choice experiment, final survey and analysing the results in various models are executed with the purpose to find an answer to the third subquestion: "Which relation could be determined between these factors, also including access travel time and parking costs, and to what extent do the weights of these factors differ between several user classes?".

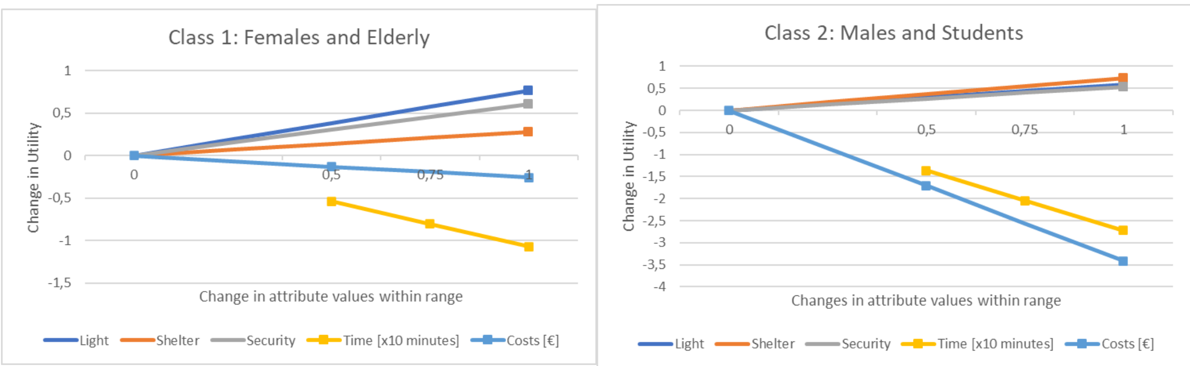
In table 7.8 the relation between factors is given of the travellers using bicycle - train using an MNL model. Since parameter weights cannot be directly compared between different models all weights are divided by the corresponding cost parameter of that model, leading to the amount of euros increasing each variable is worth. From table 7.8 it can be concluded that dissatisfiers have the most influence on station choice in the MNL model. To allow more heterogeneity between the travellers using bicycle and train, also the results are analysed using a Latent Class Choice Model. Two




(a) Costs as base parameter (Group 2)

(b) Time as base parameter (Group 2)

Figure 7.3: Pentagon Interrelation Costs and Time as base for Group 2




(a) Group 1: Changes in utilities for different attribute levels

(b) Group 2: Changes in utilities for different attribute levels

Figure 7.4: Sensitivity to parameters of both classes

distinctive classes are found. Where females, elderly and non-students have a probability of 59% to belong in the first group, and males and students have a probability of 83% to belong in the second group. A summary of both classes is visualised in figure 7.5.

| Parameter / $\beta_{Costs}$ | Weight (MNL) | Weight (Group 1) | Weight (Group 2) |
|---|---|---|---|
| $\beta_{Light}$ | -€0.61 | -€2.97 | -€0.17 |
| $\beta_{Shelter}$ | -€0.33 | -€1.09 | -€0.22 |
| $\beta_{Security}$ | -€0.51 | -€2.36 | -€0.16 |
| $\beta_{Time}$ | €0.11 | €0.41 | €0.08 |
| $\beta_{Costs}$ | -1.09 | -0.258 | -3.411 |

Table 7.8: Final weights of MNL and Latent Class model with costs as base

Table 7.8 also shows the weights of both classes for the five parameters. Here it is clearly visible that the second group is a lot more sensitive to costs compared to the first group. Increasing the satisfiers is worth much more decrease in dissatisfiers to the first group compared to the second group. This means measures influencing the satisfiers especially affect the station choice of travellers within group 1.

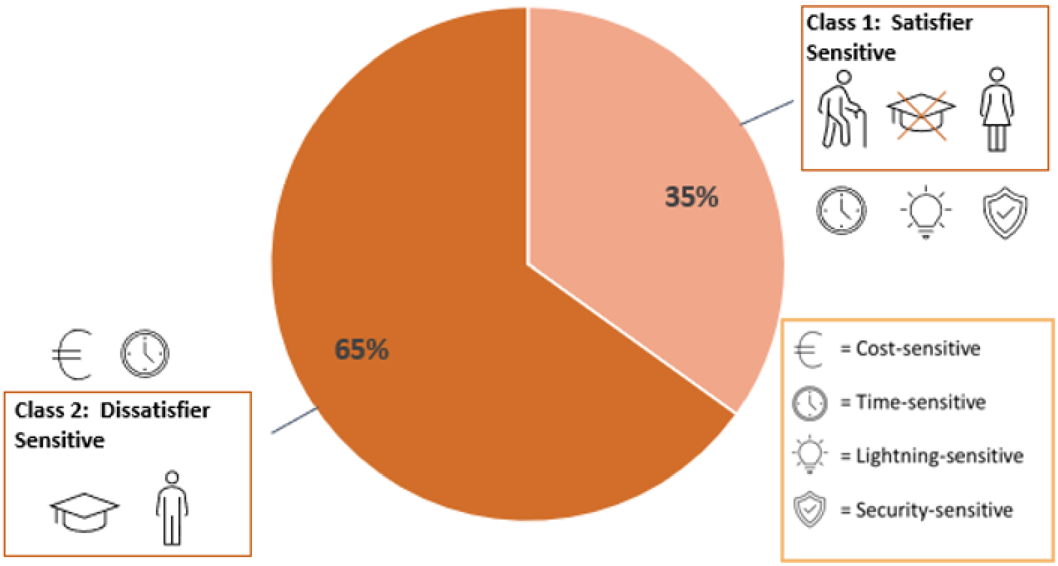



Figure 7.5: Summary of two classes and their characteristics

# 8
# Results in Practice

In this chapter a transition from theory to practice is described, using an illustrative case of Delft. Possible measures based on the researched factors are listed, and the recent renewal of Delft Campus is evaluated.

## 8.1 Possible Measures

Possible measures can be implemented to influence station choice for a local train station. These measures are categorised by the five parameters used in this research: level of lighting at the station, level of shelter at the station, security in the bicycle parking, access travel time and parking price.

**Lighting at station**
For the first group, increasing lighting at the station (and security in the bicycle parking) increases the total utility function almost enough that the dis-utility of extra parking costs and cycle time could even out. Also, for the full inclusiveness of both genders and elderly, it is necessary to provide enough lighting at all stations since the absence of lighting rejects many females and elderly.

**Shelter at station**
Introducing a shelter cubicle is most beneficial to attract the second group (men and students). Adding shelter could unfortunately not compensate for much increase in access time and parking costs. However, when parking costs are increased at the central station, and the local station is provided with a high-level shelter, the utility of the local station could compensate for some of the dis-utility caused by the longer access time. This solution is thought to be one of the cheaper measures, and few political or technical challenges are foreseen.

**Security in Bicycle Parking**
Next to the lighting at the station, another measure to attract travellers, mainly from group 1, is the presence of security in bicycle parking. Providing a human security guard might be a more expensive option. Security could also be provided by adding bicycle lockers at Delft Campus and promoting camera security. These options are not expected to have high costs and are a one-time investment, equal to adding a shelter cubicle.

**Access Travel Time**
Access travel time proved to be a sensitive parameter for station choice for both classes. Leading to the assumption that implementing a possible measure decreasing the access travel time would be most efficient for the total number of travellers. This could be done using different measures. For example, by introducing bicycle speed lanes (”snelfietsroute”) where the bicycle lane is smooth and not many intersections are present. Or to adapt the cycling infrastructure by avoiding same level crossings. For example, by constructing tunnels or bridges. However, both options are quite expensive. Cost-wise but also technical and political challenges are included when adapting an infrastructure. Maybe a more adaptable solution could be promoting and implementing a green wave

for cyclists at traffic lights. But for all options regarding access travel time, much research is to be done for route choice regarding train stations.
Another downside of improving the cycling infrastructure to a local station, also the travel time for locations near the smaller station to the central station might be improved. This means improving bicycle infrastructure could eventually lead not only to more travellers for the local station but also for changing the station choice from local to central.

**Parking Price**
Parking price can be reduced for stations that can adopt some capacity. Especially the second group, which proved to be very sensitive to parking price, would respond to a measure related to the pricing. However, urban stations often already have a free parking system. Therefore also a push measure at the central station is possible. Increasing the parking price at the central station might lead to more travellers choosing the local station.

This measure also has a downside. By increasing the bicycle parking price, one might find more travellers parking their bicycle nearby in places not intended for bicycle parking. There should be clear agreements made between the municipality and station owners regarding the approach for misplaced parking of bicycles. Another negative effect is the possible mode change of people now going by bike and train, choosing a different mode that is less sustainable. Also, the overall traveller satisfaction might decrease when increasing a parking price at all central stations.

In figure 8.1 the possible measures and their estimated effort to implement and effect on station choice for both classes are given. The effect on station utility is the ratio of resulting weights from the Latent Class model. The effort is estimated using the research of van Mil et al. (2020), who interviewed several experts.

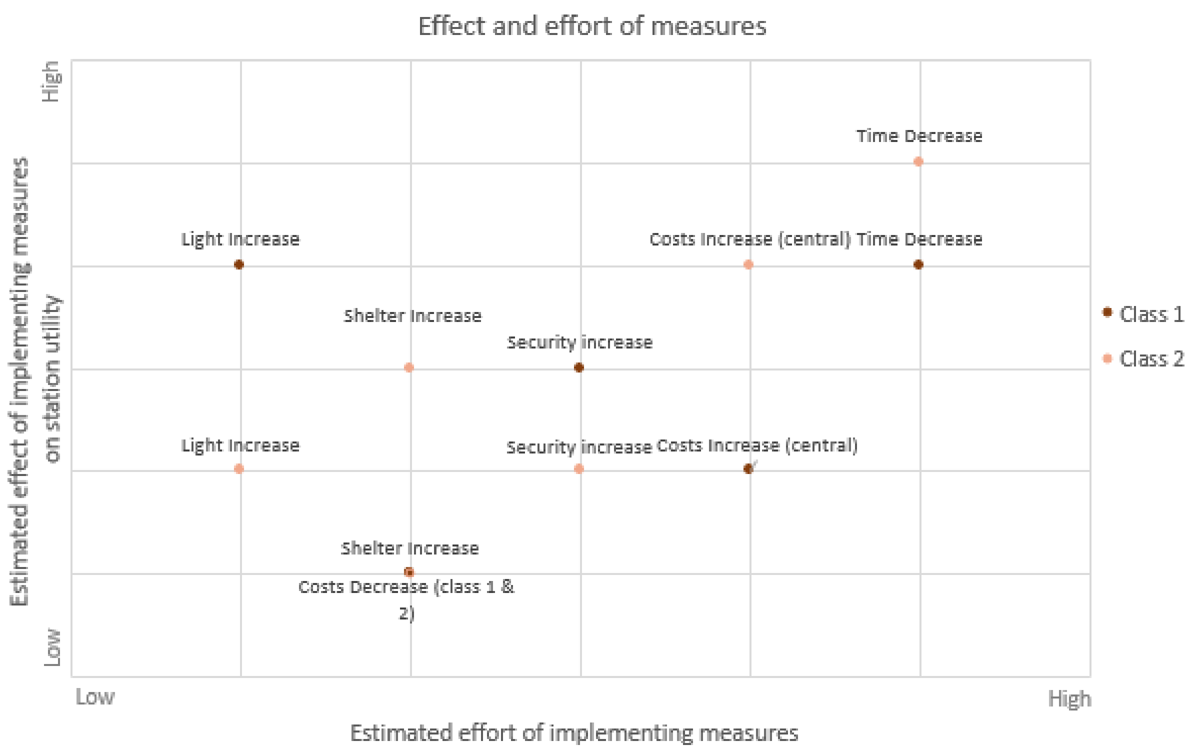



Figure 8.1: Rating of all measures for both classes

## 8.2 Illustrative Case Delft

To illustrate the resulting weights from chapter 7, the case of station choice in Delft is chosen. The resulting weights are suitable to apply to this city since there are two stations present. One belongs to the category of central station and one local station that is smaller, less luxe and is located on the city's outskirts. The city Delft also matches the respondents of this study because a university is present. In terms of the geographical location of the respondents (appendix M.2) Delft is sufficiently represented.

In December 2021, a renewed Delft Campus (previously Delft Zuid) was affected to attract more travellers. As a result, this station is the first energy-neutral train station in The Netherlands. In this illustrative case, the measures implemented by renewing Delft Campus and the found measures in this research are evaluated. Examples of adaptations of the renewed station are the modernized station entrance and increased lighting.

### 8.2.1 Inhabitants Delft

To connect the resulting weights of the two classes found in section 7.2, information is needed about the inhabitants of Delft.

| Factor | # Inhabitants | Percentage |
|---|---|---|
| Delft (total) | 104533 | 100% |
| Male | 55159 | 52% |
| Female | 49374 | 47% |
| Student | 18807 | 18% |
| Non-Student | 85726 | 82% |
| 65+ | 16767 | 16% |
| 64- | 87766 | 84% |

Table 8.1: Socio demographics of Delft (CBS, 2021; CBS, 2022)

When the socio-demographic characteristics of Delft are combined with the class membership functions used in the Latent Class Model (repeated in table 8.2), a new distribution is found of 39963 (38%) inhabitants belonging to group 1 and 64570 (62%) belonging to group 2. This is comparable to the initial 35% and 65% class membership probability.

| | **Group 1** | **Group 2** |
|---|---|---|
| Class membership probability | 34.9% | 65.1% |
| Male | 16.8% | 83.2% |
| Female | 58.8% | 41.2% |
| Student | 16.9% | 83.1% |
| Non-student | 58.6% | 41.4% |
| 65+ | 49.5% | 50.5% |
| 64- | 22.7% | 77.3% |

Table 8.2: Class membership probabilities related to socio-demographic characteristics

### 8.2.2 Delft and Delft Campus

In Delft there are two stations: Delft (central) and Delft Campus (local). On a working day in 2022, station Delft (central) has 15 intercity trains and 8 local trains per hour and approximately 40435 travellers per day in 2019 (pre-Covid). Station Delft campus (local) has 0 intercity trains and 8 local trains per hour and 2080 travellers per day in 2019 (Belderbos, 2021). The characteristics of both stations regarding this research are shown in table 8.3. Where 0 represents low and 1 high. Access time is not included since this is dependent on the origin of the traveller.

| Parameter | Delft | Delft Campus |
|---|---|---|
| Light | 1 | 1 |
| Shelter | 1 | 0 |
| Security | 1 | 0 |
| Parking Price | 0 | 0 |

Table 8.3: Characteristics of Delft and Delft Campus (based on www.stations.nl by NS Stations and ProRail)

An increase in lighting has a positive influence on both classes, but especially on the first group. When sufficient lighting is present, this could almost weigh up to 7 minutes of extra access time (section 7.2). However, this is 7 minutes extra compared to a station with insufficient lighting, which is not the case for Delft station.
Delft Campus did not include an extensive shelter option. This is unfortunate since this is the satisfier with the strongest influence of group 2. This class is highly influenced by the dis-utilities paired with extra travel time, which might be needed when cycling to Delft Campus compared to the central station of Delft.
After the renewal of Delft Campus no security is present in the bicycle parking. This is understandable due to the costs of security in bicycle parking. However, this also means missing out on the benefits of increasing the utility, mainly in group 1. Another possible measure could be introducing bicycle lockers, which travellers could unlock with their OV-chip cards.

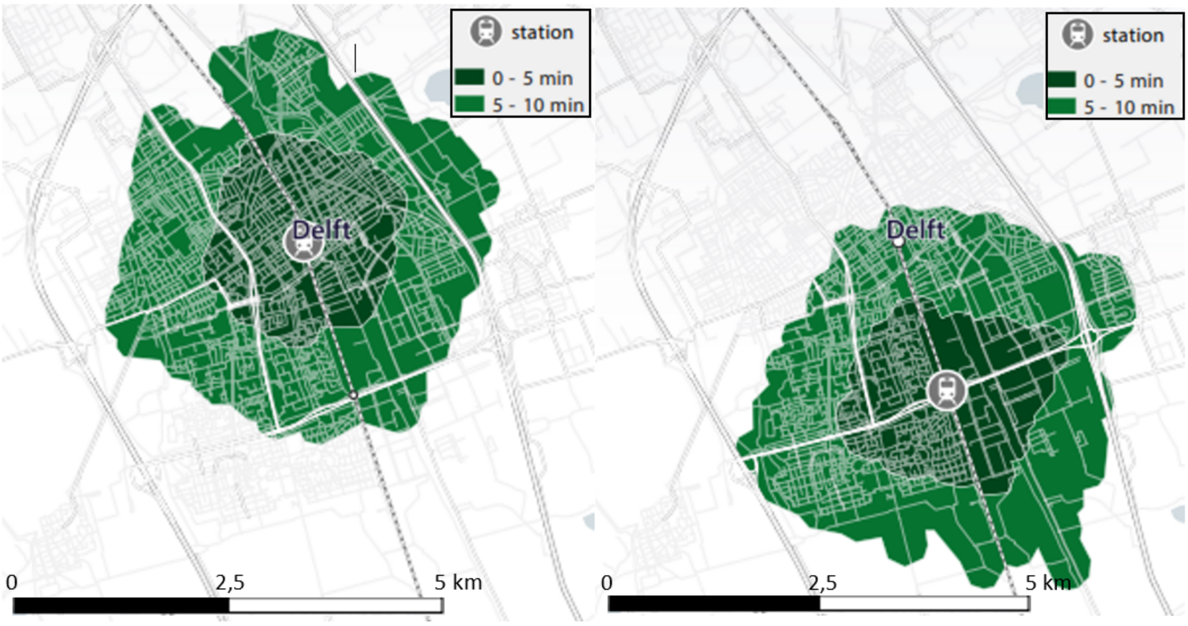
Since station Delft is located at a highly dense location, for many travellers it has less access time by bike. But for travellers coming from the south of the station, the difference between access times might not be as high. In figures 8.2a and 8.2b, both stations are indicated with their radius of access travel time by bike. As is visible, several parts of the radius overlap. Delft Campus did improve the accessibility of the station by constructing a bicycle bridge. Meaning no crossings are needed on the same level as other modes. This improves the access time of the station. Access time could also be improved by implementing the right-of-way for cyclists for a longer part of their route and by adding priority for cyclists at traffic lights. Also, the route from the centre of Delft to Delft Campus could be improved by introducing a fast bike path (”snelfietsroute”), which consists of fewer traffic crossings and a smooth cycling lane. Despite the many benefits this measure may have, it is also extremely costly to adapt the infrastructure. On top of the costs, it might also be a political challenge and technically impossible to implement a bridge or tunnel to prevent same level crossings.

In both station Delft and station Delft Campus the bicycle parking has no fees. This means no extra benefit in utility is won by free parking at Delft Campus. Using a so-called push measure is increasing the parking price at station Delft. This will possibly skim travellers at the central station of Delft to Delft Campus, mainly from the second group. However, a possible negative effect could be some travellers that are currently choosing the mode bicycle + train, who will change to a less sustainable mode once a parking fee is introduced.

## 8.3 Conclusion

All in all, the renewal of Delft Campus included several improvements compared to Delft Zuid to possibly attract more travellers. The largest is the modern station entrance and access routes for bicycle and walking near the station, which does not require same level crossings with the train. Also, lighting of the station was improved, which is important to make the station more attractive, especially for travellers in group 1. However, there are still a few recommendations for Delft Campus to be able to skim more travellers from the central station.

For females and elderly (group 1) also security in the bicycle parking and access time is important.




(a) Station Delft: Access time by bicycle

(b) Station Delft Campus: Access time by bicycle

Figure 8.2: Accessibility of Delft and Delft Campus

The security in the bicycle parking could be solved by adding human security, which is paired with high costs. Another option could be introducing lockers for bicycles which can be opened by an OV-chip card. For access travel time, the new bicycle tunnel for the station entrance has decreased it already. However, improvements can still be made regarding accessibility throughout the route from the centre of Delft to Delft Campus. For example, by introducing fast cycle lanes, minimizing same level crossings or introducing priority for cyclists at traffic lights.
Reducing the cycling access time is even more effective in increasing the utility of Delft Campus for males and students (group 2). This group proves to be very sensitive to cycle time. The most important satisfier for group 2 is shelter. The shelter is not increased in the renewal of Delft Campus. This is unfortunate since the largest group of travellers (65%) is positively sensitive to this parameter. This could increase the utility and therefore partly outweigh the dis-utility caused by the extra access time. Also, the implementation of adding a shelter booth would have rather low costs compared to adjusting the whole cycling infrastructure. Another highly effective measure would be increasing the parking fee at the central station of Delft. This has a risk of mis-parked bicycles at the central station or a mode shift from train to other modes. The advantage of increasing parking costs at Delft is expected to outweighs several minutes of extra access time for Delft Campus.

For example, when a traveller (group 2) lives 5 cycling minutes from station Delft and 10 minutes from Delft Campus, adding a parking fee of €0.50 per day at station Delft and adding extra shelter at Delft Campus. These circumstances would lead to the utilities for group 2 in equations 8.1 and 8.2. Concluding that 62% of bicycle-train travellers (group 2 in Delft) has a probability of 55% for station Delft and 45% for Delft Campus, using equation 7.3.

$$U_{Delft} = 0.588 * 1 + 0.738 * 1 + 0.532 * 1 - 0.272 * 5 + -3.411 * 0.5 = -1.21 \tag{8.1}$$

$$U_{Campus} = 0.588 * 1 + 0.738 * 1 + 0.532 * 0 - 0.272 * 10 + -3.411 * 0 = -1.39 \tag{8.2}$$

When a traveller (group 1) lives 5 and 10 minutes respectively from Delft and Delft Campus, and shelter and security are added at Delft Campus. This would lead to the utilities in equations 8.3 (Delft) and 8.4 (Delft Campus).

$$U_{Delft} = 0.765 * 1 + 0.282 * 1 + 0.610 * 1 - 0.107 * 5 - 0.258 * 0 = 1.12 \tag{8.3}$$

$$U_{Campus} = 0.765 * 1 + 0.282 * 0 + 0.610 * 1 - 0.107 * 10 - 0.258 * 0 = 0.305 \tag{8.4}$$

Resulting in the probability that 63% of 38% of train travellers will choose Delft and 37% will choose Delft Campus. This is compared to a share of 19% for Delft Campus when low shelter or no security is present. Adding shelter and security at Delft Campus and introducing a parking fee at station Delft leads to a share of 52% of travellers choosing Delft Campus. 48% of all travellers using bicycle and train will still choose station Delft.

An important disclaimer that needs to be made is that other characteristics of the train trip are not included in this analysis. The effect of a possible transfer or an intercity versus a local train is not calculated. This means that these results are only useful for the 8 local train trips which stop at both stations of Delft. When a transfer at station Delft is needed for the rest of the train trip, the hypothesis is that most travellers do not start their trip from station Delft Campus. Which makes these examples an overestimation of travellers using Delft Campus.

# 9 Conclusion, Discussion and Recommendations

The final chapter of this research provides the final conclusion, discussion and recommendations. In section 9.1 an answer is given to the main research question. These conclusions are drawn within the limitations of this study. In section 9.2 various limitations of this research are reflected on, leading to several recommendations in section 9.3.

