D.K.A. PHAM

Mobility Design Strategies for Sustainable Development in Peripheral Urban Areas

A case study in travel choice behaviour for Rijnenburg, Utrecht.

TU Delft | Gemeente Utrecht

MSc Thesis Research

Mobility Design Strategies For Sustainable Development In Peripheral Urban Areas

A case study in travel choice behaviour for Rijnenburg, Utrecht.

Author:

D.K.A. Pham

This master thesis is submitted to Delft University of Technology
in partial fulfilment of the requirements for the degree
Master of Science in Civil Engineering – track Traffic and Transport Engineering
to be defended publicly on September 2^{nd}, 2024.

Graduation Committee:

Prof.dr.ir. B. (Bart) van Arem	Chair, TU Delft
Dr. ir. N. (Niels) van Oort	First Supervisor, TU Delft
Dr. ir. A. (Alexandra) Gavriilidou	Second Supervisor, TU Delft
Ir. Leon de Graaf	External supervisor, Gemeente Utrecht
MSc. Lara Verhagen	External supervisor, Gemeente Utrecht

TU Delft | Gemeente Utrecht

Preface

In front of you is the final report of my graduation research for fulfilling the Master's degree in Civil Engineering from Delft University of Technology. While this is, in essence, an academic requirement, it also an OV-chipkaart marking the final check-out of my academic journey at TU Delft. It is hard to believe that six years have passed since I made the decision to begin my bachelor here, drawn in by a love at first sight for the university library building and a dream of designing my own bridge. This thesis is not the construction plan for my bridge, but I hope it contributes to improving sustainable accessibility in The Netherlands.

Reflecting on the entire thesis process, not every step of working on this thesis went smoothly. There were many challenges along the way, but now I can proudly say that I overcame all of them. This thesis is not perfect and may contain mistakes, but it is the best version I can offer at this moment, just as I always strive to be the best version of myself.

I want to take this opportunity to express my deepest gratitude to Gemeente Utrecht for the opportunity to work with them, for the chance to meet new people and to experience a new professional environment. I sincerely thank Lara for all her support and efforts during my thesis. I am also appreciative of Leon for his invaluable input from Rijnenbrug and feedback for making the work more practically meaningful. To my colleagues at Gemeente Utrecht and the U Ned team, thank you for your time and contributions to this project.

Additionally, I am deeply thankful to my university supervisors for their guidance, invaluable support, constructive feedback, and continuous encouragement, which helped me refine and improve my work throughout this journey. My sincere thanks go to Bart for his enthusiasm and insightful contributions that greatly enhanced the quality of my research. I am also grateful to Niels for connecting me with this incredible opportunity and for his valuable input in shaping my work. Lastly, I appreciate Alexandra for her critical feedback and technical insights, which helped me perfect my work down to the finest details.

Last but not least, I would like to extend my deepest gratitude to my family, whose unwavering support and love have been my foundation throughout this journey. I am also thankful to my dear friends for supporting and accompanying me on this thesis journey. Lars, thank you for always being there to cheer me on and for your constant support.

In conclusion, I am grateful for the opportunity to conduct this research and contribute to the advancement of knowledge in this field. I hope that this thesis serves as a valuable addition to the existing body of research and provides practical insights for sustainable development. Although this thesis marks the end of my academic journey, it also signifies the beginning of a new chapter in my life.

Kim Pham
Delft, August 2024

Abstract

In response to the housing shortage in the Netherlands, local governments are actively exploring new development plans, particularly in major cities where land capacity is a challenge due to high population density. This has led to considerations of urban expansion, even in less optimal locations near freeways, offering convenient car accessibility but contributing to increasing congestion and contradicting principles of sustainable development. To mitigate the negative impact, mobility measures regarding public transit and car planning in the development stage of the area are crucial for the development of such areas, which are referred to as peripheral areas in this research. In this research, the main objective is to define the mobility design strategy for planning public transit and car parking in peripheral areas. Accordingly, the main research question is defined as follows:

"Which mobility design strategies, regarding public transit and car parking planning, should be implemented to reduce car usage and enhance the attractiveness of public transit in a peripheral urban area?"

The results of the research are obtained by developing a Discrete Choice Model for capturing the commuting mode choice behaviour for different scenarios. This process involved several steps: defining the influencing factors, developing the hypothesised strategies, collecting mode choice preferences and estimating the model parameters based on the collected data.

First of all, all influencing factors of the commuting mode choice are identified from the literature. These are the relevant mode alternatives and the related trip characteristics, the personal characteristics of the decision-maker, and the factors of build environments and work conditions. For the context of this study, only urban transportation modes, bus, light rail, car and bike, are relevant. The corresponding trip characteristics are the in-vehicle time (also referred to as driving for cars and cycling time for bikes), travel distance, access time and distance, waiting time, transfer time, egress time, travel fare, vehicle type and the travel purpose. Regarding personal characteristics, age, educational level, employment status, income, habit, mobility, disability and availability are found to be necessary to include in this research. Additionally, the urbanisation level, parking availability and commuting reimbursement are relevant to be included in the context.

In line with the scope of this research, Rijnenburg has been chosen as a case study area for this research to provide more practical insights. Rijnenburg is an outskirt area of Utrecht's municipality, located at the interchange of two important motorways, A2 and A12. Due to this location characteristic, Rijnenburg has excellent car accessibility. In the coming years, Rijnenburg will be developed into a new urban area of Utrecht with around 20,000 to 25,000 housing units and 10,000 to 15,000 new workplaces. However, the car-promoting characteristics of the area present a risk of increased car traffic, which could lead to higher congestion and undermine sustainable mobility development in the region. Therefore, sustainable mobility planning strategies for the development of this area are required to enhance the attractiveness of public transit use and prevent the increase in car use and car ownership in Rijnenburg.

To address these challenges, strategies regarding the planning of local and global public transportation networks, local car parking planning and using reimbursement to affect the use of bikes were developed and scaled into three different levels, resulting in three design scenarios for Rijnenburg: Conventional, Sustainable, and Ambitious. These scenarios are developed based on the existing development directions of the municipality of Utrecht, reflecting different levels of expected impact of the implementation of these strategies. The design strategies suggest the consideration of stop density, line frequency, route planning, and vehicle type for public transit planning. For the car use interventions, the strategic locating of parking facilities and setting of parking costs could be effective for the reduction of car use and ownership. In addition, the commuting reimbursement form can be adapted from the conventional forms to stimulate the use of sustainable transportation modes by making the use of bikes, e-bikes and public transit economically available.

After that, a Stated Preference was conducted to gather the mode preference data for constructing a Discrete Choice Model for commuting in Rijnenburg. In this experiment, the commuting trip between Rijnenburg and Utrecht Science Park is applied to capture the mid-travel distance travel aspect of the research scope. In total, 200 valid responses were collected in around two months. The respondents mainly consist of residents of the area surrounding Rijnenburg, Utrecht employees, and university students in Randstad. Compared to the reference population, Province Utrecht, the data sample is more presented by young and highly educated individuals, indicating possible biases in the choice model. Therefore, this significant difference has to be taken into account in the interpretation of the results.

On top of the collected data, both a Multinomial Logit (MNL) model and a Panel Mixed Logit (Panel ML) model were constructed. The Multinomial Logit model is simple and widely used in travel behaviour research, while the Panel Mixed Logit Model is more complex but does capture the randomness in repeated answers of the panel data. Using a forward stepwise estimation strategy, it was found that the MNL model has lower explanatory power than the Panel Mixed Logit Model but higher accuracy. Additionally, the MNL model included more attributes, providing deeper insights into the interpretation of influencing factors. As a result, the MNL model was selected for application in analysing commuting mode choice behaviour in Rijnenburg.

The estimated values of the alternative specific constants indicated that the respondents have a high preference for bikes and they dislike using cars. In general, the alternative properties have a higher impact than the personal characteristics, indicating the importance of the proper organisation of transportation infrastructure. Besides, the transfer attribute of public transit has the largest impact on the model, while mobility disability has the lowest impact. The design variable in-vehicle time for PT has the lowest impact on the utility of public transit, and parking cost has the lowest impact on the utility of the car.

To evaluate the effectiveness of the developed design scenarios in to the commuting mode choice behaviour, the constructed MNL model is applied to calculate the modal splits of Rijnenburg. The modal split of the Conventional scenario shows that this design scenario should be avoided for the development of Rijnenburg; car use will be dominant, and car ownership will be increased in the area. The Sustainable scenario does lead to a decrease in the share of cars and an increase in PT. However, the share of cars is still higher than PT, and during the off-peak hours, it becomes dominant again. To create a more clear aversion to car use, tougher interventions are needed. Therefore, the Ambitious scenario is the most effective one. For this scenario, commuters mainly opt for public transit or the bike. The car has become the least favourite due to the restricted parking location and high parking costs. However, this scenario includes two unfeasible applicable levels for the availability of directness transit lines and e-bikes for all commuters. By performing a sensitivity analysis for the mode alternatives, a more appropriate approach is retrieved. The found approach is a combination of the Sustainable scenario and higher thresholds for car access time and parking cost.