## 9.1 Conclusion

This research investigated the station choice by cyclists concerning satisfiers, as generated by Hagen and Bron (2014). The research question of this study is: ”To what extent can satisfiers, as opposed to dissatisfiers, play a role in station choice for bicycle-train travellers regarding local railway stations?”. To answer the main research question, subquestions were constructed. This chapter first summarises the answers to the subquestions in order to eventually answer the main research question in section 9.1.4.

Four subquestions are constructed at the beginning of this research. The first two concern the identification of factors and the ranking of satisfiers. Followed by computing the weights of factors using an MNL and LC model. The fourth research question relates the findings from this research to practice.

### 9.1.1 Identification and ranking of factors

Regarding the first subquestion, thirty possible satisfier characteristics influencing station choice by bicycle are found in a literature study. For example, factors such as ’intensity of lighting’, ’crowdedness’ and ’neighbourhood reputation’ were included. In the second subquestion, all thirty satisfier factors are ranked by 21 respondents, three respondents representing each user class of the bicycle-train combination found by Shelat et al. in 2018 (table 9.1).

| Class | Description |
|---:|---|
| 1 | Middle-aged full-time professionals (26%) |
| 2 | University students living with parents (23%) |
| 3 | School children (15%) |
| 4 | Young, low income professionals (14%) |
| 5 | Middle-aged part-time professionals (10%) |
| 6 | University students living alone (9%) |
| 7 | Pensioners (2%) |

Table 9.1: Seven classes of bicycle-train travellers (Shelat et al., 2018)

Weighing the ranking results of the seven user classes to the dynamics of society led to the following overall top 3 ranking:
1. Level of shelter at the station
2. Level of lighting at the station
3. Presence of security in the bicycle station

When analysing the ranking results, the satisfiers are categorised into satisfiers influencing comfort or social safety. For example, the level of shelter is a comfort satisfier and the presence of lighting is a social safety satisfier. While the two categories were equally presented for many user groups, pensioners indicated many comfort factors as important. And between the two genders, where females often ranked social safety-related factors, and males were more drawn to the comfort factors. The safety factor 'Amount of other travellers at station' was even ranked by all females (N=10).

After ranking satisfier factors, two dissatisfiers are added to continue in the stated choice experiment. Access travel time and parking costs are included based on their essential role in choice behaviour (Ton et al., 2019; Krabbenborg et al., 2015; Weliwitiya and Eng, 2020; Young and Blainey, 2018; Givoni and Rietveld, 2007).

### 9.1.2 Weights of Factors

The factors resulting from the first and second subquestions are not expected to have an equal influence on station choice. To compute the weights of the five parameters, two models are used. An MNL model is applied first, and to further investigate the weights, a latent class choice model is also applied.

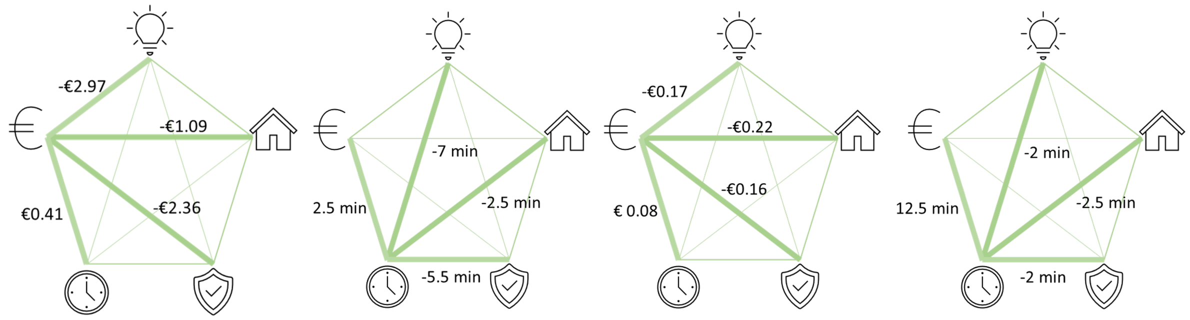
**MNL Model**

After modelling 308 responses on the final survey, where respondents are asked to choose between two stations varying in attribute values, the weights found using an MNL model are all significant at the 0.01 level, meaning the found weights are not due to coincidence. Since little research is done regarding weights of satisfiers in station choice, the found results cannot be compared. Comparing the resulting weights to the ranking of satisfiers, it is remarkable that the level of shelter was the highest ranked satisfier, whereas it has the lowest impact on station utility according to the resulting weights.

The found weights for access travel time and parking costs lead to a Value-of-Time of €0.11 per minute, which is the same as the Value-of-Time found by van Mil et al. in 2020, who researched station choice for cyclists including only dissatisfiers. When analysing each parameter's minimum and maximum values, the utility differences that parameter could cause are known. For example, access travel time varied from 5 to 10 minutes. Considering the range of influence of factors on station utility, the theory leading to the pyramid of customer needs of Hagen and Bron, 2014 proves to be correct. The dissatisfiers access time and parking costs have a higher maximum influence on station utility than the satisfiers. The utility of a station is highly influenced by access time and parking costs. However, the presence of light and security positively impact the utility, enough for each to compensate for 5.5 and 4.5 minutes of extra access time or approximately €0.50 of parking fee.

**Latent Class Model**

To allow heterogeneity between user groups, a Latent Class Model is created to capture the differences in the population. A latent class choice model is used for analysing the results to gain more insights into significant differences between the population of travellers using bicycle-train. When comparing the number of classes in the model, two classes were found. For these two groups, the socio-characteristics 'gender', 'age (65+)' and 'social participation (student')' were found statistically significant and explain class membership. The first group is most likely to consist of females, elderly and non-students. This group is sensitive to access travel time, level of lighting and security. The dissatisfier parking costs has a minimal influence on this first group compared to the MNL model. The second group (65% of the population) is most likely to consist of mainly males and students and is very sensitive to access travel time and travel costs. Both classes and their ratio weights for all parameters with parking costs and access time as a base are shown in Figures 9.1a and 9.1b. It is visible that the first group is willing to pay almost €3 to increase lighting from low to high, whereas the second group is only prepared to pay €0.17. Knowing the weights vary among the two user classes, it can be concluded that different measures are needed to appeal to both user classes to a station.




(a) Group 1 (Costs and Time as base parameter)

(b) Group 2 (Costs and Time as base parameter)

Figure 9.1: Pentagon Interrelation Costs and Time

### 9.1.3 Relation to practice

The final subquestion focused on relating the results of the latent class model to practice. For both classes, measures affecting the actual access travel time would be very efficient for increasing the utility of a local station. However, changing the infrastructure to allow a lower access time might be very complex and face technical and political challenges.

Furthermore, group 1 highly appreciates lighting and security. It is most likely that increasing lighting at the station would be lower in costs compared to changing the infrastructure. And it is expected that 35% of the population would still be influenced by increasing the lighting. Also, group 2 would be positively influenced when lighting is increased. However, their depreciation for extra travel time is much higher than lighting could compensate.

For group 2, the satisfier with the most influence is the level of shelter. This means the disutility caused by extra travel time or parking costs could be (partly) compensated when shelter is increased on the platform. This would also increase the utility for group 1, however not as much compared to lighting and security. Next to access time, parking costs greatly influence station choice for group 2. When parking at both stations is already free, decreasing the parking price for the local station to attract more passengers is not always possible. However, increasing the parking price at the central station might be effective.

The groups resulting from the latent class model are also constructed using the socio-demographic characteristics of the inhabitants of Delft. From the 104533 total inhabitants, group 1 is estimated to have a share of 38%, and group 2 has a share of 62%, based on socio-demographics. This is very close to the initially estimated 35% and 65% of the class membership probability. With recent renovations in Delft Campus, lighting was increased at the station. There is still room for improvement on increasing shelter at the station (appealing to group 2) and providing security in the bicycle parking (appealing to group 1). It is also possible to introduce a parking fee in Delft, which is expected to affect the station choice of travellers (mainly for group 2). Regarding the measure of introducing a parking fee, research is needed on the acceptance of travellers since mode shift is not desirable. When shelter and security are added in Delft Campus, and a parking fee of €0.50 is present in Delft, travellers are almost equally divided over the two stations, with an access time of 5 minutes for station Delft and 10 minutes for Delft Campus.

### 9.1.4 Main Research Question

The answers to the subquestions are used to answer the main research question: "To what extent can satisfiers, as opposed to dissatisfiers, play a role in station choice for bicycle-train travellers regarding local railway stations?".

When comparing the resulting weights of the MNL and Latent Class model, differences can easily be noticed. Meaning that heterogeneity exists in travellers using bicycle-train, and the average values are not always useful. Several different measures could be used for one user class but might not be as effective for the other. Concluding there is much added value in using a latent class model when applying new measures. This is also important to provide inclusiveness for all bicycle–train travellers. For both classes, the satisfiers and dissatisfiers have different ratios to one another (figure 9.1).

The first group (35%), mainly consisting of females and elderly, is mostly influenced by the satisfiers lighting and security and the dissatisfier access time. Access time might be complex to change, such as introducing priority for cyclists or changing the infrastructure. For this user group, the satisfier lighting could compensate for approximately 7 minutes of extra access time. Providing security in the bicycle parking could compensate for the negative utility of 5.5 minutes of extra access time.

It is concluded that satisfiers do play a role in station choice. However, when analysing the utility range, for the largest group (65%), satisfiers have less influence on station choice compared to the first group. According to the model, the dissatisfier shelter, which is the most important satisfier for the second user group, can compensate for 2.5 minutes of access travel time. Providing more lighting and security to satisfy the first group and more shelter for the second user class will also positively influence the station choice of the travellers belonging to another group, however, in a smaller ratio.
Regarding the results and possible measures, it is seen that the factors on the lower levels of the pyramid (dissatisfiers) have a more significant influence on station choice. Therefore, measures regarding the dissatisfier access time are more effective than the satisfiers for both groups. For the second group, the dissatisfier parking costs is the parameter with the most influence on station choice. Concluding that satisfiers do influence the station choice but have less influence than the access time (and parking costs for the second group).
However, when satisfiers are lacking at the local station, and the central station provides high levels of those satisfiers, it is expected that more travellers would prefer the central station over the local station. Leading to the advice for municipalities and station owners to not focus solely on central stations when renewing or designing and keep the local stations up to the standards of the travellers. Parking costs at the central station should be introduced to ensure a shift from travellers from group 2 using the central station to using the local station. However, more research is needed regarding the acceptance of measures and actual choice behaviour.

This conclusion is drawn within the scope of this study. These resulting weights are calculated only using the five factors that are included in this research. Other parameters, such as a possible transfer or the rest of the train trip, are not included in the scope. These results are accurate for train trips that stop at both stations. The results could be inaccurate when a transfer at the central station is needed for pursuing the train trip. Therefore, they could only be used as an indicator of interrelation between various factors. Other limitations are discussed in the following sections.

## 9.2 Discussion

This research shows that satisfiers, next to dissatisfiers, influence station choice. Insights regarding the weights are gathered in this research. Another scientific contribution of this research is the application of a latent class analysis on station choice, showing heterogeneity between travellers using bicycle-train. As introduced in the conclusion, there are certain limitations of this research. In this section, those limitations are reflected on. This discussion consists of two parts: the reflection on methodology and the reflection on the results found in this research.

### 9.2.1 Reflection on Methodology

In this research, more than one method is used. In this section the different methods used in this research are reflected on.

**Interviews and Ranking**

Starting with the interviews used for the ranking of satisfier factors. This was done in a very small group. As many variations of personal characteristics were included. However, it could still easily be the case that this group was not representative for the whole population of bicycle-train travellers due to its small size of N=21. Due to this small number of respondents, the resulting ranking is very sensitive. It was chosen not to elaborate further on this method due to time management since the ranking had an indicative purpose. When more or other respondents would have been included in the ranking process, other factors could have been ranked as most important. This would have led to including different factors in the stated choice experiment and a very different study would have been conducted.

Also, the fourth-ranked factor was the level of lighting during the bicycle trip. Many respondents ranked this factor as important related to station choice. It was concluded that this factor would be very intertwined with route choice, since level of lighting during the bicycle trip could depend on the route to that station and the origin node. This was not included in the scope of the research. Plus, it was the fourth-ranked factor, and the top three would be included in the stated choice experiment. Apart from that, it is still a very valid factor related to station choice.

Finally, concerning the ranking of the factors, only satisfiers were ranked. This was chosen to prioritise the satisfiers in this research. In other research almost only dissatisfiers ended up in the top 5 ranking. This choice was an important addition to this research. However, the eventual model (MNL and Latent Class) could have been more accurate if satisfiers were not prioritised and dissatisfiers were also included in the ranking. On top of that, information could have been lost while categorising factors into dissatisfiers or satisfiers. This division was based on the pyramid of customer needs by Hagen and Bron (2014). Reflecting on the ranking, the rest of this research and results are very dependent on how these factors were labelled in this step (as satisfiers or dissatisfiers).

**Stated Choice Experiment**

After the ranking, three satisfiers (lighting, shelter and security) and two dissatisfiers (access time and parking costs) are included in a stated choice experiment. Because the scope of this research was to only include the current supply of the train network, the presence of a transfer was not included in the experiment. Although this choice was motivated in the scope of the research, it also limits the interpretation of results. Van Mil (2020) found the presence of a transfer has the largest influence on station choice of his chosen parameters. The results of parameters included in both van Mil's and this research are very similar. Meaning that the transfer weight could possibly be extended to this research. The small difference between the resulting weights, for the factors access time and parking costs, of this research and van Mil et al.'s could be accounted further into the weights found for dissatisfiers such as transfer and train time (included in van Mil et al., 2020). A possible disadvantage that should be taken into account is a possible overestimation of the satisfier weights in this research caused by the scope.

Concerning the stated choice survey, also many choices are made. By using stated choice experiments, only hypothetical choice situations are asked. Compared to revealed choice experiments, where actual choices are analysed, in a stated choice it is unknown if respondents would also make the same choice in reality. Indicating that the intentions of respondents are measured instead of real behaviour, it is possible that results are slightly biased regarding the satisfier factors. In the survey, respondents had to choose between two alternatives. It was not asked to the respondents if they would make both or neither trips. Nor was there any other mode choice included in the survey. It was stated at the start that the respondent would travel by bicycle and train. This means only insight into preferences is gathered for travellers using the bicycle and train, but not if they would choose another mode for that trip. Suggestions for further research are made in section 9.3.

Another choice during the stated choice method was the use of images for the parameters lighting and shelter. This was chosen to let the respondents relate to the attributes and show the attribute levels clearly, so less confusion was caused. The use of images in a stated choice also has some disadvantages. For example, in the stated choice, lighting and shelter immediately drew the attention of the respondents. This could have led to overestimating these two factors. Concerning shelter, this weight could also have been overestimated due to the weather circumstances on the image itself. This represents drizzly weather, which is not always the situation in commuter travel. Next to overestimation, using images could also lead to unintended assumptions about the station made by the respondents. This was dealt with by ensuring no station building or many train tracks were visible in the images used, but other overlooked aspects can also have led to unintended assumptions.

**Model Analysis**
A final critical reflection on the methodology is related to the model used. This research used an MNL and Latent Class Choice Model, both assuming a linear additive utility function. The current models do not account for any possible showstoppers. For example, a respondent might never choose a station that has no sufficient lighting. In reality, this might never be compromised by a smaller travel time or sufficient shelter. It is also possible that a traveller would never choose a station further than 10 minutes of extra cycling, no matter how many facilities are provided. In the current model, all factors could be compensated by others, depending on their size, which could not be the case in reality.

### 9.2.2 Reflection on Results

It is desirable that the dissatisfier weight ratio (Value-of-Time) retrieved in this research's MNL model is equal to previous research by van Mil et al. (2020). By sharing the same Value-of-Time, two MNL models can be coupled, which could lead to conclusions regarding the satisfiers in this research, and other dissatisfiers included in van Mil et al.'s research. In addition, the latent class parameter weights indicate the effect of tailored measures on station choice for different socio-characteristics. However, a number of discussion points and limitations should be discussed regarding the retrieved results.

**Adding Factors**
The theory regarding satisfiers and dissatisfiers, and the pyramid of customer needs by Hagen and Bron (2014) is often mentioned in this research. Hagen and Bron found that dissatisfiers such as time, costs and ease form a basis for choice behaviour. All passengers prioritise a safe, reliable, easy and fast journey. These dimensions are the bedrock of transportation. When the journey is evaluated positively, the satisfiers comfort and experience can slightly influence choice behaviour (Hagen and Bron, 2014). The resulting weights in the MNL model support the theory. The weights for costs and time have the largest effect on the utility of a station when applied for the correct range of attribute levels. The satisfiers light, shelter and security also have an effect on the utility of a station, however not as large as the dissatisfiers.
When analysing the results of the latent class choice model, it is discussed that one group behaves according to the previous theory and is very sensitive to the factors access time and parking costs. However, the first group, responsible for 35% of the population, gives less prioritisation to the dissatisfiers. The first group, mainly consisting of elderly and females, is hardly influenced by costs. Access travel time has a larger influence on station choice, however presence of lighting and security could compensate for the dis-utility of extra access travel time to a great extent.

**Relation to Literature**
In the literature research it was stated by Krygsman (2004) that males would accept more travel time to a station. In the results using latent class it was found that males are more negative towards access travel time. This leads to the reflection that the findings by Krygsman might only be accurate when there is no station choice involved. It could also be understood that males are less accepting of extra access travel time compared to the nearest station, and there has to be more compromise on the chosen station than on the factors asked in the stated choice experiment.

As already stated in the reflection on methodology, excluding transfer from this research has large consequences for the interpretation of the results. Transfers are excluded since the public transport schedule is complex to change, and to solely include the current supply of public transport. The resulting weights are valid for train trips that could be taken without transfer on both central and local stations, and for travellers who are indifferent opposed to transfers using public transport. Excluding

transfers leads to the limitation that the weights might be inaccurate when a transfer is related to the station choice of that trip, which are discussed in recommendations for future research.
In the ranking, it was chosen to only provide the respondents with the satisfiers. In other research, these satisfiers often were neglected. To be able to include the satisfiers, satisfiers were prioritised in the focus of this research. Equal to the exclusion of transfers, other dissatisfiers are neglected in the process. When including a ranking of all factors (both satisfiers and dissatisfiers), the final model could be more representative of the station choice of travellers using bicycle as an access mode. The model calculated the same weight of the dissatisfiers compared to other research only including dissatisfiers, enabling a possibility to couple previous research (van Mil et al., 2020) and findings on dissatisfiers. A possible negative effect of coupling research could be that in this research the respondents have to choose between satisfiers. Meaning the influence of satisfiers could be overestimated once the model is coupled with more dissatisfiers. Therefore it would be advised to include at least one satisfier in future research allowing extra validation.

The findings in relation to literature refer to the model of Needs, Opportunities and Abilities of citizens. The results and recommendations are written to estimate the behaviour of travellers when adapting factors at a station. This research includes the ”Needs” and ”Abilities” from the NOA model. However, according to the NOA model, also ”Opportunities” and the interactions in between are needed to fully understand choice behaviour. It might be very difficult to change the behaviour of travellers. If someone travels for a long time using the central station out of habit, how will he adapt his trips when the recommendations are introduced?
It is possible that the theory and model used in this research overestimate the ease of adaptations in choice behaviour. Thus in practice, maybe respondents do have different weights or other factors not included in the model, leading to fewer changes in behaviour when several measures are implemented than expected in the results of this thesis.

**Number of Classes**
In literature it is assumed that costs and time have a negative relation regarding most utility functions. This assumption is argued by Hagen and Bron (2014), who introduced that time perception could be of more importance. Next to group size and significance, the negative relation of time and costs was a motivation for continuing with two classes. To check the accuracy of this assumption, it would be useful to check with the respondents belonging to the first group if they still support their decisions and agree with the weights assigned to their class. This is not possible in this research due to privacy reasons agreed to in the survey, still it would gather much more insight into why the third class model is not in line with the theory of costs and time having a negative relation in choice modelling. When the assumption of a negative relation for time and costs is rejected, this research could have led to more classes.

Shelat et al. (2018) found seven classes with different socio-demographic characteristics in bicycle-train travellers. In this research, those seven classes were taken into account while determining the ranking of factors and retrieving respondents for the pilot and final survey. While determining different classes regarding the factor weights, only two classes were found significant. This means that there is an agreement in station choice between travellers between different classes as found by Shelat2018. For example, the two student classes (living alone and living with parents) of Shelat et al. (2018) both have the same probability of 83% to belong in the group 2 (dissatisfier sensitive).
In the latent class model, characteristics such as income and full-time/part-time employees were not significant for class membership. This significance could be checked when a higher number of respondents is included. While 308 respondents not a small group, when dividing these respondents over several characteristics, the number of respondents per characteristic could be insufficient. Including more respondents could have led to more classes being significant.

**Bicycle - Public Transport**
From the beginning of this research, the mode-combination bicycle-train was focused on. It would be interesting to see if the results of this research also apply to other mode combinations such as bicycle – metro/bus/tram. Since many central stations of these modes are combined with the central train

stations. Even though the bicycle is less used as an access mode for these other public transport modes, it is still responsible for 1 - 25% of all access (van Goeverden and Egeter, 1993; van Uum et al., 1995; Martens, 2004; Rijsman et al., 2019; Brand et al., 2017). An insight could be retrieved when researching how much capacity of the central station parking is occupied by travellers using another mode.

Another reflection on the retrieved results is concerning the ability to extend the weights found in this research to the mode bicycle – metro/tram/bus. When a higher frequency is expected of the modes metro, bus and tram compared to train, one might be less sensitive to the station/stop parameters such as lighting and shelter. Security in bicycle parking is often absent at local stations of metro, tram and bus. However, the weight might be similar to the one found in this research regarding train. Also, the weight for access time is not expected to change for the modes metro/tram/bus compared to train. The sensitivity to costs could differ since various social characteristics determine the willingness to pay. Therefore, to calculate the possible change in weights (for example costs), the user population of the other modes should be compared with the population used in this research.
Also related to the ability to generalise this research is the possibility of using this research in other countries. This research was applied to The Netherlands due to its developed bicycle-train system. When extending the research to other countries, the weights of most parameters are expected to have a similar magnitude. However, due to less advanced bicycle infrastructure in some other countries, travellers might be more sensitive to access travel time and traffic safety.

**Mode Choice**
Finding different factor weights for two classes could also be used in more domains other than station choice, for example regarding mode choice. In this research, the respondents are representative of current bicycle-train travellers. When investigating mode choice for all travellers (including car), it is expected that directness and accessibility matter more. However, the two classes found in this research could still be valid. For a large part of the population, costs is a deciding factor. So to stimulate the use of sustainable travel modes, and to make public transport more attractive, pricing should be more competitive between public transport and private modes.

**Corona Impact**
This research did not explicitly include the impact of Covid-19. According to de Haas and Faber (2022) it is expected that when all measures for Covid-19 are cancelled, the attitude of travellers towards public transport will be restored fast. This could lead to less crowded peak hours in public transport. However, regarding the bicycle parking capacity problems, the exact arrival and departure time does not matter as much as the time span the bicycle occupies a parking place. This indicates bicycle parking to be less related to peak hours and would still have capacity issues when travel times of passengers are distributed. As an effect of Covid-19, changes are expected in travel behaviour caused by the trend of working from home (de Haas and Faber, 2022). This could lead to less expected growth of passengers, and therefore, the expected capacity issues might be less severe than pre-Covid. However, still growth of passengers is expected in the near future, and it would still be advised to anticipate early.