The research results exhibit several discussions and limitations. The discussion highlights concerns about the survey data's suitability, noting that it was slightly below the desired sample size and lacking in diversity among respondents. The choice of the MNL model over the Panel ML model, despite the latter's higher explanatory power, is also debated due to the MNL model's better accuracy and interpretability. Furthermore, factors like the type of public transit vehicle, waiting time, and transfer distance were deemed insignificant in the model, possibly due to the survey's complexity or the small sample size, which may have led respondents to overlook these aspects. Additionally, a limitation in the applicability of the study is noted, as the focus on mid-distance commuting might be too narrow, potentially reducing the generalizability of the findings, particularly when applied to other travel purposes or destinations. Besides, the design of the Stated Preference (SP) experiment may not have clearly communicated the unique characteristics of peripheral areas to respondents, leading to similar decision-making patterns as in non-peripheral areas. The survey's complexity and the number of choice tasks may have overwhelmed respondents, resulting in inconsistent answers. Finally, the study's scope was limited to commuting behaviour, overlooking other critical factors like personal perceptions, costs, and technical feasibility, which are crucial for developing practical and sustainable mobility strategies. These insights and limitations suggest that further research is needed to refine the models and broaden the scope to include a wider range of influencing factors.

In conclusion, the answer to the main question, as well as the main recommendation for practice, can be retrieved. To enhance the attractiveness and prevent the increase of cars in peripheral areas, a combination of several mobility strategies is required to be implemented in the development of the area. For **public transit planning**, minimising transfers is crucial as they significantly influence travellers' mode choice behaviour. Therefore, a wheel network form with as many direct lines as possible is preferred. Additionally, the access distance to public transit stops should be optimised, with an ideal distance of around 1000 meters between stops. For **car planning**, increasing the access distance to cars is the most effective strategy to discourage car use. Therefore, parking facilities should be located outside the residential area, with a minimum walking distance of 16.75 minutes. Additionally, a high monthly parking cost should create an economic resistance for the car ownership in the area.

The retrieved design strategies mentioned above are also the main recommendation for the mobility design in Rijnenburg. Furthermore, the research also recommends carrying out deeper studies into the eliminated factors, other travel purposes and distances to have a more complete insight into the mode choice behaviour of the area. Besides, an optimisation study that includes cost and technical constraints is also recommended to ensure the practical applicability of the strategies. For future research, when designing the Stated Preference (SP) experiment, it is recommended to have a good balance between the levels of detail and number of questions in the survey to maintain the significance of the results. Also, the pilot study with a high diversity of respondents is highly recommended to enhance the quality of the survey.

To conclude, the mobility strategies developed in this research are expected to have a significant contribution to the sustainable mobility development of Rijnenburg. Preventing the increasing car traffic and enhancing the attractiveness of sustainable transportation modes for the residents of Rijnenburg.

Contents

List of Figures

List of Tables

List of Abbreviations

AIC .... Akaike Information Criterion

ASC .... Alternative Specific Constant

BIC .... Bayesion Information Criterion

DCM .... Discrete Choice Modelling

IIA .... Independence of Irrelevant Alternatives

LCCM .... Latent Class Choice Model

LL .... Log Likelihood

MNL .... Multinomial Logit

NL .... Nested Logit

Panel ML .... Panel Mixed Logit

PT .... Public transit

RUM .... Random Utility Maximisation

SP .... Stated Preference

STOMP- .... Stappen-Trappen-Openbaar Vervoer-Mobility as a Service-Privéauto

TOD .... Transit-Oriented Development

USP .... Utrecht Science Park

1

Introduction

1.1. Research Context and Problem

The housing shortage has become a hot topic in The Netherlands in recent years. Across the country, local governments are actively exploring opportunities for new development plans to address the need for additional residential areas (Rijksoverheid, 2023). This reflects the development and growth of the country. However, there is a challenge in the capacity constraint of land, particularly in big cities where buildings and the population density are already high (Wynia, 2021). Consequently, urban expansion is necessary, even if it extends to less optimal locations.

Due to historical spatial developments, Dutch cities are predominantly delimited by highways and ring roads (Nabielek, 2012). Consequently, urban expansion typically occurs beyond these infrastructure boundaries, on the city's outskirts, namely the peripheral areas. These newly developed areas are often situated close to highways but at a considerable distance from the main city centre and essential city-scale facilities and services. The primary advantage of such locations lies in their easy car accessibility. These characteristics heavily promote car usage while limiting the feasibility of more sustainable transportation modes, making the locations develop towards car-oriented and contributing to the increase of car use and car ownership.

Currently, the most widely used passenger transportation mode in Europe is the car, for both daily commuting and recreational travel (Shabani, 2023). According to Ritchie and Roser (2021), passenger cars account for a large part of CO2 emissions in the transport sector. Developing new (sub)urban areas in the conventional way will most likely increase the number of cars further, taking examples of Leidsche Rijn (Utrecht), Nootdorp, and Leidschenveen (The Hague), as well as Nesselande (Rotterdam). Therefore, strategic spatial and mobility planning are necessary to ensure the sustainable development for new urban areas (Snellen et al., 2021).

The proximity to highways often encourages car usage over cycling and public transportation, even for short and mid-distance trips. Consequently, this contributes to a higher traffic density on the surrounding highways than in other areas. Strategies to reduce short-distance car trips could involve strategic planning of public facilities and focusing on their proximity to residential areas to attract the use of active transportation modes (Olsen et al., 2024). Also, making use of the prevalent cycling culture in the Netherlands and encouraging cycling for short journeys could be beneficial. For mid-distance trips ranging from 10 to 25 kilometres, additional approaches are required (Delice et al., 2019). These should place a stronger emphasis on promoting the use of public transportation and implementing measures which discourage the use of cars.

In The Netherlands, the highest travel kilometres are attributed to car usage, especially for commuting purposes, where cars also have the highest share (Centraal Bureau Voor de Statistiek, 2023b). According to CBS data, the average commuting distance stands at 19 kilometres (Centraal Bureau Voor de Statistiek, 2023a, Plan bureau voor leefomgeving, 2020). This preference for cars over public transit for commuting trips is often driven by the convenience and flexibility of use (Alonso-González et al., 2020). This dominance highlights the need to address commuting behaviour for having suitable mobility development strategies to prevent the increase in car use and encourage choices for more sustainable alternatives. It is worth noting that altering habitual travel behaviours for consistent commuting trips may be more feasible than influencing irregular travel purposes (Vos, Waygood, et al., 2022).

To conclude, there is a need to study deeper in mode choice behaviour within the context of urban development in peripheral areas. As the demand for sustainable transportation solutions grows, understanding the factors influencing mode choice becomes crucial. By focusing on mid-distance travel and commuting as the primary travel purpose, this study aims to provide insight into factors that could reduce car use for mid-distance tips, particularly during peak hours. Based on the insight gained, the result of this study will help urban planners and policymakers develop appropriate design strategies for public transportation in new development areas.

1.2. Research Objective and Scope

As described in the research problem, there is a need to study the commuting mode preferences and the potential of different mobility design strategies in the context of peripheral locations. Therefore, the objective of this research is to identify and develop effective urban mobility design strategies, with a specific focus on public transit and car parking planning to increase the attractiveness and usage of public transportation in peripheral urban areas. These strategies aim to contribute to defining the boundary conditions for the mobility design of peripheral areas, helping to prevent the development of car-oriented mobility in new urban areas.

The focus area in this research is to investigate and analyse the preference and behaviour of individuals in mode choice within peripheral urban environments, with a primary focus on daily commuting. The scope includes an examination of explicit trade-off considerations and the factors that drive mode choice in car-oriented settings. Additionally, the research will delve into the influence of transit network characteristics, route planning, stop density, line frequency, and vehicle type on individuals' preferences. In terms of location, the research area is scoped for new development areas with excellent freeway connectivity; Rijnenburg has been therefore chosen as a case study area for this research.

1.3. Research Questions and Approach

Based on the identified research objective and scope, the primary research question of this study can be formulated as follows:

Which mobility design strategies, regarding public transit and car parking planning, should be implemented to reduce car usage and enhance the attractiveness of public transit in a peripheral urban area?

The main research question is going to be supported by the following sub-questions:

1. *Which characteristics of travellers, public transit and car use could affect the mode choice behaviour of travellers?*
2. *Which mobility design strategies could be applied in the development of peripheral urban areas?*
3. *To what extent do these characteristics affect the mode choice behaviour of commuters in a peripheral urban area?*
4. *To what extent are these mobility design strategies effective in reducing car use among commuters in peripheral urban areas?*

The first step of the research is performing a literature study to define the mode alternatives and the possible influencing factors of the traveller's mode choice behaviour from the existing research on travel behaviour. These are the hypothesised attributes and include all the possible factors regarding personal characteristics and properties of the choice alternatives.

The literature study will be continued for sub-question 2 on the case study area, Rijnenburg in Utrecht. Based on the results of the conducted mobility research for Rijnenburg, the characteristics and future development scenarios with different mobility measures for this area are analysed. Additionally, several expert interviews will be conducted with the local government to determine the preconditions and desired attractiveness levels of public transit in this area. These components form the levels of the attributes of the travel behaviour in the case study area.