## 9.3 Recommendations

There are two types of recommendations in this section, recommendations for further research, section 9.3.1, and recommendations regarding policy and design, section 9.3.2. These sections are based on the conclusions and limitations stated in the previous chapters.

### 9.3.1 Research Recommendations

The limitations regarding the ranking of factors, such as the small number of respondents and the prioritisation, are previously discussed. This could be solved in future research by first increasing the size of the group of respondents. This would allow for more statistical significance and increases the verification of the final model.

Regarding the approach of ranking the factors influencing station choice, there are two main recommendations for further research. First, it would be useful to rank both satisfiers and dissatisfiers in one group and perform a latent class analysis on the top five factors regarding station choice. In this potential focus, there is no prioritisation towards the satisfier factors. This could make the final model more explanatory for station choice in reality. The addition to current research would lie in the Latent Class Choice Model. It is expected that the population of bicycle-train users also includes heterogeneity in the weights of the overall most important factors influencing station choice.

Another research could focus on several satisfiers but also include more aspects on the level of service of public transport. For example train travel time and the dissatisfier factor transfer. Since this is the most sensitive parameter found by van Mil et al. (2020), it would be very interesting to see how this parameter relates to various satisfier factors. To be able to relate both pieces of research, it would be useful to include at least one or two overlapping parameters.

Another recommendation for future research is to use revealed preference choices instead of stated preference. Using revealed preference is more time-consuming, but it increases the validity of the final model since it lists actual choices made and no hypothetical choices. Such revealed preference research could be executed in a city where its central station already has a parking fee, compared to a city where its central station provides free parking. It would also be interesting to compare the deviation of travellers amongst different types of stations in both of these cities. Using revealed preference would also solve the limitation of the unintended assumptions caused by using visual material in the stated choice experiment.

As noted in the discussion, the model used in this thesis could be expanded in future research regarding mode choice or allowing hard constraints. A possible improvement could be adding mode choice. The model created would become a Nested Logit model where two options could be between station choice and one option is another competing mode, for example, car. The alternatives in this model would not be unlabeled, hence the passing for an MNL model and using a Nested Logit. A Nested Logit model not only allows for alternative specific constants, but it also allows for overlapping between alternatives, which is the case when two alternatives require station choice, and the other alternative does not.
Regarding the linear additive utility functions used in the model, future research might also add the effect of possible hard constraints. This could be done by changing the mathematical assumptions and design of the utility functions. This would increase the reliability of the model and the resulting weights for all factors.

Finally, regarding implementations of this model in practice, more research should be done on the acceptance of advised measures. Based on the results of this research, the most effective measure would be increasing the parking price at the central station. However, this policy measure should not be implemented before the risks of this measure are appropriately investigated. It would be very undesirable if increasing the parking price would eventually lead to fewer public transport users or would increase the gap between socio-demographic characteristics in public transport users. In 2021, Hoskam investigated the willingness to pay of bicycle parking users at train stations. This research assumed strict control which discouraged parking the bicycle in public spaces. This is not valid for all station environments. More research is recommended on the level of supervision, willingness to pay for bicycle parking and the combination with station choice.

### 9.3.2 Policy Recommendations

Different stakeholders might be interested in station choice for bicycle-train travellers. Examples of stakeholders are station owners (NS Stations and ProRail), public transport companies (NS, Arriva e.d.), municipalities and the ministry of infrastructure and water management. Each with different levels of interest and power.
For station owners, the interest in station choice is more related to station services and the bicycle parking capacity. Currently, station services are determined by the amount of travellers for that station. A policy recommendation for station owners is to let the station services of a local station depend on the characteristics of another station nearby (if any). This research and illustrated case study show that the characteristics of station B have an impact on the amount of travellers for station

A.
Public transport companies could have interest in more insight into station choice to manage train capacity to maintain the satisfaction of travellers. These companies do not have a direct influence on station choice regarding the factors included in this research, since they cannot change any characteristic of the station directly. However, station choice also influences the needed train capacity. Therefore, for public transport companies, research is recommended on the impact of train types on station choice. Public transport companies are also advised to inform the station owners and municipalities to express their interest in insights into station choice.
A municipality could be interested in station choice regarding the access trip to stations and safety of the station environment. For municipalities it is recommended to discuss problems on station access routes with the other station owners. The municipality and station owners could cooperate in making a Cost-Benefit analysis (CBA) for specific stations and the inhabitants near those stations. A CBA provides a monetary comparison between the benefits of spreading travellers over two stations and the costs of adding extra lighting, shelter or security.
The ministry of infrastructure and water management could steer the station owners and municipalities to include satisfiers in this analysis. Also, this ministry could express the added value of including a latent class analysis to allow for heterogeneity in future policy measures.

Based on the significance of various socio-demographic parameters, it can be concluded that there is heterogeneity within station choice for travellers. When performing this research on a larger respondent group, more significant user classes could be found. An important conclusion drawn from this research is for station owners NS and ProRail, to not only implement measures calculated on the average weight for parameters of the whole population. Implementing class-specific measures could improve the diversity and inclusiveness of travellers using the mode combination bicycle - train. Analysing the results of this research, travellers from group 1 are sensitive to lighting at the station, whereas travellers in group 2 are hardly influenced by this factor. Increasing costs at the bicycle parking of the central station influences station choice of both classes, especially group 2. Next to adding lighting and introducing a parking fee, also providing security (group 1) and shelter (group 2) is expected to be effective.

When solving capacity and bicycle parking problems at the central station, satisfier factors could influence station choice. However, when the central station has a sufficient supply regarding the satisfier factors, no extra utility will be generated for the local stations by adding lighting, shelter or security, compared to the central station. However, when the local stations do not have sufficient lighting, shelter or security, the resistance of travellers using a local station is higher, and the total utility of the central station has a head start. An important policy recommendation would therefore be to not solely focus on central stations while renewing or rebuilding. But also to keep the local stations up to the standards of the travellers.

Based on the resulting weights of factors in this research, travellers in the second group are the most sensitive to introducing a parking fee. Thus when parking pressure is severe at the central station, the most effective change would be introducing a parking fee at the central station and providing free parking at the local station. As is mentioned in the recommendations for future research, this is a sensitive measure. Cooperation between municipalities and station owners is advised to gain more insights regarding the mode change of bicycle-train travellers and the effects of a parking price on misplaced bicycles. Also, the willingness to pay for bicycle parking should be investigated in to not increase a gap regarding socio-demographic characteristics in public transport usage.

Regarding the case study performed in this research, a Cost-Benefit analysis is recommended for station Delft and Delft Campus. Based on the case study in this research, it is recommended to increase shelter at the station and security in the bicycle parking when capacity problems at station Delft increase. Increasing shelter and security is expected to require less effort than adapting access infrastructure, and would still be positively rewarded by travellers. The performed case study stated an access time of 5 minutes for Delft and 10 minutes for Delft Campus. Varying the access time in a CBA allows a more accurate forecast of station choice for travellers. This would result in insightful recommendations regarding the benefits of adding extra shelter or security.

# Bibliography


Adcock, S. J. (1997). A passenger station choice model for the british rail network. *PTRC-PUBLICATIONS-P*, 141–146.

Aleskerov, F., Bouyssou, D., & Monjardet, B. (2007). *Utility maximization, choice and preference* (Second, Vol. 16). Springer Science & Business Media.

Bai, T., Li, X., & Sun, Z. (2017). Effects of cost adjustment on travel mode choice: Analysis and comparison of different logit models [World Conference on Transport Research - WCTR 2016 Shanghai. 10-15 July 2016]. *Transportation Research Procedia*, *25*, 2649–2659. https://doi.org/https://doi.org/10.1016/j.trpro.2017.05.150

Belderbos, C. (2021). Welk station laat sterkste daling zien van ns-reizigers in 2020. *Nederlandse Spoorwegen*.

Bhat, C., Eluru, N., Bhat, C. R., & Copperman, R. B. (2008). *Flexible model structures for discrete choice analysis freight transportation view project household expenditure analysis view project chapter 5: Flexible model structures for discrete choice analysis*. https://www.researchgate.net/publication/228964034

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

Birago, D., Mensah, S. O., & Sharma, S. (2017). Level of service delivery of public transport and mode choice in accra, ghana. *Transportation research part F: traffic psychology and behaviour*, *46*, 284–300.

Börjesson, M., & Eliasson, J. (2012). The value of time and external benefits in bicycle appraisal. *Transportation Research Part A: Policy and Practice*, *46*(4), 673–683. https://doi.org/https://doi.org/10.1016/j.tra.2012.01.006

Bovy, B. H. L., Bliemer, M. C. J., & van Nes, R. (2006). *Transportation modeling*. Delft University of Technology.

Bowman, L. A., & Turnquist, M. A. (1981). Service frequency, schedule reliability and passenger wait times at transit stops. *Transportation Research Part A: General*, *15*(6), 465–471.

Boxall, P. C., & Adamowicz, W. L. (2002). *Understanding heterogeneous preferences in random utility models: A latent class approach*.

Brand, J., Hoogendoorn, S., Van Oort, N., & Schalkwijk, B. (2017). Modelling multimodal transit networks integration of bus networks with walking and cycling. *2017 5th IEEE international conference on models and Technologies for Intelligent Transportation Systems (MT-ITS)*, 750–755.

Brands, T., Romph, E. D., Veitch, T., & Cook, J. (2014). Modelling public transport route choice, with multiple access and egress modes. *Transportation Research Procedia*, *1*, 12–23. https://doi.org/10.1016/j.trpro.2014.07.003

Broach, J., Dill, J., & Gliebe, J. (2012). Where do cyclists ride? a route choice model developed with revealed preference gps data. *Transportation Research Part A: Policy and Practice*, *46*(10), 1730–1740.

Brons, M., Givoni, M., & Rietveld, P. (2009). Access to railway stations and its potential in increasing rail use. *Transportation Research Part A: Policy and Practice*, *43*, 136–149. https://doi.org/10.1016/j.tra.2008.08.002

CBS. (2021). Leerlingen, deelnemers en studenten; onderwijssoort, woonregio [25 June 2021]. *Centraal Bureau voor de Statistiek*. Retrieved May 1, 2022, from https://opendata.cbs.nl/statline/#/CBS/nl/dataset/71450ned/table?ts=1651845346426

CBS. (2022). Regionale kerncijfers nederland [31 March 2022]. *Centraal Bureau voor de Statistiek*. Retrieved May 1, 2022, from https://opendata.cbs.nl/statline/#/CBS/nl/dataset/70072ned/table?ts=1651846753350

Cervero, R., Caldwell, B., & Cuellar, J. (2013). Bike-and-ride: Build it and they will come. *Journal of Public Transportation*, *16*(4), 5.

Chakour, V., & Eluru, N. (2014). Analyzing commuter train user behavior: A decision framework for access mode and station choice. *Transportation*, *41*, 211–228. https://doi.org/10.1007/s11116-013-9509-y

ChoiceMetrics. (2018). *Ngene 1.2 user manual and reference guide the cutting edge in experimental design*. www.choice-metrics.com

Clinch, J. P., & Kelly, J. A. (2004). *Planning and environmental policy research series working papers 2004 temporal variance of revealed preference on-street parking price elasticity*. www.ucd.ie/pepweb

Constructing efficient stated choice experimental designs. (2009). *Transport Reviews*, *29*(5), 587–617. https://doi.org/10.1080/01441640902827623

Cui, Y., Mishra, S., & Welch, T. F. (2014). Land use effects on bicycle ridership: A framework for state planning agencies. *Journal of Transport Geography*, *41*, 220–228.

Currie, G., & Stanley, J. (2008). Investigating links between social capital and public transport. *Transport Reviews*, *28*. https://doi.org/10.1080/01441640701817197

De la Bruheze, A. A. A., & Veraart, F. C. A. (1999). *Fietsverkeer in praktijk en beleid in de twintigste eeuw*.

Debrezion, G., Pels, E., & Rietveld, P. (2009). Modelling the joint access mode and railway station choice. *Transportation Research Part E: Logistics and Transportation Review*, *45*, 270–283. https://doi.org/10.1016/j.tre.2008.07.001

de Haas, M., & Faber, R. (2022). *De relatie tussen attitudes en reisgedrag en het verband met de coronapandemie*. Kennisinstituut voor Mobiliteitsbeleid.

Fan, K.-S., Miller, E. J., & Badoe, D. (1993). Modeling rail access mode and station choice. *Transportation Research Record*, (1413).

Ferrell, C. E., & Mathur, S. (2012). Influences of neighborhood crime on mode choice. *Transportation Research Record*, *2320*, 55–63. https://doi.org/10.3141/2320-07

Fifer, S., Rose, J., & Greaves, S. (2014). Hypothetical bias in stated choice experiments: Is it a problem? and if so, how do we deal with it? *Transportation Research Part A: Policy and Practice*, *61*, 164–177. https://doi.org/10.1016/j.tra.2013.12.010

Fox, J. (2005). Modelling park-and-ride in the prism model for the west midlands region. *European Transport Conference, 2005Association for European Transport (AET)*.

Garcia, A., Gomez, F. A., Llorca, C., & Angel-Domenech, A. (2015). Effect of width and boundary conditions on meeting maneuvers on two-way separated cycle tracks. *Accident Analysis & Prevention*, *78*, 127–137.

Givoni, M., & Rietveld, P. (2007). The access journey to the railway station and its role in passengers' satisfaction with rail travel. *Transport Policy*, *14*, 357–365. https://doi.org/10.1016/j.tranpol.2007.04.004

Goeverden, K. V., & Correia, G. (2018). Potential of peer-to-peer bike sharing for relieving bike parking capacity shortage at train stations: An explorative analysis for the netherlands. *EJTIR Issue*, *18*, 457–474. http://tlo.tbm.tudelft.nl/ejtir

Grisolía, J. M., & Willis, K. G. (2016). Consumer choice of theatrical productions: A combined revealed preference–stated preference approach. *Empirical Economics*, *50*, 933–957. https://doi.org/10.1007/s00181-015-0948-5

Hagen, M. V., & Bron, P. (2014). Enhancing the experience of the train journey: Changing the focus from satisfaction to emotional experience of customers. *Transportation Research Procedia*, *1*, 253–263. https://doi.org/10.1016/j.trpro.2014.07.025

Hasnine, M. S., & Habib, K. N. (2018). What about the dynamics in daily travel mode choices? a dynamic discrete choice approach for tour-based mode choice modelling. *Transport Policy*, *71*, 70–80. https://doi.org/10.1016/j.tranpol.2018.07.011

Hensher, D. A., & Greene, W. H. (2003). The mixed logit model: The state of practice. *Transportation*, *30*(2), 133–176.

Hess, S., Daly, A., & Batley, R. (2018). Revisiting consistency with random utility maximisation: Theory and implications for practical work. *Theory and Decision*, *84*, 181–204. https://doi.org/10.1007/s11238-017-9651-7

Hoskam, S. (2021). *The willingness to pay of various types of bike parking-users at train stations for different types of facilities and stations*. MSc Thesis, TU Delft.

Jacobs, I. (2019). Ns: Spoor nog sneller vol dan gedacht. *OVPro*.

Jonkeren, O., Harms, L., Jorritsma, P., Huibregtse, O., & Bakker, P. (2018). *Waar zouden we zijn zonder de fiets en de trein?* Kennisinstituut voor Mobiliteitsbeleid.

Jonkeren, O., Kager, ·. R., Harms, ·. L., Brömmelstroet, ·. M. T., Brömmelstroet, M. T., & Nl, M. C. G. T. (2021). The bicycle-train travellers in the netherlands: Personal profiles and travel choices. *Transportation*, *48*, 455–476. https://doi.org/10.1007/s11116-019-10061-3

Jonkeren, O., & Kager, R. (2021). Bicycle parking at train stations in the netherlands: Travellers' behaviour and policy options. *Research in Transportation Business and Management*, *40*. https://doi.org/10.1016/j.rtbm.2020.100581

Kager, R., Bertolini, L., & Brömmelstroet, M. T. (2016). Characterisation of and reflections on the synergy of bicycles and public transport. *Transportation Research Part A: Policy and Practice*, *85*, 208–219. https://doi.org/10.1016/j.tra.2016.01.015

Keijer, M. J. N., & Rietveld, P. (2000). How do people get to the railway station? the dutch experience. *Transportation Planning and Technology*, *23*(3), 215–235. https://doi.org/10.1080/03081060008717650

Krabbenborg, L., Wee, B. V., Annema, J. A., Correira, G., & Snellen, D. (2015). *Cycling to a railway station*.

Krenn, P. J., Oja, P., & Titze, S. (2014). Route choices of transport bicyclists: A comparison of actually used and shortest routes. *International journal of behavioral nutrition and physical activity*, *11*(1), 1–7.

Krygsman, S. (2004). *Activity and travel choice (s) in multimodal public transport systems*. Utrecht University.

List, J. A., & Gallet, C. A. (2001). *What experimental protocol influence disparities between actual and hypothetical stated values? evidence from a meta-analysis*.

Liu, Y., Yang, D., Timmermans, H. J., & de Vries, B. (2020). Analysis of the impact of street-scale built environment design near metro stations on pedestrian and cyclist road segment choice: A stated choice experiment. *Journal of transport geography*, *82*, 102570.

Mackenbach, J. D., Randal, E., Zhao, P., & Howden-Chapman, P. (2016). The influence of urban land-use and public transport facilities on active commuting in wellington, new zealand: Active transport forecasting using the wilute model. *Sustainability*, *8*(3), 242.

Martens, K. (2004). The bicycle as a feedering mode: Experiences from three european countries. *Transportation Research Part D: Transport and Environment*, *9*, 281–294. https://doi.org/10.1016/j.trd.2004.02.005

Martens, K. (2006). Promoting bike-and-ride: The dutch experience. *Transportation Research Part A: Policy and Practice*, *41*, 326–338. https://doi.org/10.1016/j.tra.2006.09.010

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

Metz, F. (2014). *Eén reizigerskilometer minder met de auto: Wat levert dat op?* Colloquium Vervoesplanologisch Speurwerk.

Middelkoop, A. (2017). *Sluit ns 's nachts stations af?* https://nieuws.ns.nl/sluit-ns-s-nachts-stations-af/

Molin, E., & Maat, K. (2015). Bicycle parking demand at railway stations: Capturing price-walking trade offs. *Research in Transportation Economics*, *53*, 3–12. https://doi.org/10.1016/j.retrec.2015.10.014

Molin, E. (2019). Introduction to experimental designs [powerpoint - slides]. *Technical, Policy and Management, TU Delft*.

Müller, S., Tscharaktschiew, S., & Haase, K. (2008). Travel-to-school mode choice modelling and patterns of school choice in urban areas. *Journal of Transport Geography*, *16*, 342–357. https://doi.org/10.1016/j.jtrangeo.2007.12.004

Mumford, L. (1955). The roaring traffic boom. *New York*, *31*, 92–95.

Murphy, J. J., Stevens, T., & Weatherhead, D. (2004). *Is cheap talk effective at eliminating hypothetical bias in a provision point mechanism?* http://www.umass.edu/resec/workingpapers

NS. (2019). *Ns annual report*. Nederlandse Spoorwegen.

Ohler, T., Le, A., Louviere, J., & Swait, J. (2000). Attribute range effects in binary response tasks. *Marketing Letters*, *11*(3), 249–260.

PanelClix. (2021). *Panelclix - het grootste panel van nederland*. retrieved%20from%20https://www.panelclix.nl/

Park, S., Kang, J., & Choi, K. (2014). Finding determinants of transit users' walking and biking access trips to the station: A pilot case study. *KSCE Journal of Civil Engineering*, *18*(2), 651–658.
ProRail. (2019). *Fietsparkeren bij stations*. https://www.prorail.nl/reizen/stations/fietsen
Pucher, J., & Dijkstra, L. (2000). *Making walking and cycling safer: Lessons from europe*.
Puello, L. L. P., & Geurs, K. T. (2016). Integration of unobserved effects in generalised transport access costs of cycling to railway stations. *EJTIR Issue*, *16*, 385–405.
Puello, L. L. P., & Geurs, K. (2015). Modelling observed and unobserved factors in cycling to railway stations: Application to transit-oriented-developments in the netherlands. *European journal of transport and infrastructure research*, *15*(1).
Replogle, M. A. (1987). Bicycles on transit: A review of international experience. *Transportation research record*, *1141*, 26–36.
Rietveld, P. (2000a). The accessibility of railway stations: The role of the bicycle in the netherlands. *Transportation Research Part D: Transport and Environment*, *5*(1), 71–75. https://doi.org/https://doi.org/10.1016/S1361-9209(99)00019-X
Rietveld, P. (2000b). *Non-motorised modes in transport systems: A multimodal chain perspective for the netherlands*. www.elsevier.com/locate/trd
Rietveld, P., & Daniel, V. (2004). Determinants of bicycle use: Do municipal policies matter? *Transportation Research Part A: Policy and Practice*, *38*(7), 531–550.
Rijsman, L., van Oort, N., Ton, D., Hoogendoorn, S., Molin, E., & Teijl, T. (2019). Walking and bicycle catchment areas of tram stops: Factors and insights. *2019 6th International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*, 1–5. https://doi.org/10.1109/MTITS.2019.8883361
Sáez, D., Cortés, C. E., Milla, F., Núñez, A., Tirachini, A., & Riquelme, M. (2012). Hybrid predictive control strategy for a public transport system with uncertain demand. *Transportmetrica*, *8*, 61–86. https://doi.org/10.1080/18128601003615535
Scheltema, E. (2012). Recycle city: Strengthening the bikeability from home to the dutch railway station.
Semler, C., & Hale, C. (2010). Rail station access–an assessment of options.
Shelat, S., Huisman, R., & van Oort, N. (2018). Analysing the trip and user characteristics of the combined bicycle and transit mode. *Research in Transportation Economics*, *69*, 68–76. https://doi.org/10.1016/j.retrec.2018.07.017
Shen, J. (2006). A review of stated choice method. *nternational Public Policy Studies*, *10*(2), 97–121.
Shen, J. (2009). Latent class model or mixed logit model? a comparison by transport mode choice data. *Applied Economics*, *41*, 2915–2924. https://doi.org/10.1080/00036840801964633
Sherwin, H., Parkhurst, G., Robbins, D., & Walker, I. (2011). Practices and motivations of travellers making rail–cycle trips. *Proceedings of the Institution of Civil Engineers-Transport*, *164*(3), 189–197.
Stanko, E. A. (2008). The case of fearful women. *Women & Criminal Justice*, *4*(1), 117–135. https://doi.org/10.1300/J012v04n01_06
Stanley, J. K., Hensher, D. A., & Loader, C. (2011). Road transport and climate change: Stepping off the greenhouse gas. *Transportation Research Part A: Policy and Practice*, *45*(10), 1020–1030.
Stinson, M. A., Bhat, C. R. et al. (2005). A comparison of the route preferences of experienced and inexperienced bicycle commuters. *TRB 84th annual meeting compendium of papers*, (05-1434).
Sweet, R. J. (1997). *An aggregate measure of travel utility* (5).
Ton, D., Nijënstein, S., & Shelat, S. (2019). Fietsen naar de tramhalte: Simultane modellering van voortransport-en haltekeuze. *2019 Colloquium Vervoersplanologisch Speurwerk*, 1–15.
Ton, D., Shelat, S., Nijënstein, S., Rijsman, L., Oort, N. V., & Hoogendoorn, S. (2020). Understanding the role of cycling to urban transit stations through a simultaneous access mode and station choice model. *Transport Research Record*. https://doi.org/10.1177/0361198120925076
Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge university press.
Van Wee, B., Rietveld, P., & Meurs, H. (2006). Is average daily travel time expenditure constant? in search of explanations for an increase in average travel time. *Journal of transport geography*, *14*(2), 109–122.
van Boggelen, O. (2008). Het fietsparkeren bij vier stations onder een vergrootglas. *Fietsverkeer*, *7*, 26–31.

van Goeverden, C. D., & Egeter, B. (1993). *Gecombineerd gebruik van fiets en openbaar vervoer: Verwachte effecten opde vervoerwijzekeuze van optimale fietsbeschikbaarheid in voor- en natransport.* TU Delft, Faculteit de Civiele Techniek.

van Hagen, M., Govers, B., & de Haan, M. (2012). *Robuust sturen op keuzegedrag van mobilisten*. Colloquium Vervoersplanologisch Speurwerk.

van Mil, J. F. P., Leferink, T. S., Annema, J. A., & van Oort, N. (2020). Insights into factors affecting the combined bicycle-transit mode. *Public Transport*. https://doi.org/10.1007/s12469-020-00240-2

van Oort, N., van Wee, B., & Hoogendoorn, S. (2017). Nieuwe lessen over de potentie van fiets en ov. *CVS Congres, Gent*.

van Uum, J. R. G., Salverda, J. C., & Veling, I. H. (1995). De rol van de fiets in het verbindend stads- en streekvervoer. *Traffic Test, Veenendaal*.