The outcome of the conducted literature studies and the expert interviews are used to develop a survey for the Stated Preference (SP) experiment. This experiment is a data collection paradigm aiming to identify the preferences of modes using predicated attributes and attribute levels. The collected survey data from the choice experiment will serve as input for calibrating the discrete choice model (DCM) model parameters, as addressed in sub-question 3. Also, these collected data are used for the model validation.

After that, in sub-question 4, the constructed discrete choice model will be applied to the case study area. This aims to predict future commuters' travel behaviour and assess the effectiveness of mobility measures found in sub-question 2. On top of this, the mobility strategies for developing Rijnenburg can be derived.

1.4. Research Relevance

The research is expected to have significant contributions to both research literature and practical implementation. Generally, the results can be helpful for urban planners, policymakers, and researchers who focus on passengers' travel behaviour, aiming to decrease the proportion of car usage.

For society, this research aims to contribute to the improvement of urban accessibility and sustainable urban mobility development. The research's outcome provides policymakers with strategic insights for organising infrastructure in new development areas. By analysing travel choice behaviour, the study offers practical recommendations for designing transportation systems. This ensures that infrastructure developments are well-planned and capable of supporting the anticipated demand, leading to a more efficient and effective transportation network.

Additionally, this research serves as a critical assessment tool for proposed policy measures related to transportation and urban development. By understanding the potential impacts of various policies on travel behaviour, policymakers can evaluate the effectiveness of their strategies. This assessment helps in refining and optimising policies to better achieve desired outcomes, such as reducing traffic congestion, promoting sustainable transportation modes, and enhancing overall urban mobility.

Furthermore, by incorporating the travel preferences of potential users, the research ensures that the human aspects of societal satisfaction are prioritised in the planning process. By understanding future residents' specific needs and preferences, mobility solutions can be developed to meet these needs, making them more likely to be adopted. This participatory approach not only enhances the likelihood of successful implementation but also promotes community involvement. Ensuring that the transportation infrastructure aligns with user preferences leads to higher satisfaction levels and a better quality of life for residents.

From a scientific perspective, this study contributes to the academic understanding of travel choice behaviour, particularly in the context of peripheral development areas. It expands the existing body of knowledge by exploring the specific factors that influence mode choice in a peripheral developing urban setting, providing empirical data and insights that can be utilised in future studies. The research also included a Stated Preference experiment, which provides robust empirical data on the factors influencing commuting mode choice. This data enriches existing theoretical models, enabling more accurate predictions and analyses of travel behaviour.

Overall, this research bridges the gap between theoretical inquiry and practical application, providing insights valuable to both academia and practitioners in urban development and mobility planning.

1.5. Report Outline

As the starting point of the research, a literature review is conducted in Chapter 2 to provide an in-depth analysis of the problem and the methodology. Chapter 3 describes the hypothesised factors that could influence the mode choice behaviour of commuters. In Chapter 4, the case study area of Rijnenburg is analysed in depth to develop potential mobility design strategies. The findings from Chapters 3 and 4 inform the design of the Stated Preference experiment and the construction of a discrete choice model, detailed in Chapter 5. The results of these analyses are presented in Chapter 6. Following this, the choice model is applied to commuters in Rijnenburg to assess the effectiveness of the proposed mobility design strategies, with the outcomes discussed in Chapter 7. Chapter 8 offers a discussion of the results and reflections on the limitations of the study. Finally, Chapter 9 concludes the report, providing answers to the research questions and recommendations for the development of Rijnenburg and future research.

2
Literature Review

A literature study is carried out in this chapter to delve into the existing literature on travel choice behaviour in peripheral development areas. First of all, the research context will be studied in depth by performing a state-of-art about car-oriented development in peripheral areas. Herein, the definition, the characteristics and the threat to sustainable urban mobility development will be explored. After that, a review of the methods that could be used to tackle this problem will be conducted in section 2.3. This chapter finishes with a description of the research gap that briefly indicates what the main problem is and which method will be used to tackle the problem in this research.

2.1. Methodology for Literature Review

The literature review for this research serves two primary goals: to define the problem that the study seeks to address and to determine the appropriate research methodology. The review will begin by exploring existing studies related to mode choice behaviour, particularly within the context of peripheral urban areas, to clearly articulate the problem statement of this research. This involves understanding the main mobility challenges and the need to study the choice behaviour of travellers in these areas. Additionally, the reviewing literature provides a deeper understanding in Discrete Choice Modelling, serving as the foundation to develop the research methodology.