Verkerk, S. (2019). *Aantal in- en uitstappers per station bij ns 2013 - 2018*.

Vlek, C. (2000). Essential psychology for environmental policy making. *International Journal of Psychology*, *35*, 153–167. https://doi.org/10.1080/002075900399457

Vos, M., Hagen, M. V., & Spoorwegen, N. (2015). *Licht op treinstations*. CVS.

Wang, D., & Li, J. (2002). Handling large numbers of attributes and/or levels in conjoint experiments. *Geographical Analysis*, *34*(4), 350–362.

Weliwitiya, H., & Eng, B. (2020). *Bicycle train intermodality: Exploring mode choice decisions and mode shift potential*.

Young, M., & Blainey, S. (2018). Railway station choice modelling: A review of methods and evidence. *Transport Reviews*, *38*(2), 232–251.

Zhang, N., Chen, F., Zhu, Y., Peng, H., Wang, J., & Li, Y. (2020). A study on the calculation of platform sizes of urban rail hub stations based on passenger behavior characteristics. *Mathematical Problems in Engineering*, *2020*. https://doi.org/10.1155/2020/3689760

Zhao, F., & Ubaka, I. (2004). Transit network optimization-minimizing transfers and optimizing route directness. *Journal of Public Transportation*, *7*(1), 4.

# A

# Scientific Paper

# Station Choice for Cyclists Regarding Perceived Safety and Comfort

**A.A. Barneveld**

*To obtain the degree of Master of Science in Transport, Infrastructure & Logistics,*
*Delft University of Technology, Delft, The Netherlands*

E-mail: anne.barneveld@live.nl

*1 May 2022*

**Abstract** Investigating station choice is important for possibly solving various capacity issues such as bicycle parking. Current literature is often restricted to dissatisfier parameters such as travel time, presence of transfer and travel costs. In order to fully understand station choice, this research includes factors regarding perceived safety and comfort, lighting at the station, shelter at the station and security in the bicycle parking, next to access travel time and bicycle parking costs. In this research, a choice experiment is carried out, and the retrieved data is analysed using a Multinominal Logit (MNL) model and a Latent Class (LC) model. The weights retrieved by the MNL model align with the assumptions that travel time and costs are most important for station choice. When applying an LC Model, two classes are found. One class (65%) is highly influenced by travel time and parking costs, which is equal to the MNL results. Males and students are very likely to belong in this class. Whereas the other class (35%) is also highly influenced by lighting and security in the bicycle parking and parking costs have a very low impact on the station choice for this class. Females and elderly (65+) are most likely to belong in this class. Thus for 35% of the population, satisfiers play a strong role in station choice. This leads to the conclusion that satisfiers play a role in station choice. The extent of influence by satisfiers on station choice is dependent on various socio-demographic characteristics.

**Key words:** Station Choice, Cyclist, Bicycle-train, Mode-combination, Satisfiers, Comfort and Perceived Safety, Latent Class Model

## Introduction

Passenger transport plays a major role in emitting greenhouse gasses, which should be reduced significantly in the near future (Stanley et al., 2011). Besides sustainability goals, ongoing urbanisation takes place. To accommodate the growth of transport in a sustainable way, the bicycle-train combination could play a large role since it is able to compete with the mode car on speed and accessibility (Kager et al., 2016). The bicycle-train combination combines the speed of the train with the connectivity of the bicycle. Currently, over 50% of the train passengers use their bicycle as an access mode for the train (ProRail, 2019). The largest bicycle parking facility in The Netherlands already accommodates more than 12.500 bicycles. The expanding popularity of the bicycle as a connection mode results in high parking demands at stations (van Boggelen, 2008). Increasing the capacity of bicycle parking facilities at these stations requires large investments and is not always possible since these facilities have to be built in dense city centres. However, within urban regions, travellers could often find more than one train station within 20 minutes of cycling.

To investigate if the needed capacity of bicycle parking for travellers can be accommodated at local stations, understanding of the incentives why people cycle to which station is needed. Existing research focuses mostly on dissatisfier factors, such as travel time and costs, presence of transfers and waiting time (van Mil et al., 2020). These dissatisfiers belong in the lower layers of the customer pyramid of needs (figure 1). However, van Hagen concluded that people do not have a sense of time. He states that the factor travel time is influenced by time perception, which is influenced by other factors. Next to the factor access time, the factor travel costs is highly dependent on personal characteristics, e.g. income and age. So time perception could be different for various user groups. Therefore, the research goal is to investigate the influence of satisfier characteristics, such as (perceived) safety and comfort, referring to the upper layers of the pyramid (figure 1), on station choice.

The research question of this study is:

*"To what extent can satisfiers, as opposed to dissatisfiers, play a role in station choice for bicycle-train travellers regarding local railway stations?"*

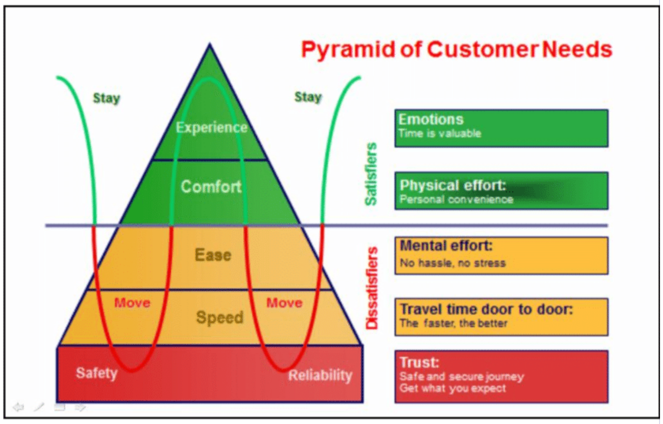



**Figure 1** Pyramid of Customer Needs (Hagen and Bron, 2014)

## Methodology

To answer this research question, several methods are used in this research. First, a literature study is performed to list all factors influencing station choice and their positive or negative relation to station choice. When all factors are stated, a ranking of all satisfier factors is performed by respondents. The three most important satisfier factors are included in the stated choice experiment, together with access travel time and parking costs. A stated choice experiment is used to determine the weight of all five factors. To compute the factor weights, two models are used. At first a Multinominal Logit model is applied to retrieve the factors weights for the whole group of respondents. In addition, a Latent Class model is applied to find weights for similar respondents in the population. To relate the results to practice, an illustrative case is performed regarding the city of Delft.

## Satisfier Factors and Bicycle-Train Travellers

To answer the research question, all possible characteristics influencing station choice for cyclists are found in a literature study. Most research stated the satisfiers in their factor lists but did not include any satisfier further in their method. In table 6 all factors are shown including their relation to the pyramid of customer needs. For each factor it is clear if it is a satisfier, dissatisfier, both or neither.

In 2018, seven classes are defined by Shelat et al. These classes (table 1) are incorporated further in this research to have a wide variation in respondents and the classes created in the latent class model are reflected on the research of Shelat et al.

| Class | Description |
|---|---|
| 1 | Middle-aged full-time professionals (26%) |
| 2 | University students living with parents (23%) |
| 3 | School children (15%) |
| 4 | Young, low income professionals (14%) |
| 5 | Middle-aged part-time professionals (10%) |
| 6 | University students living alone (9%) |
| 7 | Pensioners (2%) |

**Table 1** Seven classes of bicycle-train travellers (Shelat et al., 2018)

## Ranking of Factors

These factors are ranked by 21 respondents. The respondents were chosen such that all seven user classes of the bicycle-train combination, as found by Shelat et al. (2018), were represented by three respondents. After weighing the results to the dynamics of society, the following overall top 3 ranking is:
1. Level of shelter at station
2. Level of lighting at station
3. Presence of security in the bicycle station

Due to the small respondent group (N=21), this ranking step is indicative and might not be statistically significant. This means the final ranking might not be fully representative for the whole population. This step is solely taken to gain explorative insights into the prioritisation of factors that influence station choice.

In this research, station choice is prioritised above investigating satisfiers, as station choice is not only made in the upper layers of the customer needs pyramid. For that reason, two important dissatisfiers related to station choice are included in the rest of this research: access travel time and bicycle parking costs (Ton et al., 2019; Krabbenborg et al., 2015; Weliwitiya and Eng, 2020; Young and Blainey, 2018; Givoni and Rietveld, 2007).

## Multinominal Logit Model Results

The top five factors do not have an equal influence on station choice. This was further investigated by using a stated choice experiment. First, a pilot survey is taken by 38 respondents to retrieve prior values. While retrieving respondents for the pilot survey, all seven classes of Shelat et al. are included. Prior values are used to create an efficient design for the final stated choice survey (ChoiceMetrics, 2018). In this final stated choice survey, 308 respondents were asked to choose between station A and station B, having different values for the parameters: level of lighting, level of shelter, presence of security, cycling access

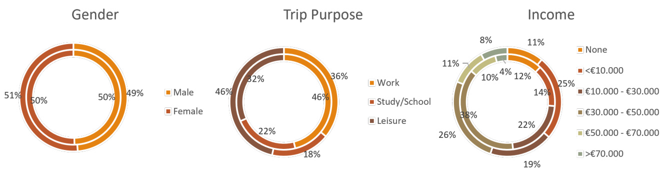



**Figure 2** Comparison data-sets final survey (outer) and ODIN 2019 (inner)

time and bicycle parking costs. All respondents varied on several socio-demographic factors, such as gender, age, income, education, social participation and household composition, train frequency and trip purpose. The distribution of personal factors of respondents in the final survey was very similar to the full population of bicycle-train travellers retrieved from the dataset ODIN 2019. Three examples of the comparison of the final and ODIN dataset are shown in figure 2, an expanded version can be found in figure 14.

The distribution of weights by respondents is first modelled using an MNL model, with the utility function as presented in equation 1.

$$U_i = \beta_{Light} * Light_i + \beta_{Shelter} * Shelter_i + \beta_{Security} * Security_i + \beta_{Time} * Time_i + \beta_{Price} * Costs_i \quad (1)$$

After modelling the responses on the final survey regarding station choice using an MNL model, meaning the average is taken of all respondents, the following weights are found (table 2). The number of attribute levels are also shown in table 2. When an attribute has two levels, 0 and 1 are implied. A value of zero means a low level of lighting, shelter and security, and a value of one means a high level. For access time and parking costs three levels are used. These levels are 5, 7.5 or 10 minutes access time and €0, €0.50 or €1. The minimum and maximum column of table 2, refers to the utility difference caused by that parameter. This indicates the range of the effect on the total utility function for each parameter.
The weights indicate a change of utility when the parameter changes by one. Meaning when lighting is increased, the utility could increase by 0.667, and when parking costs are increased by €1, the utility of that station decreases by -1.09. Regarding the range of attributes, it is visible that the average traveller using bicycle-train is the most sensitive to changes in the access time or parking costs parameter.

Parameter weights can only be compared with different models when a ratio is computed, for example the Value-of-Time. In order to compare the results with other models, and to show the results more clearly, all parameter ratios with the dissatisfiers costs and time are computed. The ratio values are in the unit of the base parameter: euros (figure 3) and minutes (figure 4). It can be concluded respondents are willing to pay

**Table 2** Overall Parameter Weights

| Factor | Value | # levels | minimum | maximum |
|---|---|---|---|---|
| Light | 0.667 | 2 | 0 | 0.667 |
| Shelter | 0.364 | 2 | 0 | 0.364 |
| Security | 0.554 | 2 | 0 | 0.554 |
| Access time | -0.122 | 3 | -1.22 | -0.610 |
| Parking costs | -1.09 | 3 | -1.09 | 0 |

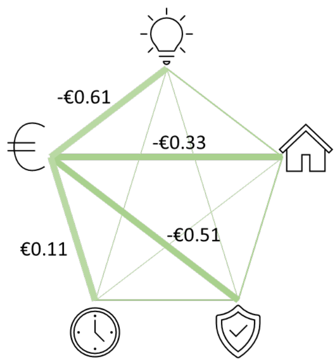



**Figure 3** Costs as base (MNL)

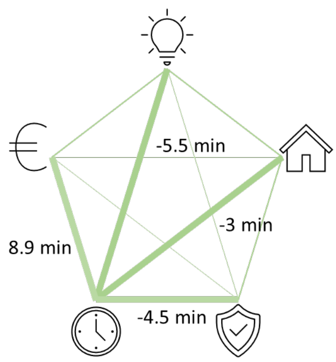



**Figure 4** Time as base (MNL)

€0.61 parking costs or to cycle 5.5 minutes extra for a well-lighted station.

## Latent Class Model Results

To allow heterogeneity between user groups also a Latent Class Model is created to capture the differences in the population. A latent class choice model is used for analysing the results to gain more insight if there are large differences between bicycle-train travellers. Two classes are found based similar weights of respondents. Afterwards, it is analysed which socio-demographic characteristics explain class-membership. The characteristics of the two user groups are presented in figure 5.

Figure 5 shows that females, elderly and non-students are likely to belong to group 1. This first group is sensitive to the parameters lighting, security and access travel time. Group 1 has an overall size of 35% of all travellers, which makes group 2 65% of the population. Males and students have a probability of 83% for belonging to the second class. Group 2 is highly sensitive to access travel time and parking costs. Each class has different values for the parameter weights, shown in table 3 and 4. In figures 6 and 7, the effect of all parameter weights for both classes on the total utility is shown.
As already stated, parameter weights cannot be directly compared to results from other models. To be able to compare the two groups, four pentagon figures are designed (figures 8-11), showing the ratio of all parameters with time and costs as a base parameter for both groups (G1 and G2). Figures 8

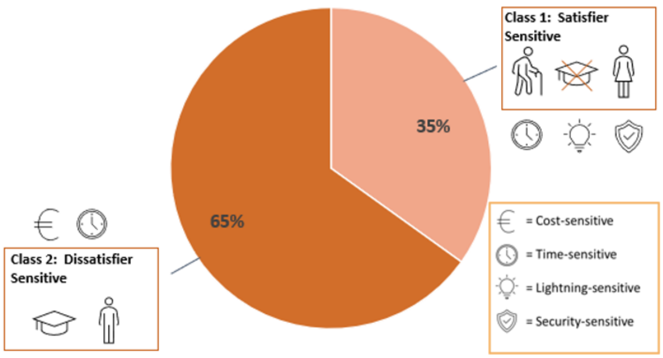



**Figure 5** Summary of two classes

**Table 3** Group 1 Parameters

| Parameter | Weight | Minimum | Maximum |
|---|---|---|---|
| $\beta_{Light}$ | 0.765 | 0 | 0.765 |
| $\beta_{Shelter}$ | 0.282 | 0 | 0.282 |
| $\beta_{Security}$ | 0.610 | 0 | 0.610 |
| $\beta_{Time}$ | -0.107 | -1.07 | -0.545 |
| $\beta_{Costs}$ | -0.258 | -0.258 | 0 |

**Table 4** Group 2 Parameters

| Parameter | Weight | Minimum | Maximum |
|---|---|---|---|
| $\beta_{Light}$ | 0.588 | 0 | 0.588 |
| $\beta_{Shelter}$ | 0.738 | 0 | 0.738 |
| $\beta_{Security}$ | 0.532 | 0 | 0.532 |
| $\beta_{Time}$ | -0.272 | -2.72 | -0.750 |
| $\beta_{Costs}$ | -3.411 | -3.411 | 0 |


**Figure 6** Utility Effect Group 1

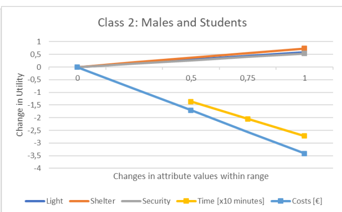



**Figure 7** Utility Effect Group 2

and 9 show that respondents belonging to the first group are willing to pay €2.97 parking costs or add 7 minutes of access time for a well-lighted station. Whereas the second group is only willing to pay €0.17 or 2 minutes of access time for a well-lighted station.

These differences between the two groups are very interesting. Overall it can be noted that the first group is more sensitive to the satisfiers compared to the second group.

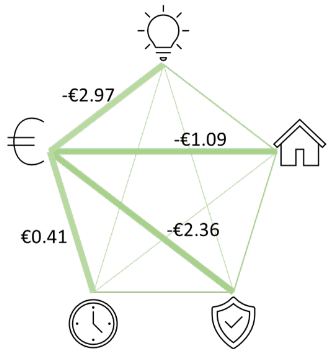



**Figure 8** G1: Costs as base

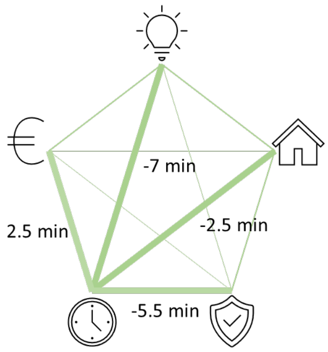



**Figure 9** G1: Time as base

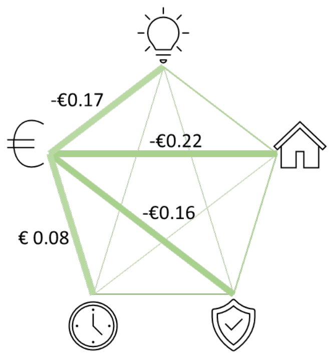



**Figure 10** G2: Costs as base

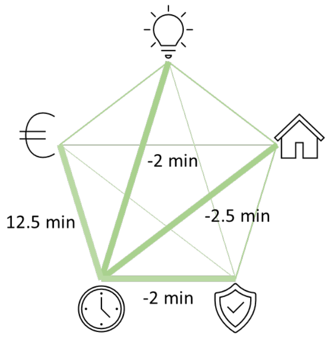



**Figure 11** G2: Time as base

## Application in Practice

Knowing this distribution of weights among various socio-demographic characteristics it is visible that different measures are needed to appeal to both the user classes. The ratio of socio-demographic characteristics of the municipality Delft is given in table 5 (CBS, 2021; CBS, 2022). Leading to the computation that 39963 inhabitants (38%) would belong in group 1, and 64570 inhabitants (62%) would belong in group 2.

| Factor | # Inhabitants | Percentage |
|---|---|---|
| Delft (total) | 104533 | 100% |
| Male | 55159 | 52% |
| Female | 49374 | 47% |
| Student | 18807 | 18% |
| Non-Student | 85726 | 82% |
| 65+ | 16767 | 16% |
| 65- | 87766 | 84% |

**Table 5** Socio demographics of Delft (CBS, 2021; CBS, 2022)

Measures that would affect the access travel time would be most effective for increasing the utility of a local station for both groups. Unfortunately, these measures are also expected to be the highest of costs and most technically and politically challenging. In figure 12 and 13 a map is illustrated to show the bicycle access time of 5 and 10 minutes for station Delft and Delft Campus. It is visible that both stations overlap several areas with 10 minutes of cycling time. For these areas, the resulting weights of the Latent Class model are accurate.

On a working day, 15 intercity trains and 8 local trains stop at Station Delft per hour. Only the 8 local trains also stop at Delft Campus. Since no transfer is taken into account, the results are applicable on the 8 local trains that stop at both stations.

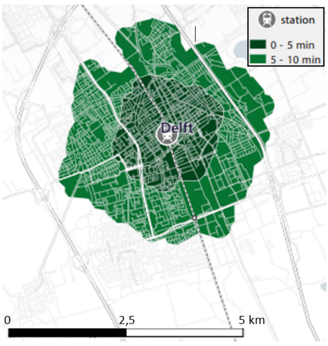



**Figure 12** Access time Station Delft

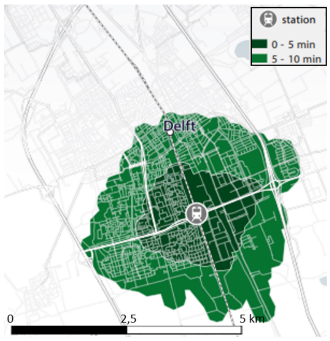



**Figure 13** Access time Station Delft Campus

Furthermore, group 1 highly appreciates lighting at the station and security in the bicycle parking. It could be imagined that increasing lighting at the station and security in the bicycle parking has fewer costs and challenges compared to changing the infrastructure. For group 2, the satisfier shelter at the station has the highest weight of all satisfiers. This means possible extra dis-utilities of access time and parking costs could be partly compensated by placing a full shelter cubicle on the platform. Providing more shelter will also affect the first group, and providing more lighting and security is also of influence for the second group, however, in a smaller ratio. Next to actual access travel time, also the dissatisfier parking costs is very important for the second class. Decreasing the parking price for the local station is not always possible, for example when parking is free at both stations. In that case, increasing the parking price at the central station could lead to more travellers choosing the local station.

For example, when a traveller (group 2) lives 5 cycling minutes from station Delft and 10 minutes from Delft Campus, adding a parking fee of €0.50 per day at station Delft and adding extra shelter at Delft Campus. These circumstances would lead to the utilities for group 2 in equations 2 and 3. Concluding that 62% of bicycle-train travellers (group 2 in Delft) has a probability of 55% for station Delft and 45% for Delft Campus, using equation 6.

$$U_{Delft} = 0.588 * 1 + 0.738 * 1 + 0.532 * 1 \\ - 0.272 * 5 + -3.411 * 0.5 = -1.21 \tag{2}$$

$$U_{Campus} = 0.588 * 1 + 0.738 * 1 + 0.532 * 0 \\ - 0.272 * 10 + -3.411 * 0 = -1.39 \tag{3}$$

When a traveller (group 1) lives 5 and 10 minutes respectively from Delft and Delft Campus, and shelter and security is added at Delft Campus. This would lead to the utilities in equation 4 (Delft) and 5 (Delft Campus).

$$U_{Delft} = 0.765 * 1 + 0.282 * 1 + 0.610 * 1 \\ - 0.107 * 5 - 0.258 * 0 = 1.12 \tag{4}$$

$$U_{Campus} = 0.765 * 1 + 0.282 * 0 + 0.610 * 1 \\ - 0.107 * 10 - 0.258 * 0 = 0.305 \tag{5}$$

Resulting in the probability that 63% of 38% train travellers will choose for Delft and 37% will choose Delft Campus. This is compared to a share of 19% for Delft Campus when low shelter or no security is present.

$$P(Y = i) = \frac{e^{V_i}}{\sum e^{V_j}} \tag{6}$$

## Conclusion

This research shows that satisfiers can influence station choice. The extent of the effect of satisfiers on station choice is dependent on socio-demographic characteristics such as gender, age and social participation (student/non-student). The first group (35%) is found to be quite sensitive to lighting and security in bicycle parking. Meaning this group is likely to respond to measures concerning these satisfiers. Regarding the results and possible measures, it is seen that for the largest group (65%), the dissatisfiers have a larger influence on station choice than the satisfiers. For this group, the implementation of measures regarding the dissatisfiers are therefore more effective compared to the satisfiers. Concluding that for the second group, satisfiers have an influence on station choice, however, this could be overshadowed by the dissatisfiers.

## Discussion and Recommendations

This research provides new insights into the role of satisfiers on station choice. However, there are several ideas to improve and expand this research. As already stated in this paper, the ranking was done on a very small test group (N=21). This step was only indicative in this research, nonetheless it would be very interesting to perform this ranking on a larger group and compare the resulting top 3. If other factors are ranked most important, these could be included in a new stated choice experiment.

Regarding dissatisfiers, access time and parking costs are researched. This research did not include other dissatisfiers such as the presence of transfers within the train trip and relation to station choice. According to van Mil et al. (2020) transfers are of high importance in the station choice for travellers. Excluding transfers means that the results are applicable for train trips that can be taken from both central and local stations, or for travellers indifferent against transfers. For further research it is recommended to include other dissatisfiers combined with satisfiers in a stated choice experiment.

It is also recommended to further develop research using Latent Class Modelling. The differences between the resulting weights from the MNL model and the LC model in station choice are clearly visible. Showing that there are two significant classes that can be explained by socio-demographic factors. It shows that it is important to allow for heterogeneity in the population and not only take the average value. This leads to results being better applicable in reality.

Practical advice for municipalities and station owners based on this research is to not solely focus on central stations when renewing or designing and keeping the local stations up to the standards of the travellers. It could also be recommended to make a Cost-Benefit analyses for each station specific. In this paper an indication is given how satisfiers influence travel behavior. However, in this example of Delft access time is stated as 5 minutes for Delft and 10 minutes for Delft Campus. Varying the access time and applying a station specific Cost-Benefit analyses allows a more accurate forecast of station choice by travellers and a Cost-Benefit analyses would result much insight in the advantages and disadvantages of increasing station lighting, shelter and security for a specific station.

## References


Birago, D., Mensah, S. O., & Sharma, S. (2017). Level of service delivery of public transport and mode choice in accra, ghana. *Transportation research part F: traffic psychology and behaviour*, *46*, 284–300.

Bowman, L. A., & Turnquist, M. A. (1981). Service frequency, schedule reliability and passenger wait times at transit stops. *Transportation Research Part A: General*, *15*(6), 465–471.

Broach, J., Dill, J., & Gliebe, J. (2012). Where do cyclists ride? a route choice model developed with revealed preference gps data. *Transportation Research Part A: Policy and Practice*, *46*(10), 1730–1740.

CBS. (2021). Leerlingen, deelnemers en studenten; onderwijssoort, woonregio [25 June 2021]. *Centraal Bureau voor de Statistiek*. Retrieved May 1, 2022, from https://opendata.cbs.nl/statline/#/CBS/nl/dataset/71450ned/table?ts=1651845346426

CBS. (2022). Regionale kerncijfers nederland [31 March 2022]. *Centraal Bureau voor de Statistiek*. Retrieved May 1, 2022, from https://opendata.cbs.nl/statline/#/CBS/nl/dataset/70072ned/table?ts=1651846753350

Cervero, R., Caldwell, B., & Cuellar, J. (2013). Bike-and-ride: Build it and they will come. *Journal of Public Transportation*, *16*(4), 5.

ChoiceMetrics. (2018). *Ngene 1.2 user manual and reference guide the cutting edge in experimental design*. www.choice-metrics.com

Cui, Y., Mishra, S., & Welch, T. F. (2014). Land use effects on bicycle ridership: A framework for state planning agencies. *Journal of Transport Geography*, *41*, 220–228.

Ferrell, C. E., & Mathur, S. (2012). Influences of neighborhood crime on mode choice. *Transportation Research Record*, *2320*, 55–63. https://doi.org/10.3141/2320-07

Garcia, A., Gomez, F. A., Llorca, C., & Angel-Domenech, A. (2015). Effect of width and boundary conditions on meeting maneuvers on two-way separated cycle tracks. *Accident Analysis & Prevention*, *78*, 127–137.

Givoni, M., & Rietveld, P. (2007). The access journey to the railway station and its role in passengers' satisfaction with rail travel. *Transport Policy*, *14*, 357–365. https://doi.org/10.1016/j.tranpol.2007.04.004

Hagen, M. V., & Bron, P. (2014). Enhancing the experience of the train journey: Changing the focus from satisfaction to emotional experience of customers. *Transportation Research Procedia*, *1*, 253–263. https://doi.org/10.1016/j.trpro.2014.07.025

Kager, R., Bertolini, L., & Brömmelstroet, M. T. (2016). Characterisation of and reflections on the synergy of bicycles and public transport. *Transportation Research Part A: Policy and Practice*, *85*, 208–219. https://doi.org/10.1016/j.tra.2016.01.015

Keijer, M. J. N., & Rietveld, P. (2000). How do people get to the railway station? the dutch experience. *Transportation Planning and Technology*, *23*(3), 215–235. https://doi.org/10.1080/03081060008717650

Krabbenborg, L., Wee, B. V., Annema, J. A., Correira, G., & Snellen, D. (2015). *Cycling to a railway station*.

Krenn, P. J., Oja, P., & Titze, S. (2014). Route choices of transport bicyclists: A comparison of actually used and shortest routes. *International journal of behavioral nutrition and physical activity*, *11*(1), 1–7.

Liu, Y., Yang, D., Timmermans, H. J., & de Vries, B. (2020). Analysis of the impact of street-scale built environment design near metro stations on pedestrian and cyclist road segment choice: A stated choice experiment. *Journal of transport geography*, *82*, 102570.

Mackenbach, J. D., Randal, E., Zhao, P., & Howden-Chapman, P. (2016). The influence of urban land-use and public transport facilities on active commuting in wellington, new zealand: Active transport forecasting using the wilute model. *Sustainability*, *8*(3), 242.

Martens, K. (2004). The bicycle as a feedering mode: Experiences from three european countries. *Transportation Research*

*Part D: Transport and Environment*, *9*, 281–294. https://doi.org/10.1016/j.trd.2004.02.005

Park, S., Kang, J., & Choi, K. (2014). Finding determinants of transit users' walking and biking access trips to the station: A pilot case study. *KSCE Journal of Civil Engineering*, *18*(2), 651–658.

ProRail. (2019). *Fietsparkeren bij stations*. https://www.prorail.nl/reizen/stations/fietsen

Puello, L. L. P., & Geurs, K. T. (2016). Integration of unobserved effects in generalised transport access costs of cycling to railway stations. *EJTIR Issue*, *16*, 385–405.

Puello, L. L. P., & Geurs, K. (2015). Modelling observed and unobserved factors in cycling to railway stations: Application to transit-oriented-developments in the netherlands. *European journal of transport and infrastructure research*, *15*(1).

Replogle, M. A. (1987). Bicycles on transit: A review of international experience. *Transportation research record*, *1141*, 26–36.

Rietveld, P. (2000a). The accessibility of railway stations: The role of the bicycle in the netherlands. *Transportation Research Part D: Transport and Environment*, *5*(1), 71–75. https://doi.org/https://doi.org/10.1016/S1361-9209(99)00019-X

Rietveld, P. (2000b). *Non-motorised modes in transport systems: A multimodal chain perspective for the netherlands*. www.elsevier.com/locate/trd

Rietveld, P., & Daniel, V. (2004). Determinants of bicycle use: Do municipal policies matter? *Transportation Research Part A: Policy and Practice*, *38*(7), 531–550.

Semler, C., & Hale, C. (2010). Rail station access–an assessment of options.

Shelat, S., Huisman, R., & van Oort, N. (2018). Analysing the trip and user characteristics of the combined bicycle and transit mode. *Research in Transportation Economics*, *69*, 68–76. https://doi.org/10.1016/j.retrec.2018.07.017

Stanley, J. K., Hensher, D. A., & Loader, C. (2011). Road transport and climate change: Stepping off the greenhouse gas. *Transportation Research Part A: Policy and Practice*, *45*(10), 1020–1030.

Stinson, M. A., Bhat, C. R. et al. (2005). A comparison of the route preferences of experienced and inexperienced bicycle commuters. *TRB 84th annual meeting compendium of papers*, (05-1434).

Ton, D., Nijënstein, S., & Shelat, S. (2019). *Fietsen naar de tramhalte: Simultane modellering van voortransport-en haltekeuze*.

Van Wee, B., Rietveld, P., & Meurs, H. (2006). Is average daily travel time expenditure constant? in search of explanations for an increase in average travel time. *Journal of transport geography*, *14*(2), 109–122.

van Boggelen, O. (2008). Het fietsparkeren bij vier stations onder een vergrootglas. *Fietsverkeer*, *7*, 26–31.

van Hagen, M., Govers, B., & de Haan, M. (2012). *Robuust sturen op keuzegedrag van mobilisten*. Colloquium Vervoersplanologisch Speurwerk.

van Mil, J. F. P., Leferink, T. S., Annema, J. A., & van Oort, N. (2020). Insights into factors affecting the combined bicycle-transit mode. *Public Transport*. https://doi.org/10.1007/s12469-020-00240-2

Weliwitiya, H., & Eng, B. (2020). *Bicycle train intermodality: Exploring mode choice decisions and mode shift potential*.

Young, M., & Blainey, S. (2018). Railway station choice modelling: A review of methods and evidence. *Transport Reviews*, *38*(2), 232–251.

Zhao, F., & Ubaka, I. (2004). Transit network optimization-minimizing transfers and optimizing route directness. *Journal of Public Transportation*, *7*(1), 4.

| Factor | Relation | Type | Source |
|---|---|---|---|
| **Access trip** | | | |
| distance | - | dissatisfier | (Van Wee et al., 2006; Keijer and Rietveld, 2000; Martens, 2004) |
| travel time | - | dissatisfier | (Weliwitiya and Eng, 2020) |
| cycling infrastructure | + | dissatisfier | (Cervero et al., 2013) |
| accessibility to station | + | dissatisfier | (Cui et al., 2014; Rietveld, 2000b) |
| right of way | + | satisfier | (Stinson, Bhat, et al., 2005) |
| presence of obstacles | - | satisfier | (Garcia et al., 2015) |
| human scale | + | satisfier | (Park et al., 2014; Krabbenborg et al., 2015) |
| topography | +/- | satisfier | (Semler and Hale, 2010; Weliwitiya and Eng, 2020) |
| vegetation | + | satisfier | (Krenn et al., 2014; Puello and Geurs, 2016; Mackenbach et al., 2016) |
| vehicle volumes | +/- | dissatisfier | (Cui et al., 2014) |
| crowdedness | +/- | both | (Krabbenborg et al., 2015; Puello and Geurs, 2015) |
| speed limit | - | dissatisfier | (Park et al., 2014) |
| social safety | + | satisfier | (Hagen and Bron, 2014) |
| traffic safety | + | dissatisfier | (Rietveld and Daniel, 2004) |
| street lighting | + | satisfier | (Puello and Geurs, 2015) |
| route preference | +/- | both | (Rietveld, 2000a; Broach et al., 2012; Liu et al., 2020; van Hagen et al., 2012) |
| **Station Characteristics** | | | |
| size of node | + | both | |
| train types | + | both | |
| bicycle parking | + | satisfier | (Keijer and Rietveld, 2000; Givoni and Rietveld, 2007; Cervero et al., 2013) |
| security | + | both | (Replogle, 1987; Givoni and Rietveld, 2007) |
| overview | + | satisfier | |
| parking price | - | dissatisfier | (van Mil et al., 2020) |
| transfer time | - | dissatisfier | (van Mil et al., 2020) |
| station environment | + | satisfier | |
| vegetation | + | satisfier | (Krenn et al., 2014) |
| neighbourhood reputation | + | satisfier | (Ferrell and Mathur, 2012) |
| level of maintenance | + | satisfier | (Birago et al., 2017) |
| tidiness | + | satisfier | (Birago et al., 2017) |
| level of lighting | + | satisfier | |
| station services | + | satisfier | (Puello and Geurs, 2016; Cervero et al., 2013) |
| toilets | + | satisfier | (Puello and Geurs, 2016) |
| shopping facilities | + | satisfier | (Puello and Geurs, 2016; Weliwitiya and Eng, 2020) |
| eating facilities | + | satisfier | |
| level of shelter | + | satisfier | |
| seating at platform | + | satisfier | |
| seating at station | + | satisfier | |
| spaciousness | + | satisfier | |
| crowdedness | - | satisfier | (van Mil et al., 2020) |
| **Train trip** | | | |
| directness | + | dissatisfier | (Zhao and Ubaka, 2004;van Mil et al., 2020) |
| travel time | - | dissatisfier | |
| train type | +/- | both | |
| reliability | + | dissatisfier | |
| frequency | + | dissatisfier | (Bowman and Turnquist, 1981) |
| price | - | dissatisfier | |
| **General factors** | | | |
| age | +/- | | |
| gender | +/- | | |
| income | +/- | | |
| car ownership | +/- | | |
| time of day | +/- | | |
| trip type | +/- | | |
| trip frequency | +/- | | |
| weather | +/- | satisfier | |

**Table 6** Factors of Station Choice related to Pyramid of Customer Needs (Hagen and Bron, 2014)

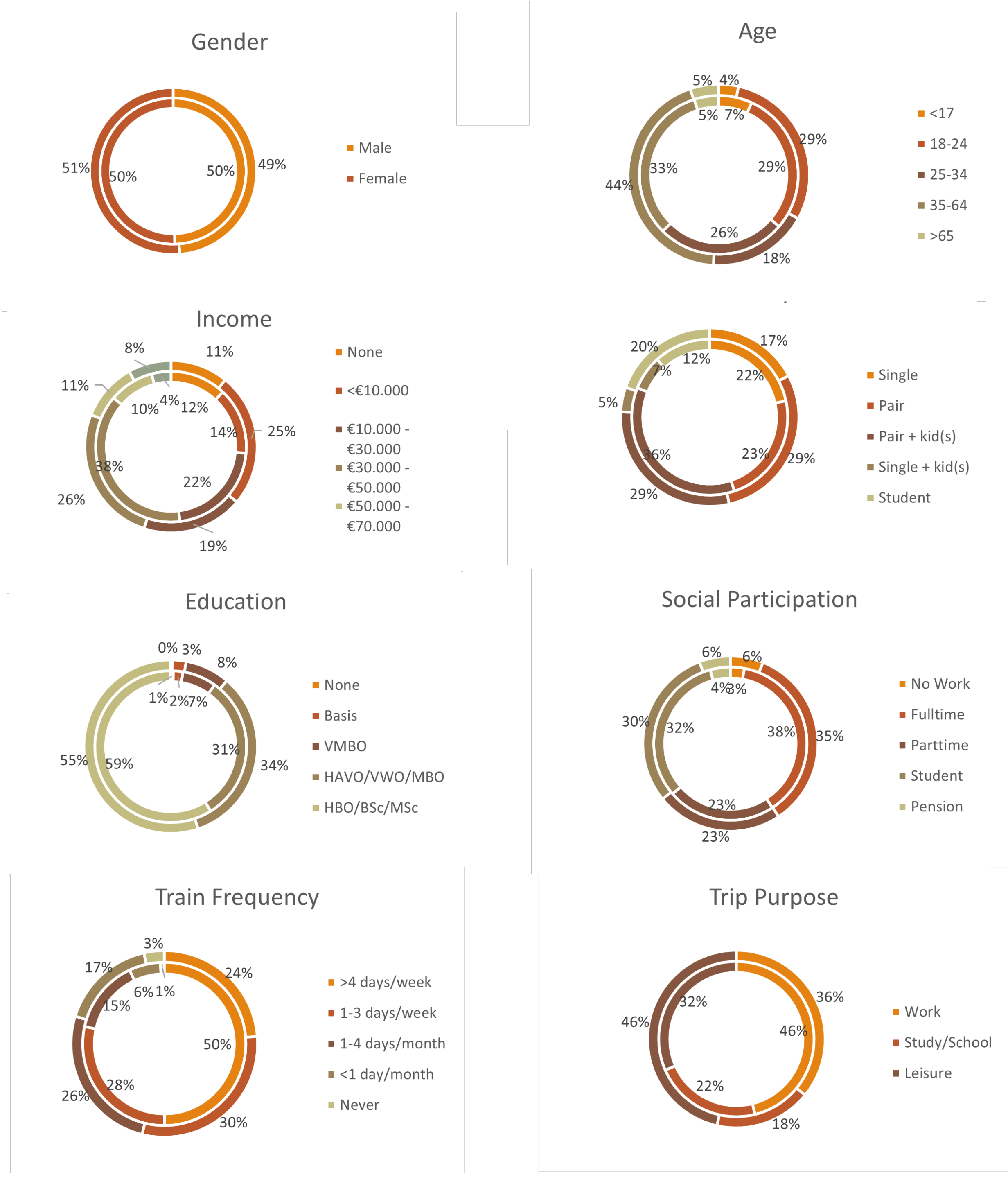



**Figure 14** Comparison final respondents and ODIN data-set on socio-demographics

# B

# 7-cluster model

| Cluster number | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| **Cluster size** | 26.50% | 22.84% | 15.10% | 14.06% | 9.64% | 9.47% | 2.39% |
| **Label** | Middle-aged full-time professionals | University students living with parents | School children | Young, low income professionals | Middle-aged part-time professionals | University students living alone | Pensioners |
| **Household indicators** | | | | | | | |
| **Household composition** | | | | | | | |
| Single | 4% | 1% | 1% | **64%** | 2% | **58%** | **39%** |
| Couple | **37%** | 1% | 0% | **24%** | **31%** | 11% | **59%** |
| Couple + kid(s) | **55%** | **84%** | **86%** | 1% | **64%** | 2% | 0% |
| Couple + kid(s) + other(s) | 0% | 1% | 0% | 0% | 1% | 0% | 0% |
| Couple + other(s) | 0% | 0% | 0% | 0% | 0% | 1% | 0% |
| Single parent + kid(s) | 2% | 10% | 12% | 8% | 3% | 3% | 1% |
| Single parent + kid(s) + other(s) | 0% | 0% | 0% | 0% | 0% | 1% | 0% |
| Other composition | 1% | 2% | 0% | 2% | 0% | 25% | 1% |
| **Degree of urbanization** | | | | | | | |
| Extremely urbanised | **28%** | 7% | 6% | **50%** | **21%** | **61%** | **23%** |
| Strongly urbanised | **36%** | **22%** | 19% | **30%** | **33%** | **29%** | **25%** |
| Moderately urbanised | **22%** | **26%** | **21%** | 13% | **24%** | 6% | **34%** |
| Hardly urbanised | 11% | **30%** | **25%** | 4% | 14% | 3% | 9% |
| Not urbanised | 4% | 15% | **29%** | 2% | 8% | 1% | 10% |
| **Household disposable income** | | | | | | | |
| Lowest - 1st 10%ile | 1% | 5% | 7% | 9% | 5% | **71%** | 3% |
| 2nd 10%ile | 2% | 4% | **10%** | **10%** | 5% | 10% | **10%** |
| 3rd 10%ile | 3% | 7% | **9%** | **10%** | 5% | 5% | 7% |
| 4th 10%ile | 7% | 9% | **12%** | **13%** | 7% | 1% | **12%** |
| 5th 10%ile | 6% | 10% | **9%** | **13%** | 10% | 2% | **10%** |
| 6th 10%ile | 11% | **13%** | **10%** | **10%** | 11% | 3% | **10%** |
| 7th 10%ile | **12%** | **14%** | **11%** | **12%** | **15%** | 1% | **11%** |
| 8th 10%ile | **18%** | **17%** | **10%** | 9% | **12%** | 1% | **12%** |
| 9th 10%ile | **21%** | **14%** | **12%** | 6% | **17%** | 3% | 8% |
| Highest - 10th 10%ile | **20%** | 7% | **9%** | 6% | **13%** | 2% | **16%** |
| Unknown | 0% | 0% | 0% | 1% | 0% | 1% | 0% |
| **Individual** | | | | | | | |
| **Gender** | | | | | | | |
| Male | **70%** | 52% | 47% | 42% | 15% | 39% | 45% |
| Female | 30% | 48% | 53% | 58% | **85%** | 61% | 55% |
| **Age group** | | | | | | | |
| 1 to 17 | 0% | 1% | **100%** | 0% | 0% | 0% | 0% |
| 18 to 24 | 2% | **98%** | 0% | 4% | 5% | **84%** | 0% |
| 25 to 34 | 24% | 1% | 0% | **49%** | 17% | 15% | 0% |
| 35 to 64 | **74%** | 0% | 0% | **46%** | **78%** | 1% | 12% |
| > =65 | 0% | 0% | 0% | 0% | 0% | 0% | **88%** |
| **Highest education** | | | | | | | |
| None | 0% | 1% | 2% | 0% | 0% | 0% | 0% |
| Primary | 0% | 2% | **25%** | 1% | 0% | 0% | 1% |
| Secondary | 3% | 19% | **26%** | 7% | 9% | 3% | 16% |
| High school | 19% | **70%** | **18%** | 24% | **29%** | **55%** | 19% |
| University | **76%** | 7% | 0% | **67%** | **62%** | **41%** | **63%** |
| Other | 1% | 1% | 1% | 1% | 0% | 1% | 0% |
| Unknown | 0% | 0% | 0% | 0% | 0% | 0% | 1% |
| N/A (< 15 years) | 0% | 0% | 29% | 0% | 0% | 0% | 0% |
| **Transport** | | | | | | | |
| **Car availability** | | | | | | | |
| Other | 0% | 0% | 0% | 0% | 0% | 0% | 0% |
| Always available | 43% | 10% | 0% | 21% | 28% | 6% | 42% |
| Limited availability | 46% | 49% | 2% | 0% | 52% | 7% | 23% |
| No car in household | 5% | 0% | 0% | 49% | 4% | 59% | 11% |
| Younger than 18 | 0% | 0% | 98% | 0% | 0% | 0% | 0% |
| No license to drive | 6% | 41% | 0% | 30% | 16% | 28% | 24% |
| **Transit use frequency** | | | | | | | |
| Daily | **60%** | **70%** | **66%** | **53%** | 14% | **41%** | 3% |
| Once a week | **32%** | 27% | **19%** | **33%** | **65%** | **51%** | **52%** |
| Once a month | 6% | 2% | 8% | 13% | 13% | 7% | **34%** |
| < Once a month | 3% | 0% | 5% | 1% | 7% | 0% | 11% |
| Almost never | 0% | 0% | 2% | 0% | 1% | 1% | 0% |

Figure B.1: 7-cluster model (Shelat et al., 2018)

# C

# Factor Ranking by Usergroups

## C.1 Middle-aged full-time professionals

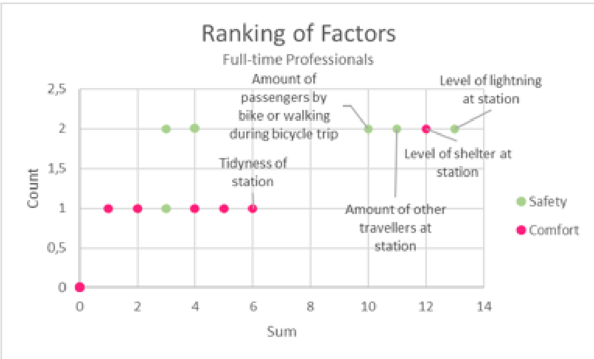



Figure C.1: Ranking Middle-aged full-time professionals

The three respondents that represent the largest group of the bicycle-train population were two men (53 and 39 years) and one woman (46 years). This distribution of gender is chosen to represent the distribution of Shelat et al. (2018, where 70% was male. The top 5 for this group includes the factors 1) 'Level of lightning at station', 2) 'Level of shelter at station', 3) 'Amount of other travellers at station', 4) 'Amount of passengers by bike or walking during bicycle trip' and 5) 'Tidiness of station' (figure C.1). Three factors of the top five are labeled as a safety-factor and two as a comfort-factor.