To retrieve relevant literature for the review, Scopus, Google Scholar and Google search engine are used. Keywords such as “Mode Choice Behaviour In Peripheral Areas”, ”Car-oriented Urban Areas”, ”Peripheral Mobility Development”, ”Mobility Development Strategies”, “Discrete choice modelling” have been used in the search queries to locate relevant literature. In addition to the articles found through the initial Scopus database search, further literature was identified using a ’snowballing’ technique. This involved reviewing references and related papers cited in the initially selected articles.

Specific inclusion and exclusion criteria were applied to ensure the relevance and quality of the literature included in this review. The inclusion criteria focused on selecting articles written in English that directly addressed mode choice behaviour for commuting purposes, particularly in the context of peripheral urban or suburban areas, or those that contributed to understanding Discrete Choice Modelling. Studies that primarily addressed other travel purposes, such as tourism, leisure, or long-distance travel, were excluded, as they did not align with the core focus on daily commuting.Additionally, papers were excluded if they focused solely on urban areas of main city centres or isolated regions or if they discussed transportation modes or strategies that were not applicable to the study area. These criteria helped to refine the literature to those studies most relevant to the research objectives of understanding mode choice behaviour and mobility strategy development in peripheral urban areas. Additionally, papers focusing on destination choice modelling from a supply-side perspective were excluded since this study assumes a fixed destination. Through this process, a well-defined and focused set of 35 articles was identified for further analysis.

The methodology for determining the relevance of an article involves a multi-step process. First, the title of each paper is reviewed to assess whether the overall theme aligns with the research objectives. If the title suggests relevance, the abstract is then read to make a more informed decision on whether to include or exclude the paper from the review. This approach ensures that only the most relevant studies are considered. As the search progresses, additional keywords may be introduced to refine the search results and narrow the focus to specific domains, thereby increasing the relevance of the retrieved papers. This iterative process allows for the continuous refinement of search queries to ensure comprehensive coverage of the relevant literature.

2.2. Car-oriented in Peripheral Development Areas

Peripheral development areas are regions situated on the outskirts of a big city. These areas are characterised by low-density development, a significant distance from the urban core, and often limited infrastructure (Woltjer, 2014). These regions usually encompass less densely populated zones, underdeveloped land, and suburban sprawl. Peripheral areas often have lower land costs, making them attractive for urban expansion (Ola Adetokunboh MRICS, 2023).

Typically, peripheral areas are located at a distance from the major city centre and current public transportation infrastructures, the area is mostly attached to the city by existing motorways. This makes the area more accessible by car than other modes of transportation and more vulnerable to becoming a car-oriented location (Sarwar, 2021). Because of this, these development of peripheral areas often face challenges in integrating into the broader urban fabric, particularly concerning sustainable mobility and accessibility (Moyano et al., 2023, European Commission, 2024).

Various studies have shown that there is a clear relationship between the land use planning (location and structure) and the travel behaviour of people (Chatman, 2009, van Wee and Handy, 2016). In the article 'How to fit the car in urban areas', Terlien (2018) has recited four ways for a relationship between the car and the city. **Car-centric**, people need a car, is prevalent in both rural and urban settings, necessitate car dependence for daily activities due to limited alternatives. **Car-oriented**, people want a car, is when there are alternatives but cars are still more comfortable for some jobs and urban services, zoning rules and (single) zoning areas prevail. When cars become unnecessary due to superior transportation alternatives, the urban becomes **multi-modal**, when people do not need a car. Finally, when cars are unwanted and unnecessary due to the high attractiveness of the other transportation alternatives, the urban area will become **walkable**.

Since the Second World War, transport and land use planning have been planned with a focus on private cars, particularly in Western countries (Rye and Hrelja, 2020). That has led to an oversupply of car-oriented urban structures and, thus, a constant increase in the number of cars (Moyano et al., 2023). Current levels of car ownership and use contribute significantly to road congestion, resulting in a notable decrease in accessibility to essential economic zones and negative impacts on both humans and the environment (Haustein and Kroesen, 2022 and van Exel et al., 2010). Because of the negative impacts, a reduction in the number of cars owned and used is necessary and urgent in the current society.

However, as Terlien (2018) mentioned in his article, there are not enough options for multi-modal and much too few options for walkable urban areas in the most sub-urban areas. Therefore, (urban) planners should pay more attention to the potential to turn the car-oriented into a multi-modal or even walkable city. This concern has been underscored by notable examples of urban development in peripheral areas in The Netherlands such as Leidsche Rijn (Utrecht), Nootdorp, Leidschenveen (The Hague), and Nesselande (Rotterdam) (Jonkeren et al., 2019, Woninck, 2015). The dominance of car use in these examples has shown that without careful planning strategies, peripheral urban areas can easily prioritise car transportation for residents and evolve into car-dependent communities.