## C.2 University students living with parents

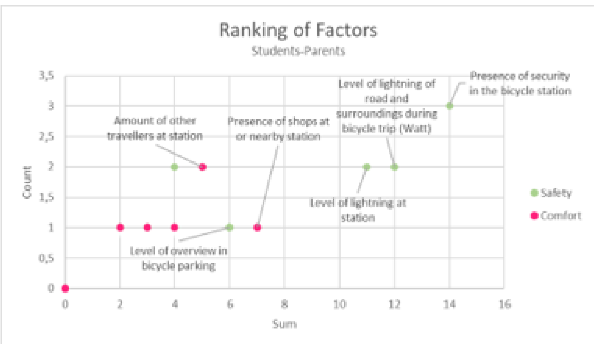



Figure C.2: Ranking Students living with their parents

The university students living with their parents are represented by two women (age 19 and 21) and one men (age 21). In this user group the most important factors are 1) Presence of security in the bicycle parking, 2) Level of lightning during bicycle trip, 3) Level of lightning at station, 4) Presence of shops in or nearby the station. For the fifth most important factor a priority needs to be stated. It is decided that the count has a slightly higher priority than the sum of each factor. This is explained by the fact that when a factor is chosen by multiple interviewees to be in their ranking, this could be more relevant for the whole population. Therefor in this usergroup, the factors on the fifth place are 5) Amount of other travellers at station and Level of shelter at the station. Both factors have the exact same sum and count, which allows the fifth place to be shared by two factors. Figure C.2 shows the ranking of factors for this usergroup. The three most important factors are all labeled as safety. The fourth and fifth ranked factors are related to comfort.

## C.3 School children

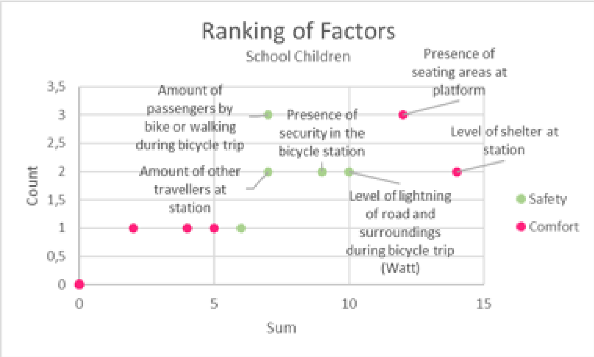



Figure C.3: Ranking School Children

The third group is represented by school children. In this usergroup two women and three men were interviewed. The top five for this group includes the factors 1) Level of shelter at station, 2) Presence of seating areas at platform, 3) Level of lightning during bicycle trip, 4) Presence of security in the bicycle station and 5) Amount of passengers by bike or walking during bicycle trip. This is also illustrated in figure C.3. The factor amount of other travellers at station is excluded from the top five according to the prioritisation of count over sum. The two most important factors in this usergroup are comfort-related, whereass the third to fifth factors are labeled as a safaty-factor.

## C.4 Young low income professionals

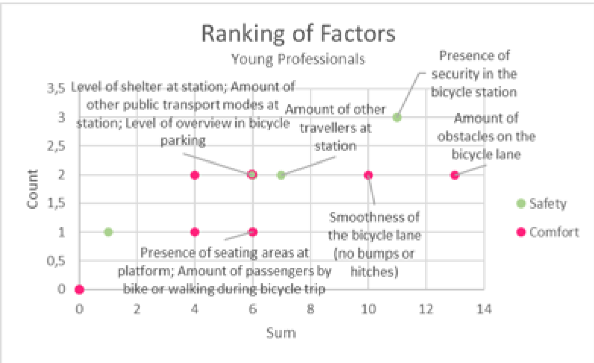



Figure C.4: Ranking Young low income professionals

The Young professionals questioned in the interviews were two males (age 26 & 31) and one female (age 27). One of these interviewees has a car. In this top 5, there are three factors with the same ranking in fifth place. The ranking in this group is shown in figure C.4. The factors included in this top five are 1) Amount of obstacles on the bicycle lane, 2) Presence of security in the bicycle station, 3) Smoothness of the bicycle lane, 4) Amount of other travellers at the station and 5)Level of shelter, amount of other public transport modes and level of overview in the bicycle parking. These factors are evenly distributed between safety and comfort. Even the factors at the fifth place are represented by both safety and comfort.

## C.5 Middle-aged part-time professionals

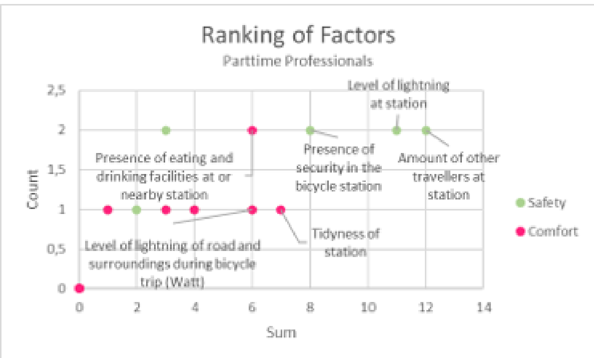



Figure C.5: Ranking Middle-aged part-time professionals

The middle-aged part-time professional interviewees are represented by two females (age 52 and 59) and one male (age 43). The gender distribution of this group found by Shelat et al. (2018) was 85% female. The ranking of this group consists of the following factors: 1) Amount of other travellers at station, 2) Level of lightning at station, 3) Presence of security in the bicycle parking, 4) Tidyness of station and 5) Presence of eating and drinking facilities at the station. The ranking of factors are illustrated in figure C.5. The three most important factors ranked by this group are related to safety, whereas the fourth and fifth place are regarding comfort.

## C.6 University students living alone

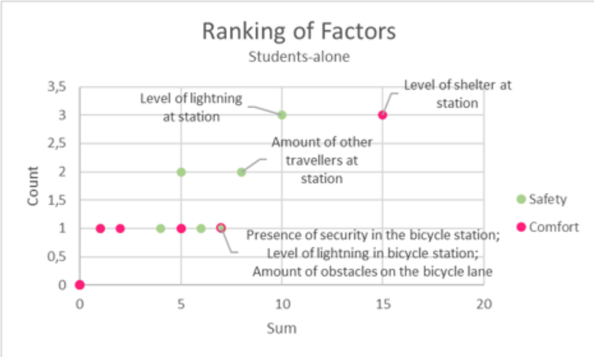



Figure C.6: Ranking University students living alone

Other than university students living with their parents, the students living alone are represented by one female and two males, to account for the total equal gender distribution. It is interesting to find that the student groups compared have a lot of comparative factors in their top five. The most important factors of this usergroups are 1) Level of shelter at station, 2) Level of lightning at station, 3) Amount of other travellers at station and 4) Presence of security in bicycle parking, level of lightning at bicycle parking and amount of obstacles on the bicycle lane. Since three factors have the same count and sum for the fourth ranking place, the fifth place is neglected. This is visualised in figure C.6. In this usergroup the safety and comfort categories are equally represented.

## C.7 Pensioners

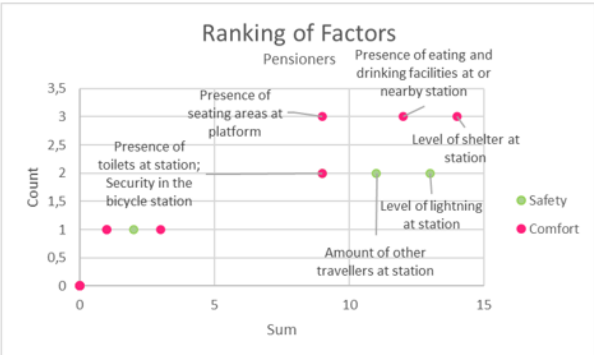



Figure C.7: Ranking Pensioners

The group with the smallest share in bicycle - train users are the pensioners. The gender of the three respondents were two males (age 66 and 71) and one female (age 67). The top five ranked factors of this usergroup included 1) Level of shelter at station, 2) Level of lightning at station 3) Presence of eating and drinking facilities at station, 4) Amount of other travellers at station and 5) Presence of seating areas at platform (figure C.7). As expected, more comfort factors obtained a higher ranking than average. This could be explained due to physical effort of travelling by bicycle and train. Remarkable is the homogeneity in the ranking of this group. Three factors are ranked by all respondents, and three other factors are ranked by two respondents.

# D

# Ngene code pilot

```
design
;alts = alt1, alt2
;rows = 12
;orth = seq
;model:
U(alt1) = b1 * light[0,1] + b2 * shelter[0,1] + b3*security[0,1] + b4*time[5,10,15,20]+
b5*price[0,0.5,1,1.5]/
U(alt2) = b1 * light + b2*shelter + b3*security + b4*time +b5*price
$
```

# E

# Pilot choice set generation

| Design | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Choice situation | alt1.light | alt1.shelter | alt1.security | alt1.time | alt1.price | alt2.light | alt2.shelter | alt2.security | alt2.time | alt2.price |
| 1 | 0 | 1 | 0 | 10 | 0.5 | 0 | 0 | 0 | 15 | 1.5 |
| 2 | 0 | 0 | 1 | 10 | 0 | 0 | 1 | 0 | 10 | 0.5 |
| 3 | 1 | 0 | 1 | 5 | 0.5 | 1 | 0 | 0 | 15 | 0 |
| 4 | 1 | 0 | 0 | 15 | 0 | 0 | 0 | 1 | 10 | 0 |
| 5 | 1 | 1 | 0 | 20 | 0 | 1 | 1 | 1 | 5 | 1 |
| 6 | 0 | 1 | 1 | 15 | 0.5 | 0 | 1 | 0 | 5 | 1 |
| 7 | 1 | 0 | 0 | 10 | 1.5 | 0 | 0 | 1 | 20 | 1 |
| 8 | 0 | 1 | 0 | 5 | 1 | 1 | 1 | 1 | 20 | 1.5 |
| 9 | 1 | 1 | 1 | 5 | 1 | 1 | 0 | 0 | 10 | 1.5 |
| 10 | 0 | 0 | 0 | 15 | 1.5 | 1 | 0 | 1 | 5 | 0.5 |
| 11 | 0 | 0 | 1 | 20 | 1 | 0 | 1 | 1 | 15 | 0.5 |
| 12 | 1 | 1 | 1 | 20 | 1.5 | 1 | 1 | 0 | 20 | 0 |

Figure E.1: Example of dominant choice set

| Correlations (Pearson Product Moment) | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Attribute | alt1.light | alt1.shelter | alt1.security | alt1.time | alt1.price | alt2.light | alt2.shelter | alt2.security | alt2.time | alt2.price |
| alt1.light | 1 | 0 | 0 | 0 | 0 | 0.333333 | -0.333333 | 0 | 0.149071 | -0.298142 |
| alt1.shelter | 0 | 1 | 0 | 0 | 0 | 0.333333 | 0.333333 | -0.333333 | 0 | 0.596285 |
| alt1.security | 0 | 0 | 1 | 0 | 0 | 0 | 0.333333 | -0.666667 | 0 | -0.298142 |
| alt1.time | 0 | 0 | 0 | 1 | 0 | -0.149071 | 0.447214 | 0.298142 | -0.266667 | -0.4 |
| alt1.price | 0 | 0 | 0 | 0 | 1 | 0.298142 | -0.149071 | 0.149071 | 0.466667 | 0.066667 |
| alt2.light | 0.333333 | 0.333333 | 0 | -0.149071 | 0.298142 | 1 | 0 | 0 | 0 | 0 |
| alt2.shelter | -0.333333 | 0.333333 | 0.333333 | 0.447214 | -0.149071 | 0 | 1 | 0 | 0 | 0 |
| alt2.security | 0 | -0.333333 | -0.666667 | 0.298142 | 0.149071 | 0 | 0 | 1 | 0 | 0 |
| alt2.time | 0.149071 | 0 | 0 | -0.266667 | 0.466667 | 0 | 0 | 0 | 1 | 0 |
| alt2.price | -0.298142 | 0.596285 | -0.298142 | -0.4 | 0.066667 | 0 | 0 | 0 | 0 | 1 |

Figure E.2: Correlations between attributes

# F

# Pilot Survey

## Stationskeuze

Voor mijn afstudeeronderzoek aan de TU Delft naar de keuze voor een treinstation heb ik de volgende vragenlijst opgesteld. Dit kan worden gebruikt om kleinere stations meer aantrekkelijk te maken en zo verschillende capaciteitsproblemen bij centrale stations op te lossen. In dit onderzoek ligt er een focus op factoren als veiligheidsgevoel en comfort. De uitkomsten van dit onderzoek geven meer inzicht in welke factoren van invloed zijn bij stationskeuze.

De vragenlijst bestaat uit drie onderdelen. Eerst vraag ik wat algemene vragen over persoonskenmerken. Deze vragen helpen mij om verschillende groepen fiets-trein reizigers te vinden. Vervolgens leg ik u 12 keuzes van stations voor en tot slot een aantal afsluitende vragen over uw keuzes.
Het invullen van deze vragenlijst zal ongeveer 5-7 minuten duren. De enquête is volledig anoniem en zal alleen voor dit onderzoek gebruikt worden. Het invullen van deze enquête is geheel vrijwillig en u kunt altijd uit de enquête stappen door de link af te sluiten.

Ik wil u alvast hartelijk bedanken voor uw medewerking.

Mocht u vragen hebben of op de hoogte wil blijven van mijn onderzoek kunt u contact opnemen met A.A.Barneveld@student.tudelft.nl

Vriendelijke groet,
Anne Barneveld

Persoonskenmerken

Wat is uw leeftijd?

- ◯ 17 jaar of jonger
- ◯ 18 t/m 24 jaar
- ◯ 25 t/m 34 jaar
- ◯ 35 t/m 64 jaar
- ◯ 65 jaar of ouder

Met welk geslacht identificeert u zichzelf?

- ◯ Man
- ◯ Vrouw
- ◯ Zeg ik liever niet
- ◯ Anders: ______

Wat is uw hoogst afgeronde opleiding?

- ◯ Geen
- ◯ Basisschool
- ◯ VMBO/MAVO/LBO
- ◯ HAVO/VWO
- ◯ MBO
- ◯ HBO/WO Bachelor
- ◯ WO Master
- ◯ PhD
- ◯ Zeg ik liever niet
- ◯ Anders: ______

Wat is de samenstelling van uw huishouden?

- ◯ Alleenstaand
- ◯ Tweepersoons zonder kinderen
- ◯ Gezin met 1 of meer thuiswonende kinderen
- ◯ Eenoudergezin met 1 of meer thuiswonende kinderen
- ◯ Studentenhuishouden
- ◯ Zeg ik liever niet
- ◯ Anders: ______

Wat is uw huidige werksituatie?

- ◯ Werkloos
- ◯ Fulltime (36 - 40 uur per week)
- ◯ Parttime (minder dan 36 uur per week)
- ◯ Student / Scholier
- ◯ Gepensioneerd
- ◯ Zeg ik liever niet

Hoeveel dagen reist u gemiddeld met de trein (voor Covid-19)?

- ◯ 4 dagen per week of vaker
- ◯ 1-3 dagen per week
- ◯ 1-4 dagen per maand
- ◯ Minder dan 1 dag per maand
- ◯ Nooit

Wat is meestal het doel van uw treinreis?

- ◯ Werk
- ◯ Studie/School
- ◯ Vrije tijd
- ◯ Zeg ik liever niet
- ◯ Anders: ______

Welk vervoermiddel gebruikt u meestal om vanuit huis bij het station te komen?

- ◯ Lopend
- ◯ Fiets
- ◯ Auto
- ◯ OV (bus, tram, metro)
- ◯ Zeg ik liever niet
- ◯ Anders: ____________

Wat is jaarlijks uw bruto inkomen?

- ◯ Geen inkomen
- ◯ Lager dan €10.000
- ◯ €10.000 - €30.000
- ◯ €30.000 - €50.000
- ◯ €50.000 - €70.000
- ◯ Meer dan €70.000
- ◯ Zeg ik liever niet / Weet ik niet

Wat zijn de vier cijfers van uw postcode?

Jouw antwoord

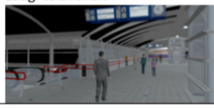
## Keuzesets

In dit onderdeel volgen 12 vragen. In deze vragen kiest u tussen twee stations om naartoe te fietsen. Het is gegeven dat u met de fiets naar het station komt. Vanaf het gekozen station reist u met de trein verder naar uw eindbestemming. Ieder te kiezen station heeft andere kenmerken. De kenmerken waar de te kiezen stations op verschillen zijn:
- Lichtintensiteit (schemering of nageboostst daglicht)
- Beschutting op het perron (Lage mate of hoge mate van beschutting. Lage beschutting betekent een voor en achter spatscherm en is de ruimte niet winddicht. Hoge beschutting betekent een volledig dichte ruimte. In beide opties is er genoeg ruimte en een dak)
- Beveiliging in de fietsenstalling (wel of geen menselijke beveiliging aanwezig)
- Fietstijd naar het station (de snelste route naar het station in minuten)
- Prijs van de fietsenstalling (euro per dag).

U mag aannemen dat verder de stations hetzelfde zijn.

Denkt u bij het maken van deze keuzes aan een dagelijkse rit in een herfstperiode. Waarbij het donker/schemering is en kans op neerslag en kou. U heeft geen grote bagage bij u. Denk hierbij aan een periode zonder Corona.

Keuzeset 1 *

| | Station A | Station B |
|---|---|---|
| **Licht** | Laag: Schemerlicht  | Hoog: Nabootsing Daglicht  |
| **Beschutting** | Hoog: Volledig dichte abri  | Laag: Alleen voor en achter spatscherm  |
| **Beveiliging Fietsenstalling** | Nee | Nee |
| **Fietstijd** | 10 minuten | 15 minuten |
| **Prijs Fietsenstalling** | €0,50 | €0,00 |

- Station A
- Station B

Keuzeset 2 *

| | Station A | Station B |
|---|---|---|
| **Licht** | Laag: Schemerlicht  | Hoog: Nabootsing Daglicht  |
| **Beschutting** | Laag: Alleen voor en achter spatscherm  | Hoog: Volledig dichte abri  |
| **Beveiliging Fietsenstalling** | Ja | Ja |
| **Fietstijd** | 10 minuten | 20 minuten |
| **Prijs Fietsenstalling** | €0,00 | €1,50 |

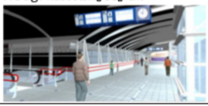
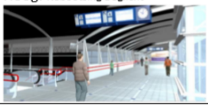
- Station A
- Station B

Keuzeset 3 *

| | Station A | Station B |
|---|---|---|
| **Licht** | Hoog: Nabootsing Daglicht  | Laag: Schemerlicht  |
| **Beschutting** | Laag: Alleen voor en achter spatscherm  | Hoog: Volledig dichte abri  |
| **Beveiliging Fietsenstalling** | Ja | Nee |
| **Fietstijd** | 5 minuten | 5 minuten |
| **Prijs Fietsenstalling** | €0,50 | €1,00 |

- Station A
- Station B

Keuzeset 4 *

| | Station A | Station B |
|---|---|---|
| **Licht** | Hoog: Nabootsing Daglicht  | Hoog: Nabootsing Daglicht  |
| **Beschutting** | Laag: Alleen voor en achter spatscherm  | Laag: Alleen voor en achter spatscherm  |
| **Beveiliging Fietsenstalling** | Nee | Ja |
| **Fietstijd** | 15 minuten | 5 minuten |
| **Prijs Fietsenstalling** | €0,00 | €0,50 |

◯ Station A

◯ Station B

Keuzeset 5 *

| | Station A | Station B |
|---|---|---|
| **Licht** | Hoog: Nabootsing Daglicht  | Laag: Schemerlicht  |
| **Beschutting** | Hoog: Volledig dichte abri  | Hoog: Volledig dichte abri  |
| **Beveiliging Fietsenstalling** | Nee | Ja |
| **Fietstijd** | 20 minuten | 15 minuten |
| **Prijs Fietsenstalling** | €0,00 | €0,50 |

◯ Station A

◯ Station B

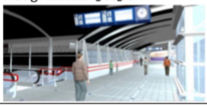
Keuzeset 6 *

| | Station A | Station B |
|---|---|---|
| **Licht** | Laag: Schemerlicht  | Hoog: Nabootsing Daglicht  |
| **Beschutting** | Hoog: Volledig dichte abri  | Hoog: Volledig dichte abri  |
| **Beveiliging Fietsenstalling** | Ja | Ja |
| **Fietstijd** | 15 minuten | 5 minuten |
| **Prijs Fietsenstalling** | €0,50 | €1,00 |

◯ Station A

◯ Station B

Keuzeset 7 *

| | Station A | Station B |
|---|---|---|
| **Licht** | Hoog: Nabootsing Daglicht  | Laag: Schemerlicht  |
| **Beschutting** | Laag: Alleen voor en achter spatscherm  | Laag: Alleen voor en achter spatscherm  |
| **Beveiliging Fietsenstalling** | Nee | Ja |
| **Fietstijd** | 10 minuten | 20 minuten |
| **Prijs Fietsenstalling** | €1,50 | €1,00 |

◯ Station A

◯ Station B

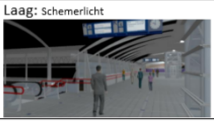
Keuzeset 8 *

| | Station A | Station B |
|---|---|---|
| **Licht** | Laag: Schemerlicht  | Hoog: Nabootsing Daglicht  |
| **Beschutting** | Hoog: Volledig dichte abri  | Laag: Alleen voor en achter spatscherm  |
| **Beveiliging Fietsenstalling** | Nee | Nee |
| **Fietstijd** | 5 minuten | 10 minuten |
| **Prijs Fietsenstalling** | €1,00 | €1,50 |

◯ Station A

◯ Station B

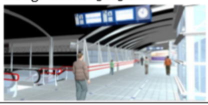
Keuzeset 9 *

| | Station A | Station B |
|---|---|---|
| **Licht** | Hoog: Nabootsing Daglicht  | Laag: Schemerlicht  |
| **Beschutting** | Hoog: Volledig dichte abri  | Laag: Alleen voor en achter spatscherm  |
| **Beveiliging Fietsenstalling** | Ja | Ja |
| **Fietstijd** | 5 minuten | 10 minuten |
| **Prijs Fietsenstalling** | €1,00 | €0,00 |

◯ Station A

◯ Station B

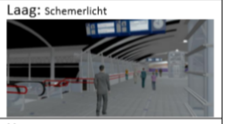
Keuzeset 10 *

| | Station A | Station B |
|---|---|---|
| Licht | Laag: Schemerlicht  | Hoog: Nabootsing Daglicht  |
| Beschutting | Laag: Alleen voor en achter spatscherm  | Hoog: Volledig dichte abri  |
| Beveiliging Fietsenstalling | Nee | Nee |
| Fietstijd | 15 minuten | 20 minuten |
| Prijs Fietsenstalling | €1,50 | €0,00 |

◯ Station A

◯ Station B

Keuzeset 11 *

| | Station A | Station B |
|---|---|---|
| Licht | Laag: Schemerlicht  | Laag: Schemerlicht  |
| Beschutting | Laag: Alleen voor en achter spatscherm  | Hoog: Volledig dichte abri  |
| Beveiliging Fietsenstalling | Ja | Nee |
| Fietstijd | 20 minuten | 10 minuten |
| Prijs Fietsenstalling | €1,00 | €0,50 |

◯ Station A

◯ Station B

Keuzeset 12 *

| | Station A | Station B |
|---|---|---|
| Licht | Hoog: Nabootsing Daglicht  | Laag: Schemerlicht  |
| Beschutting | Hoog: Volledig dichte abri  | Laag: Alleen voor en achter spatscherm  |
| Beveiliging Fietsenstalling | Ja | Nee |
| Fietstijd | 20 minuten | 15 minuten |
| Prijs Fietsenstalling | €1,50 | €1,50 |

◯ Station A

◯ Station B

**Afsluiting**

Heeft u vanaf uw woning twee of meer treinstations op maximaal 20 minuten fietsen?