To solve this problem, multiple mobility development trends are ongoing and focus on promoting sustainable urban growth in peripheral areas. The mobility transition emphasises shifting from car use to sustainable transport modes like walking, cycling, and public transit (PT) (Goudappel, n.d). It encourages integrated urban planning and multi-modal transportation networks. Implement the STOMP principle in urban mobility design, which prioritises walking (Stappen) and cycling (Trappen), followed by public transport (Openbaar vervoer), Mobility as a Service, and lastly, private car use (Privéauto) (CROW-KpVV, 2021). This approach aims to create more sustainable and liveable urban environments by reducing car dependency. Transit-oriented development (TOD) focuses on creating high-density, mixed-use communities centred around public transit hubs (Ibraevaa et al., 2020). This model aims to reduce car dependency by providing accessible public transport options and promoting sustainable living environments. Car-free urban areas are initiatives that restrict car use in certain zones, creating pedestrian-friendly and bike-friendly environments (Loo, 2018). These initiatives enhance urban livability by reducing pollution and congestion, making cities more sustainable.

The challenges of developing a peripheral area sustainably are multifaceted, involving limited public transport, car dependency, inadequate cycling and walking infrastructure, and potential social isolation. Car-oriented development worsens these problems, leading to environmental degradation, urban sprawl, traffic congestion, and reduced quality of life. Addressing these issues requires a comprehensive approach to urban planning that prioritises sustainable development principles. These identified problems justify the need for research focused on modelling travel choice behaviour in peripheral development areas. Understanding the factors that influence travel choices can inform the design of effective mobility strategies that promote the use of public transport and reduce car dependency, ultimately contributing to sustainable urban growth.

2.3. Methodology for Travel Choice Behaviour Research

Transport modelling can help in understanding the effects of the different mobility measures (Goudappel, n.d). A critical component of transport models is the behavioural foundation: how individual travellers are likely to behave. Travel behaviour research can help understand mobility's effects, which refers to the decision-making processes individuals use to select various transportation options. It has been widely applied to assess the effects or implications of intended policy measures in designing.

In the travel choice research, there are two data collection paradigms: revealed preferences and stated preferences (Kroesen, 2024). Revealed preference (RP) methods rely on observed choices individuals make in real-world situations, and it is not subject to hypothetical bias since they are based on observed actions rather than stated intentions. It reflects actual behaviour in real-world contexts, providing insights into how people make travel choices based on their experiences. This data is typically collected through surveys, travel diaries, or direct observation of travel behaviour (Train, 2002).

Stated Preference (SP) methods involve presenting respondents with hypothetical scenarios and asking them to choose their preferred option. These scenarios are carefully designed to capture various factors influencing travel choices, such as travel time, cost, and convenience. SP allows researchers to control the attributes and levels of the alternatives, enabling the study of new or hypothetical transport options not yet available in the real world (Train, 2002). Therefore, it is possible to test respondents' reactions to a wide range of hypothetical policy measures or infrastructure changes (Kroesen, 2024). Using a Stated Preference experiment allows for collecting hypothetical scenario data for developing a discrete choice model to predict travel behaviour in peripheral areas (Molin, 2024).

The Discrete Choice Modelling (DCM) has been widely applied as a modelling method in most travel behaviour research to describe decision-makers' choices among alternatives (Ben-Akiva and Bierlaire, 2003). This method is used to understand the reasoning behind human choices, predict behaviour, and potentially influence decision-making. Four key elements of Discrete choice theory comprise:

- Decision Maker: The decision maker can be an individual or a group considered a single entity, ignoring potential internal negotiations. Characteristics such as gender and age help capture heterogeneity in choices (Ben-Akiva and Bierlaire, 2003).
- Alternatives: Alternatives come from a finite set of exhaustive and exclusive options known and feasible for the decision-maker. In continuous choice problems, discretisation is necessary (Ben-Akiva and Bierlaire, 2003). To analyse decision-making, both chosen and unchosen options have

to be understood. Assumptions about available alternatives, known as the choice set, are essential (Ben-Akiva and Bierlaire, 2003).

- Attributes: Each alternative in the choice set is defined by a set of attributes, which can be continuous or categorical (Ben-Akiva and Bierlaire, 2003). Some attributes may be common across all alternatives, while others are specific to certain options. While some attribute values are known with certainty, others are uncertain, and their variance can be included as an additional attribute (Ben-Akiva and Bierlaire, 2003).
- Decision Rule: This is the process the decision maker uses to evaluate alternatives and make a choice. Most travel behaviour models are based on utility theory, which assumes that preferences are captured by a utility value (Ben-Akiva and Bierlaire, 2003). The decision maker selects the alternative with the highest utility in the choice set. By far, the most widely used discrete choice model is the Logit family (Train, 2002).