- ◯ Ja
- ◯ Nee
- ◯ Weet ik niet / Zeg ik liever niet

Op welke factor(en) heeft u uw keuze in deze vragenlijst het meeste gebaseerd? Meerdere antwoorden zijn mogelijk.

- ☐ Licht op het station en perron
- ☐ Beschutting op het perron
- ☐ Beveiliging in de fietsenstalling
- ☐ Fietstijd naar het station
- ☐ Prijs
- ☐ Ik heb alle factoren gelijk overwogen

Heeft u vragen of opmerkingen?

Jouw antwoord

Figure F.1: Pilot Survey

# G

# Descriptive Statistics Pilot Survey

| **Age** | **Count** | **Percentage** |
|---|---|---|
| <17 | 3 | 8% |
| 18-24 | 9 | 24% |
| 25-34 | 10 | 26% |
| 35-64 | 11 | 29% |
| >65 | 5 | 13% |
| | | |
| **Gender** | | |
| Male | 20 | 53% |
| Female | 18 | 47% |
| | | |
| **Education** | | |
| None | 2 | 5% |
| Elementary | 1 | 3% |
| VMBO/MAVO | 2 | 5% |
| HAVO/VWO | 5 | 13% |
| MBO | 10 | 26% |
| HBO/WO Bachelor | 12 | 32% |
| WO Master | 5 | 13% |
| PhD | 1 | 3% |
| | | |
| **Household** | | |
| Single | 5 | 13% |
| Pair | 13 | 34% |
| Pair + kid(s) | 8 | 21% |
| Single + kid(s) | 2 | 5% |
| Student | 8 | 21% |
| Unknown | 2 | 5% |
| | | |
| **Trip Purpose** | | |
| Work | 11 | 29% |
| Study | 8 | 21% |
| Leisure | 19 | 50% |

| **Income** | **Count** | **Percentage** |
|---|---|---|
| None | 4 | 11% |
| <€10.000 | 6 | 16% |
| €10.000 - €30.000 | 7 | 18% |
| €30.000 - €50.000 | 10 | 26% |
| €50.000 - €70.000 | 5 | 13% |
| >€70.000 | 2 | 5% |
| Unknown | 4 | 11% |
| | | |
| **Social Participation** | | |
| None | 3 | 8% |
| Fulltime | 14 | 37% |
| Parttime | 9 | 24% |
| Student | 11 | 29% |
| Retired | 1 | 3% |
| | | |
| **Train Frequency** | | |
| >4 days/week | 9 | 24% |
| 1-3 days/week | 11 | 29% |
| 1-4 days/month | 10 | 26% |
| <1 days/month | 8 | 21% |
| Never | 0 | 0% |
| | | |
| **Access Mode** | | |
| Walking | 4 | 11% |
| Bicycle | 29 | 76% |
| Car | 2 | 5% |
| Public Transport | 3 | 8% |

Figure G.1: Descriptive statistics of respondents in pilot survey

# H

# Biogeme Model code pilot

```
[Choice]
CHOICE

[Beta]
// Name Value LowerBound UpperBound status (0=variable, 1=fixed)
BETA_LIGHT 0 -10000 10000 0
BETA_SHELTER 0 -10000 10000 0
BETA_SECURITY 0 -10000 10000 0
BETA_TIME 0 -10000 10000 0
BETA_PRICE 0 -10000 10000 0

[Utilities]
1 A av1 $NONE
2 B av2 $NONE

[GeneralizedUtilities]
// Id Name utility function
1 BETA_LIGHT * LIGHTA + BETA_SHELTER * SHELTERA + BETA_SECURITY * SECURITYA +
BETA_TIME * TTA + BETA_PRICE * TCA
2 BETA_LIGHT * LIGHTB + BETA_SHELTER * SHELTERB + BETA_SECURITY * SECURITYB +
BETA_TIME * TTB + BETA_PRICE * TCB

[PanelData]
ID

[Expressions]
av1 = 1
av2 = 1

[Model]
$MNL
```

# NGENE code final

```
design
;alts = alt1, alt2
;rows = 10
;eff = (mnl,d)
;model:
U(alt1) = b1[0.941] * light[0,1] + b2[0.306] * shelter[0,1] + b3[0.761] *security[0,1] + b4[-0.126] *time[5,7.5,10]+ b5[-1.16] *price[0,0.5,1]/
U(alt2) = b1 * light + b2*shelter + b3*security + b4*time +b5*price
$
```

# J

# MNL Probabilities

| Choice situation | Station A | Station B |
|---|---|---|
| 1 | 0.341639 | 0.658361 |
| 2 | 0.233977 | 0.766023 |
| 3 | 0.474062 | 0.525938 |
| 4 | 0.500998 | 0.499002 |
| 5 | 0.362309 | 0.637691 |
| 6 | 0.456113 | 0.543887 |
| 7 | 0.73835 | 0.26165 |
| 8 | 0.597245 | 0.402755 |
| 9 | 0.445715 | 0.554285 |
| 10 | 0.846966 | 0.153034 |

Table J.1: MNL probabilities

| Prior | $\beta_{light}$ | $\beta_{shelter}$ | $\beta_{security}$ | $\beta_{time}$ | $\beta_{costs}$ |
|---|---|---|---|---|---|
| Fixed prior value | 0.941 | 0.306 | 0.761 | -0.126 | -1.16 |
| Sp estimates | 4.06 | 23.15 | 4.59 | 9.05 | 3.58 |
| Sp t-ratios | 0.97 | 0.41 | 0.92 | 0.65 | 1.03 |

Table J.2: MNL efficiency measures

# K

# Final Survey

*Introductie.*

Voor mijn afstudeeronderzoek aan de TU Delft heb ik deze vragenlijst opgesteld. Mijn onderzoek gaat over de keuze van reizigers voor een bepaald treinstation. Dit kan worden gebruikt om verschillende capaciteitsproblemen van centrale stations op te lossen. In dit onderzoek ligt de focus op factoren als veiligheidsgevoel en comfort. De uitkomsten van dit onderzoek geven meer inzicht in welke factoren van invloed zijn bij stationskeuze.
De vragenlijst bestaat uit drie onderdelen. Eerst stel ik u een aantal algemene vragen over uw persoonskenmerken. Dit helpt mij om verschillende groepen fiets-trein reizigers te onderscheiden. Vervolgens leg ik u 10 keuzes voor van stations. Tot slot stel ik enkele afsluitende vragen.

Het invullen van deze vragenlijst duurt ongeveer **7** minuten. De enquête is **volledig anoniem** en zal alleen voor dit onderzoek gebruikt worden. Het invullen van deze enquête is geheel vrijwillig en u kunt altijd uit de enquête stappen door de link af te sluiten.

Ik wil u alvast hartelijk bedanken voor uw medewerking.
Mocht u vragen hebben of op de hoogte wil blijven van mijn onderzoek kunt u contact opnemen via A.A.Barneveld@student.tudelft.nl

Vriendelijke groet,
Anne Barneveld

*Vraag 1.* Wat was uw leeftijd in 2019 (pre-Covid)?

- 17 jaar of jonger
- 18 t/m 24 jaar
- 25 t/m 34 jaar
- 35 t/m 64 jaar
- 65 jaar of ouder

*Vraag 2.* Met welk geslacht identificeert u zich?

- Man
- Vrouw
- Zeg ik liever niet
- Anders, namelijk:

*Vraag 3.* Wat was in 2019 (pre-Covid) uw hoogst afgeronde opleiding (of vergelijkbaar)?

- Geen
- Basisschool
- VMBO/MAVO/LBO
- HAVO/VWO
- MBO
- HBO / WO Bachelor / Kandidaats
- WO Master / Doctoraal
- PhD
- Zeg ik liever niet
- Anders, namelijk:

*Vraag 4.* Wat was in 2019 (pre-Covid) de samenstelling van uw huishouden?

- Alleenstaand zonder (thuiswonende) kinderen
- Tweepersoons zonder (thuiswonende) kinderen
- Tweeoudergezin met 1 of meer thuiswonende kinderen
- Eenoudergezin met 1 of meer thuiswonende kinderen
- Studentenhuishouden
- Zeg ik liever niet
- Anders, namelijk:

*Vraag 5.* Wat was uw werksituatie in 2019 (pre-Covid)?

- Werkloos
- Fulltime (36 - 40 uur per week)
- Parttime (minder dan 36 uur per week)
- Student / Scholier
- Gepensioneerd
- Zeg ik liever niet
- Anders, namelijk:

*Vraag 6.* Hoeveel dagen reisde u gemiddeld met de trein in 2019 (pre-Covid)?

- 4 dagen per week of vaker
- 1 t/m 3 dagen per week
- 1 t/m 4 dagen per maand
- Minder dan 1 dag per maand
- Nooit (ga verder naar vraag 9)

*Vraag 7.* Wat was meestal het doel van uw treinreis in 2019 (pre-Covid)?

- Werk
- Studie / School
- Vrije tijd
- Zeg ik liever niet
- Anders, namelijk:

*Vraag 8.* Welk vervoermiddel gebruikte u vooral om vanuit uw huis bij het station te komen (pre-Covid)?

- Lopend
- Fiets
- Auto
- OV (bus, tram, metro)
- Zeg ik liever niet
- Anders, namelijk:

*Vraag 9.* Wat was uw jaarlijks bruto inkomen in 2019 (pre-Covid)?

- Geen inkomen
- Lager dan € 10.000
- € 10.000 - € 30.000
- € 30.000 - € 50.000
- € 50.000 - € 70.000
- Meer dan € 70.000
- Zeg ik liever niet / Weet ik niet

*Vraag 10.* Wat zijn de vier cijfers van uw Postcode (pre-Covid)?

*Keuzesets.*

Nu volgen 10 vragen. In deze vragen kiest u steeds tussen twee denkbeeldige stations.
Ik vraag u om u bij iedere keuze de volgende situatie in te beelden:
- u gaat steeds met de fiets naar het station.
- het gaat om een dagelijkse rit in de herfst, waarbij het **donker/schemerig** is en er **kans** is op **neerslag en kou**.
- u reist in een periode zonder Corona zonder grote bagage.
- u reist vanaf het gekozen station met de trein verder naar uw eindbestemming. De treinreis (zoals duur, overstap en kosten) wordt niet beïnvloed door uw stationskeuze.

Ieder te kiezen station heeft andere kenmerken. De stations verschillen in:
- **Lichtintensiteit** (schemerlicht of nagebootst daglicht)
- **Beschutting** op het perron (Lage mate of hoge mate van beschutting. Lage beschutting betekent een voor en achter spatscherm en is de ruimte niet winddicht. Hoge beschutting betekent een volledig dichte ruimte. In beide opties is er genoeg ruimte en een dak)
- **Beveiliging** in de fietsenstalling (wel of geen menselijke beveiliging aanwezig)
- **Fietstijd** naar het station (de snelste route naar het station in minuten)
- **Prijs** van de fietsenstalling (euro per dag).

*Keuzeset 1.*

| Keuzeset 1 | **Station A** | **Station B** |
|---|---|---|
| **Licht** | Laag: Schemerlicht  | Hoog: Nabootsing Daglicht  |
| **Beschutting** | Hoog: Volledig dichte abri  | Laag: Alleen voor en achter spatscherm  |
| **Beveiliging Fietsenstalling** | Ja | Nee |
| **Fietstijd** | 10 minuten | 5 minuten |
| **Prijs Fietsenstalling / dag** | €0,00 | €1,00 |

Welk station heeft uw voorkeur?

Station A

Station B

*Keuzeset 2.*

| Keuzeset 2 | **Station A** | **Station B** |
|---|---|---|
| **Licht** | Hoog: Nabootsing Daglicht  | Laag: Schemerlicht  |
| **Beschutting** | Hoog: Volledig dichte abri  | Laag: Alleen voor en achter spatscherm  |
| **Beveiliging Fietsenstalling** | Nee | Ja |
| **Fietstijd** | 7,5 minuten | 7,5 minuten |
| **Prijs Fietsenstalling / dag** | €0,00 | €0,50 |

Welk station heeft uw voorkeur?

Station A

Station B

*Keuzeset 3.*

| Keuzeset 3 | **Station A** | **Station B** |
|---|---|---|
| **Licht** | Hoog: Nabootsing Daglicht  | Laag: Schemerlicht  |
| **Beschutting** | Laag: Alleen voor en achter spatscherm  | Hoog: Volledig dichte abri  |
| **Beveiliging Fietsenstalling** | Ja | Nee |
| **Fietstijd** | 5 minuten | 10 minuten |
| **Prijs Fietsenstalling /dag** | €1,00 | €0,00 |

Welk station heeft uw voorkeur?

Station A

Station B

*Keuzeset 4.*

| Keuzeset 4 | **Station A** | **Station B** |
|---|---|---|
| **Licht** | Hoog: Nabootsing Daglicht  | Laag: Schemerlicht  |
| **Beschutting** | Hoog: Volledig dichte abri  | Laag: Alleen voor en achter spatscherm  |
| **Beveiliging Fietsenstalling** | Nee | Ja |
| **Fietstijd** | 5 minuten | 7,5 minuten |
| **Prijs Fietsenstalling /dag** | €1,00 | €0,00 |

Welk station heeft uw voorkeur?

Station A

Station B

*Keuzeset 5.*

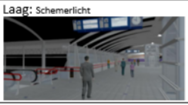
| Keuzeset 5 | **Station A** | **Station B** |
|---|---|---|
| **Licht** | Laag: Schemerlicht  | Hoog: Nabootsing Daglicht  |
| **Beschutting** | Hoog: Volledig dichte abri  | Laag: Alleen voor en achter spatscherm  |
| **Beveiliging Fietsenstalling** | Ja | Nee |
| **Fietstijd** | 5 minuten | 10 minuten |
| **Prijs Fietsenstalling / dag** | €0,50 | €0,50 |

Welk station heeft uw voorkeur?

Station A

Station B

*Keuzeset 6.*

| Keuzeset 6 | **Station A** | **Station B** |
|---|---|---|
| **Licht** | Hoog: Nabootsing Daglicht  | Laag: Schemerlicht  |
| **Beschutting** | Laag: Alleen voor en achter spatscherm  | Hoog: Volledig dichte abri  |
| **Beveiliging Fietsenstalling** | Nee | Ja |
| **Fietstijd** | 10 minuten | 5 minuten |
| **Prijs Fietsenstalling / dag** | €0,00 | €1,00 |

Welk station heeft uw voorkeur?

Station A

Station B

*Keuzeset 7.*

| Keuzeset 7 | **Station A** | **Station B** |
|---|---|---|
| **Licht** | Laag: Schemerlicht  | Hoog: Nabootsing Daglicht  |
| **Beschutting** | Laag: Alleen voor en achter spatscherm  | Hoog: Volledig dichte abri  |
| **Beveiliging Fietsenstalling** | Ja | Nee |
| **Fietstijd** | 7,5 minuten | 7,5 minuten |
| **Prijs Fietsenstalling / dag** | €0,50 | €0,00 |

Welk station heeft uw voorkeur?

Station A

Station B

*Keuzeset 8.*

| Keuzeset 8 | **Station A** | **Station B** |
|---|---|---|
| **Licht** | Hoog: Nabootsing Daglicht  | Laag: Schemerlicht  |
| **Beschutting** | Laag: Alleen voor en achter spatscherm  | Hoog: Volledig dichte abri  |
| **Beveiliging Fietsenstalling** | Ja | Nee |
| **Fietstijd** | 10 minuten | 5 minuten |
| **Prijs Fietsenstalling / dag** | €1,00 | €0,00 |

Welk station heeft uw voorkeur?

Station A

Station B

*Keuzeset 9.*

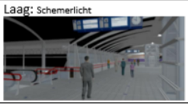
| Keuzeset 9 | **Station A** | **Station B** |
|---|---|---|
| **Licht** | Laag: Schemerlicht  | Hoog: Nabootsing Daglicht  |
| **Beschutting** | Laag: Alleen voor en achter spatscherm  | Hoog: Volledig dichte abri  |
| **Beveiliging Fietsenstalling** | Nee | Ja |
| **Fietstijd** | 5 minuten | 10 minuten |
| **Prijs Fietsenstalling / dag** | €0,00 | €1,00 |

Welk station heeft uw voorkeur?

Station A

Station B

*Keuzeset 10.*

| Keuzeset 10 | **Station A** | **Station B** |
|---|---|---|
| **Licht** | Laag: Schemerlicht  | Hoog: Nabootsing Daglicht  |
| **Beschutting** | Hoog: Volledig dichte abri  | Laag: Alleen voor en achter spatscherm  |
| **Beveiliging Fietsenstalling** | Nee | Ja |
| **Fietstijd** | 7,5 minuten | 5 minuten |
| **Prijs Fietsenstalling / dag** | €0,50 | €0,50 |

Welk station heeft uw voorkeur?

- Station A
- Station B

*Vraag 11.* U heeft zojuist keuzes gemaakt uit verschillende stations. Welke factoren waren voor u het belangrijkst? Meerdere antwoorden zijn mogelijk.

- Licht op het station
- Beschutting op het perron
- Beveiliging in de fietsenstalling
- Fietstijd naar het station
- Prijs van de fietsenstalling
- Alle factoren gelijk overwogen

*Vraag 12.* Heeft u vanaf uw woning twee of meer treinstations op maximaal 15 minuten fietsen (pre-Covid)?

- Ja
- Nee
- Weet ik niet / Zeg ik liever niet

*Vraag 13.* Op wat voor soort fiets fietst u meestal naar het station?

- Gewone fiets met een (geschatte) waarde lager dan €150
- Gewone fiets met een (geschatte) waarde hoger dan €150 (niet elektrisch)
- Sportfiets (racefiets of mountainbike)
- Elektrische fiets
- Ik heb geen fiets
- Ik fiets nooit naar het station
- Anders, namelijk:

*Vraag 14.* Heeft u vragen of opmerkingen?

L

# Estimation outcomes interaction effects

Significant results with a p-value < 0.05 are shown in green.

| Socio-demographic characteristics | Lightning | P-value | Shelter | P-value | Security | P-value | Cycle time | P-value | Parking Costs | P-value |
|---|---|---|---|---|---|---|---|---|---|---|
| **GENDER** | | | | | | | | | | |
| *Female* | 0.746 | 0.00 | 0.366 | 0.00 | 0.592 | 0.00 | -0.104 | 0.00 | -0.868 | 0.00 |
| *Male* | 0.584 | 0.03 | 0.368 | 0.98 | 0.515 | 0.43 | -0.144 | 0.04 | -1.349 | 0.00 |
| **AGE** | | | | | | | | | | |
| *<17 years* | 0.342 | 0.60 | 0.471 | 0.48 | 0.783 | 0.35 | -0.086 | 0.93 | -0.734 | 0.60 |
| *18/24* | 0.769 | 0.28 | 0.597 | 0.01 | 0.548 | 0.87 | -0.196 | 0.00 | -2.13 | 0.00 |
| *25/35 years* | 0.784 | 0.29 | 0.292 | 0.87 | 0.616 | 0.53 | -0.142 | 0.08 | -1.49 | 0.00 |
| *35/64 years* | 0.631 | 0.00 | 0.310 | 0.00 | 0.530 | 0.00 | -0.091 | 0.00 | -0.549 | 0.00 |
| *>65* | 0.680 | 0.84 | 0.300 | 0.95 | 0.597 | 0.78 | -0.111 | 0.59 | -0.104 | 0.13 |
| **EDUCATION** | | | | | | | | | | |
| *None* | 1.106 | 0.06 | 2.752 | 0.07 | 1.63 | 0.05 | 0.058 | 0.00 | 1.289 | 0.19 |
| *Basis* | 0.418 | 0.10 | 0.643 | 0.46 | 0.785 | 0.51 | -0.133 | 0.82 | -1.27 | 0.75 |
| *VMBO* | 0.444 | 0.06 | 0.255 | 0.22 | 0.427 | 0.45 | -0.039 | 0.00 | -0.435 | 0.00 |
| *HAVO/VWO* | 0.781 | 0.95 | 0.323 | 0.29 | 0.654 | 0.53 | -0.118 | 0.42 | -1.317 | 0.46 |
| *MBO* | 0.517 | 0.06 | 0.294 | 0.15 | 0.432 | 0.30 | -0.056 | 0.00 | -0.603 | 0.00 |
| *HBO/WO Bachelor* | 0.787 | 0.00 | 0.436 | 0.00 | 0.557 | 0.00 | -0.142 | 0.00 | -1.46 | 0.00 |
| *WO Master* | 0.692 | 0.52 | 0.391 | 0.68 | 0.634 | 0.65 | -0.205 | 0.04 | -1.04 | 0.03 |
| *PhD* | -0.056 | 0.95 | 0.302 | 0.28 | 0.787 | 0.03 | -0.118 | 0.04 | -2.707 | 0.17 |
| **HOUSEHOLD** | | | | | | | | | | |
| *Single* | 0.614 | 0.50 | 0.427 | 0.06 | 0.498 | 0.32 | -0.098 | 0.35 | -1.054 | 0.86 |
| *Two-parents no kids* | 0.591 | 0.31 | 0.331 | 0.29 | 0.596 | 0.75 | -0.079 | 0.05 | -0.655 | 0.03 |
| *Two-parents 1+ kids* | 0.723 | 0.00 | 0.220 | 0.00 | 0.636 | 0.00 | -0.127 | 0.00 | -1.02 | 0.00 |
| *Single 1+ kids* | 0.376 | 0.22 | 0.649 | 0.04 | 0.152 | 0.01 | -0.147 | 0.71 | -0.598 | 0.07 |
| *Student* | 0.956 | 0.15 | 0.674 | 0.00 | 0.514 | 0.27 | -0.254 | 0.00 | -2.39 | 0.00 |
| **PARTICIPATION** | | | | | | | | | | |
| *No job* | 0.853 | 0.36 | 0.463 | 0.45 | 0.391 | 0.40 | -0.016 | 0.05 | -0.273 | 0.01 |
| *Fulltime* | 0.652 | 0.00 | 0.327 | 0.00 | 0.564 | 0.00 | -0.117 | 0.00 | -0.976 | 0.00 |
| *Parttime* | 0.581 | 0.59 | 0.239 | 0.33 | 0.511 | 0.68 | -0.100 | 0.49 | -0.605 | 0.03 |
| *Student* | 0.838 | 0.17 | 0.603 | 0.01 | 0.620 | 0.64 | -0.211 | 0.00 | -2.205 | 0.00 |
| *Pension* | 0.436 | 0.03 | 0.367 | 0.03 | 0.647 | 0.18 | -0.067 | 0.28 | -0.118 | 0.00 |
| **DAYS TRAIN** | | | | | | | | | | |
| *4+ per week* | 0.547 | 0.38 | 0.381 | 0.51 | 0.576 | 0.55 | -0.145 | 0.88 | -1.368 | 0.66 |
| *1-3 per week* | 0.681 | 0.00 | 0.449 | 0.00 | 0.513 | 0.00 | -0.151 | 0.00 | -1.300 | 0.00 |
| *1-4 per month* | 0.707 | 0.85 | 0.388 | 0.54 | 0.488 | 0.83 | -0.116 | 0.22 | -1.145 | 0.37 |
| *<1 per month* | 0.930 | 0.13 | 0.190 | 0.04 | 0.800 | 0.07 | -0.087 | 0.04 | -0.571 | 0.00 |
| *Never* | 0.19 | 0.44 | 0.431 | 0.04 | 0.368 | 0.13 | -0.036 | 0.29 | -0.175 | 0.59 |
| **TRIP PURPOSE** | | | | | | | | | | |
| *Work* | 0.642 | 0.64 | 0.297 | 0.23 | 0.547 | 0.84 | -0.143 | 0.17 | -0.950 | 0.80 |
| *School* | 0.671 | 0.89 | 0.423 | 0.87 | 0.560 | 0.96 | -0.112 | 0.97 | -1.726 | 0.00 |
| *Leisure* | 0.694 | 0.00 | 0.402 | 0.00 | 0.568 | 0.00 | -0.111 | 0.00 | -0.989 | 0.00 |
| **ACCESS MODE** | | | | | | | | | | |
| *Walk* | 0.924 | 0.02 | 0.433 | 0.40 | 0.702 | 0.25 | -0.135 | 0.55 | -0.905 | 0.10 |
| *Bicycle* | 0.579 | 0.00 | 0.340 | 0.00 | 0.546 | 0.00 | -0.117 | 0.00 | -1.220 | 0.00 |
| *Car* | 0.561 | 0.81 | 0.058 | 0.04 | 0.499 | 0.54 | -0.091 | 0.46 | -0.669 | 0.01 |
| *PT* | 0.889 | 0.08 | 0.682 | 0.01 | 0.513 | 0.77 | -0.165 | 0.14 | -1.033 | 0.35 |
| **INCOME/YEAR** | | | | | | | | | | |
| *None* | 0.575 | 0.42 | 0.523 | 0.09 | 0.879 | 0.08 | -0.199 | 0.13 | -2.320 | 0.00 |
| *<€10.000* | 0.830 | 0.50 | 0.551 | 0.01 | 0.558 | 0.89 | -0.153 | 0.49 | -1.755 | 0.00 |
| *€10.000-€30.000* | 0.563 | 0.30 | 0.248 | 0.89 | 0.487 | 0.73 | -0.056 | 0.02 | -0.747 | 0.28 |
| *€30.000-€50.000* | 0.724 | 0.00 | 0.233 | 0.00 | 0.539 | 0.00 | -0.131 | 0.00 | -0.960 | 0.00 |
| *€50.000-€70.000* | 0.639 | 0.66 | 0.283 | 0.64 | 0.650 | 0.55 | -0.146 | 0.69 | -0.818 | 0.55 |
| *>€70.000* | 0.705 | 0.93 | 0.432 | 0.18 | 0.687 | 0.48 | -0.159 | 0.54 | -0.451 | 0.04 |
| *Unknown* | 0.580 | 0.43 | 0.558 | 0.04 | 0.356 | 0.32 | -0.067 | 0.10 | -0.516 | 0.04 |
| **TWO STATIONS** | | | | | | | | | | |
| *Yes* | 0.636 | 0.00 | 0.419 | 0.00 | 0.530 | 0.00 | -0.127 | 0.00 | -1.24 | 0.00 |
| *No* | 0.714 | 0.44 | 0.299 | 0.14 | 0.590 | 0.53 | -0.111 | 0.67 | -0.923 | 0.01 |
| **BICYCLE TYPE** | | | | | | | | | | |
| *Bicycle <€150* | 0.834 | 0.00 | 0.519 | 0.00 | 0.525 | 0.00 | -0.169 | 0.00 | -1.62 | 0.00 |
| *Bicycle >€150* | 0.695 | 0.29 | 0.134 | 0.00 | 0.624 | 0.42 | -0.099 | 0.01 | -0.771 | 0.00 |
| *Sport* | 0.018 | 0.00 | 0.329 | 0.42 | 0.628 | 0.73 | -0.001 | 0.01 | -1.065 | 0.17 |
| *Electric* | 0.385 | 0.00 | 0.346 | 0.13 | 0.640 | 0.46 | -0.066 | 0.00 | -0.457 | 0.00 |
| *No bicycle* | 0.043 | 0.90 | 0.991 | 0.05 | 0.235 | 0.48 | -0.003 | 0.44 | -0.306 | 0.48 |
| *Never to station* | 0.742 | 0.67 | 0.297 | 0.14 | 0.581 | 0.80 | -0.136 | 0.47 | -0.264 | 0.00 |

Figure L.1: Interaction effects on socio-demographic characteristics including their p-values

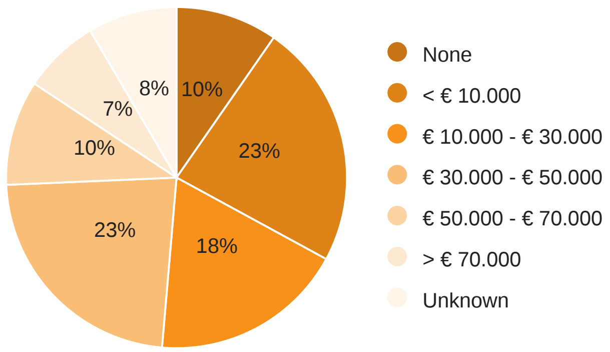
# M

# Descriptive Statistics

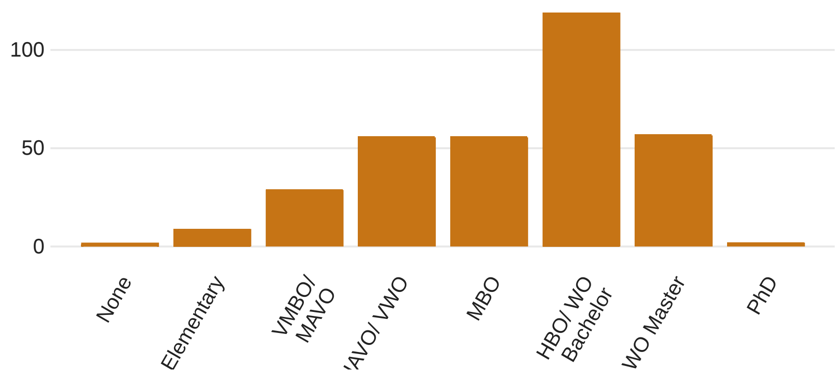
## M.1 Final survey

# Descriptive Statistics

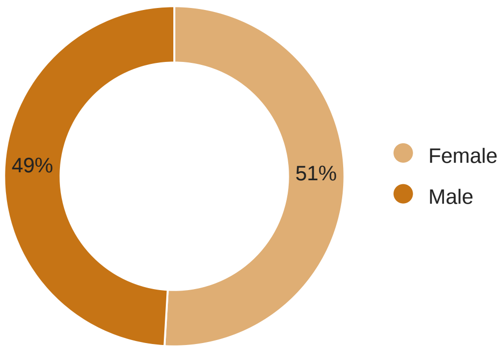
## Income



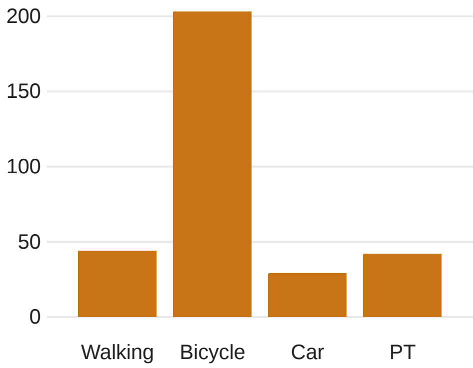

## Education Level



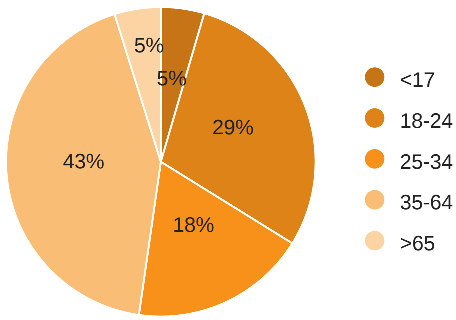

## Gender



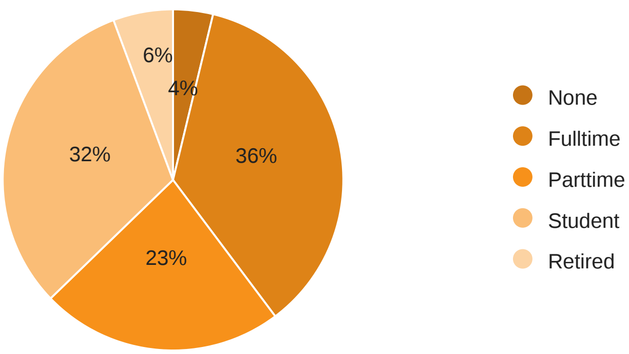

## Access Mode



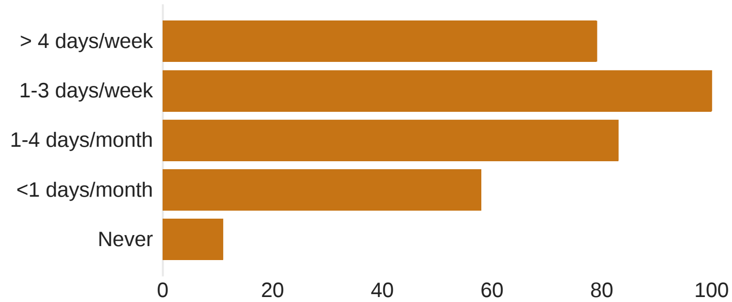

## Age



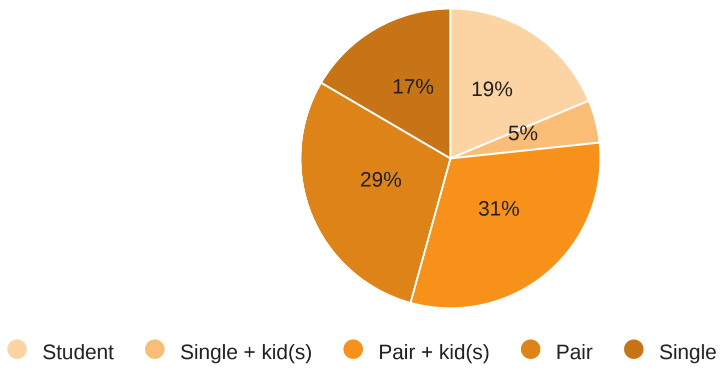

## Participation



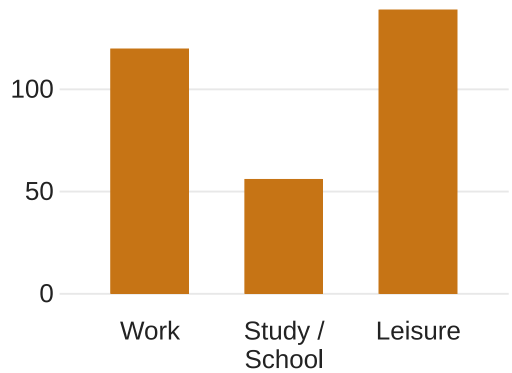

## Train Frequency




## Household




## Trip Purpose



## M.2 Geographical spread of respondents

Figure M.1: Geographical location of respondents

# N

# Matlab Code

```
% MODEL MNL / LATENT CLASS CHOICE MODEL
% A. A. Barneveld | 3 - 3- 2022
%clear;close;clc;

global VARS COVAR Y AV NCS NALT NAT PANEL NIND

%% Load the data set
df = readtable('finaldat4_4.dat');

%% Define the global variables
NCS = height(df);                % No. Choice Observations
NIND = length(unique(df.ID));    % No. Respondents
NALT = 2;                        % No. Alternatives
NAT = 5;                         % No. Attributes
Y = dummyvar(df.CHOICE);         % Indicator array with the choices
AV = ones(NCS,NALT);             % Availability

%% Identify the columns in the dataframe df used for estimation.
% VARS is a dataframe containing the attributes of the alternatives
% COVAR is a dataframe containing covariates, such as age, income, etc. used in the membership model
VARS = df(:,{'LightA','ShelterA','SecurityA','TimeA','CostsA','LightB','ShelterB','SecurityB','TimeB','CostsB'});
COVAR = df(:,{'SEX1','WORK2','WORK3','WORK5','WORK4','AGE','INC','TP','AM','EDU', 'BT', 'TS1'});
%COVAR = df(:,{'Education', 'Gender','Age', 'Income', 'biketype', 'Household', 'Participation', 'Daystrain', 'Trippurpose', 'Accesmode', 'twostations'});

%% Define the panel structure of the data
% PANEL is a NCS x 2 array with the start & end row numbers belonging to an individual
[~,T]=count_unique(df.ID); % Count the number of choice task per respondent
TT = [0;T(1:end-1)];
PANEL = [cumsum(TT)+1 cumsum(TT)+T]; % Construct the PANEL variable

%% Set the hyperparameters of the optimisation algorithm | Do not change
paramtol = 1e-6;     % parameter tolerance
lltol = 1e-6;        % objective function tolerance
```

```
options = optimset('LargeScale','off','Display','iter','GradObj','off',...
'MaxFunEvals',10000,'MaxIter',1000,'TolX',paramtol,'TolFun',lltol,'DerivativeCheck','off');

%% Select the model to estimate
model_names = {'1-class RUM-MNL','2-class LC model: RUM-MNL RUM-MNL. Cross sectional','2-class LC model: RUM-MNL RUM-MNL. Panel'};
% 1: 1-class RUM-MNL
% 3: 2-class LC model: RUM-MNL RUM-MNL. Panel

model = 3;  % Change to estimate a different model
disp(['Estimate model ',num2str(model),': ',model_names{model}]); % Print to the command window which model is estimated
tic;        % Starts a timer to monitor estimation time
if model == 1
    % Specify the variable names for model 1
    param_names = {'Light','Shelter','Security','Time', 'Costs'};
    % Starting values
    coef = zeros(size(param_names,2),1);
    % Estimate a RUM-MNL model
    [paramhat,fval,exitflag,output,grad,hessian] = fminunc(@AnneRumMnl,coef,options);

elseif model == 3
    % Specify the variable names for model 3
    param_names = {'Light1','Shelter1','Security1','Time1','Costs1','Light2','Shelter2','Security2','Time2','Costs2','delta_2','gender_2','student_2','ag
    % Starting values
    coef = 0.1.*rand(size(param_names,2),1);
     % Estimate LC_RUM_RUM-MNL. Cross sectional
    [paramhat,fval,exitflag,output,grad,hessian] = fminunc(@Loglik_LC_2,coef,options);
end

% Display and save the output
disp(' ');
save_results_to_file = 2;   % 1: save results to txt file; 2 do not save results to txt file
writeoutput(paramhat,fval,exitflag,grad,hessian,output,AV,NCS,NIND,param_names,model,model_names{model},lltol,paramtol,save_results_to_file);
```

```
function [ ll ] = loglik_LC_2_RUM_RUM_panel(coef)
% This function computes the log-likelihood of a LC RUM-MNL RUM-MNL model
% (two classes) panel
% Set global variables
global NCS NAT NIND AV Y VARS COVAR PANEL

%% Get parameters
coef_s1 = coef(1:NAT);          % parameters of class 1
coef_s2 = coef(NAT+1:2*NAT);    % parameters of class 2

% Class coefficients
delta_s1 = 0;
delta_s2 = coef(end-3);         % class membership parameter
gender_2 = coef(end-2);           % class membership parameter
student_2 = coef(end-1);
age_2 = coef(end);
%% Class probabilities
r1 = repmat(delta_s1,[NCS 1]);
r2 = repmat(delta_s2,[NCS 1]) + gender_2.*COVAR.SEX1 + student_2.*COVAR.WORK4 + age_2.*COVAR.AGE ;%+ inc_2.*COVAR.EDU;
pi_s1 = exp(r1)./(exp(r1)+ exp(r2));
pi_s2 = exp(r2)./(exp(r1)+ exp(r2));

%% Compute observed utilities for each class
% RUM CLASS 1
v1_s1 = coef_s1(1).* VARS.LightA + coef_s1(2).*VARS.ShelterA + coef_s1(3).*VARS.SecurityA + coef_s1(4).*VARS.TimeA +coef_s1(5).*VARS.CostsA ;
v2_s1 = coef_s1(1).* VARS.LightB + coef_s1(2).*VARS.ShelterB + coef_s1(3).*VARS.SecurityB + coef_s1(4).*VARS.TimeB +coef_s1(5).*VARS.CostsB ;
v_s1 = [v1_s1 v2_s1] ;

% To avoid numerical overflow
v_s1(v_s1>700) =  700;
v_s1(v_s1<-700)= -700;

% RUM CLASS 2
```

```
v1_s2 = coef_s2(1).* VARS.LightA + coef_s2(2).*VARS.ShelterA + coef_s2(3).*VARS.SecurityA + coef_s2(4).*VARS.TimeA +coef_s2(5).*VARS.CostsA ;
v2_s2 = coef_s2(1).* VARS.LightB + coef_s2(2).*VARS.ShelterB + coef_s2(3).*VARS.SecurityB + coef_s2(4).*VARS.TimeB +coef_s2(5).*VARS.CostsB ;
v_s2 = [v1_s2 v2_s2] ;

% To avoid numerical overflow
v_s2(v_s2>700) =  700;
v_s2(v_s2<-700)= -700;

%% Take exponent of the observed utilities / regrets
ev_s1 = exp(v_s1);
ev_s2 = exp(v_s2);

%% Compute the probability of observing each choice given the set of parameters
p_s1  = sum(Y.*ev_s1,2)./sum(AV.*ev_s1,2);
p_s2  = sum(Y.*ev_s2,2)./sum(AV.*ev_s2,2);


%% Panel probabilities
p_s1_seq = zeros(NIND,1);
p_s2_seq = zeros(NIND,1);
for n = 1:NIND
    p_s1_seq(n) = prod(p_s1(PANEL(n,1):PANEL(n,2))); % take the product over the choice probs for each individual cond. on class 1
    p_s2_seq(n) = prod(p_s2(PANEL(n,1):PANEL(n,2))); % take the product over the choice probs for each individual cond. on class 2
end

%% Construct class member probabilities of size NIND x 1 (thus per individual)
pi_s1_ind = pi_s1(PANEL(:,1));  % NIND x 1
pi_s2_ind = pi_s2(PANEL(:,1));  % NIND x 1

%% Unconditional choice probabilities
P_LC = pi_s1_ind.*p_s1_seq + pi_s2_ind.*p_s2_seq;

%% Compute the log-likelihood
ll=-sum(log(P_LC));  % Negative since neg of ll is minimized
```

# O

# Other Latent Class Models and Weights

| | Class 1 | Class 2 | Class 3 |
|---|---|---|---|
| **Light** | 0,794 | 2,237 | 0,546 |
| **Shelter** | 0,179 | 0,609 | 0,864 |
| **Security** | 0,785 | 1,097 | 0,565 |
| **Time** | -0,011 | -0,737 | -0,205 |
| **Costs** | 0,048 | -2,194 | -2,954 |
| | | | |
| delta | 0 | -1,386 | 0,949 |
| gender | 0 | 0,615 | 1,047 |
| age | 0 | -0,127 | -0,556 |
| student | 0 | 1,473 | 1,351 |
| | | | |
| class membership probability | | | |
| **Class mer** | 26% | 7% | 67% |
| man | 11% | 5% | 83% |
| vrouw | 49% | 7% | 44% |
| | | | |
| student | 8% | 9% | 83% |
| non-studer | 58% | 3% | 39% |

| | Class 1 | Class 2 | Class 3 | Class 4 |
|---|---|---|---|---|
| **Light** | 0,504 | 0,08 | 1,422 | 2,102 |
| **Shelter** | 0,791 | 0,188 | 0,142 | 0,746 |
| **Security** | 0,56 | 2,267 | 0,169 | 1,406 |
| **Time** | -0,17 | -0,012 | -0,036 | -0,828 |
| **Costs** | -2,551 | 0,391 | 0,017 | -2,135 |
| | | | | |
| delta | 0 | -2,152 | -1,744 | -1,933 |
| gender | 0 | 0,954 | 1,046 | 0,398 |
| student | 0 | -1,982 | -2,421 | -0,364 |
| | | | | |
| | | | | |
| class membership probability | | | | |
| **Class mer** | 70% | 8% | 12% | 10% |
| man | 50% | 15% | 25% | 11% |
| vrouw | 2% | 35% | 59% | 4% |
| | | | | |
| student | 88% | 1% | 1% | 9% |
| non-studer | 25% | 21% | 49% | 5% |

Figure O.1: Weights for three and four user classes (red means insignificant value)