According to Ben-Akiva and Bierlaire (2003), the four commonly used Logit models for Discrete Choice Models (DCM) include the Multinomial Logit Model (MNL), Nested Logit Model (NL), Panel Mixed Logit Model (Panel ML), and Latent Class Choice Model (LCCM):

- The MNL is popular due to its simplicity and ease of calculating choice probabilities, which makes it computationally efficient. However, a key limitation of this model is its assumption of independence of irrelevant alternatives (IIA), meaning that the relative odds of choosing between any two options remain unaffected by the presence of additional alternatives. This assumption might not hold true in real-world situations, limiting the model's applicability in some cases.
- The NL model addresses the IIA limitation by allowing for correlations of similar alternatives. This flexibility makes it more suitable for situations where choices are naturally grouped, such as different modes of transportation that are similar in nature (e.g., buses and trains). By relaxing the IIA assumption, the NL model provides a more accurate representation of decision-making when grouped choices are involved (Train, 2002).
- This model is known for its high flexibility and ability to approximate any random utility model. It overcomes the limitations of the MNL by accommodating random variations in individual preferences, allowing for unrestricted substitution patterns among alternatives, and accounting for correlations in unobserved factors over time. The Panel ML model is particularly useful when dealing with repeated choices by the same individuals, as it captures individual-specific preferences and provides a more realistic representation of heterogeneous decision-making behaviour (Train, 2002).
- The LCCM segments the population into different classes or groups based on distinct choice behaviors. This model is valuable for identifying heterogeneity within the population, as it allows researchers to recognize and analyze varying preferences across different market segments. By tailoring policies or products to these specific groups, the LCCM can provide insights into more targeted and effective interventions. (Ben-Akiva and Bierlaire, 2003).

In conclusion, the research on mode choice behaviour in peripheral areas, the Stated Preference experiment should be applied as the data collection paradigm. This approach allows for creating hypothetical scenarios to capture various factors influencing travel choices, which is essential for evaluating hypothesised mobility measures. Regarding the Discrete Choice Modelling, the Multinomial Logit Model (MNL) is an excellent starting point due to its simplicity and computational efficiency, making it ideal for establishing a baseline model. This baseline can then be extended into the Panel Mixed Logit Model (Panel ML), which builds on the MNL framework by capturing individual-specific preferences and repeated choices over time, providing a more realistic understanding of travel behaviour dynamics. The Nested Logit Model (NL) is unnecessary for this study because it focuses on grouped alternatives, which adds complexity without aligning with the study's goal of analysing individual-specific preferences. The Latent Class Choice Model (LCCM) is also less suitable as it emphasises population segmentation, whereas this research aims to understand overall mode choice behaviour across the entire population, making the Panel Mixed Logit Model (Panel ML) a more appropriate choice.

2.4. Conclusion on Literature Review

This literature review has delved into the relationship between peripheral urban areas and car-oriented mobility development. It has highlighted how peripheral areas, often characterised by low-density development and significant distance from urban cores, are prone to becoming car-dependent. This dependency poses challenges to sustainable urban mobility and broader environmental goals.

A key insight from this review is the critical importance of understanding mode choice behaviour in peripheral areas within the Dutch context. The experiences of similar such as Leidsche Rijn (Utrecht), Nootdorp, Leidschenveen (The Hague), and Nesselande (Rotterdam) illustrate the risks of car dependency in these regions. Without mobility planning and intervention, these areas can evolve into car-dominated communities, exacerbating traffic congestion, environmental degradation, and reduced livability. As the Netherlands continues to expand its urban peripheries, studying how travellers make transportation decisions becomes essential for developing strategies that promote sustainable mobility.

Discrete choice modelling is a widely used method in travel behaviour research, aiming to analyse the factors that influence the underlying decision-making process of travellers. By employing the Stated Preference experiment, various hypothetical scenarios with different mobility measures for urban development can be evaluated. The Panel Mixed Logit Model (Panel ML), which is an extension on the Multinomial Logit Model (MNL), could be especially effective in capturing individual-specific preferences and accounting for repeated choices over time, offering a comprehensive understanding of travel behaviour dynamics.

Using the literature findings and the research objective from section1.2, a conceptual framework for the construction of the discrete choice model to capture the mode choice behaviour of the commuters in peripheral urban areas is developed and visualised in figure 2.1.

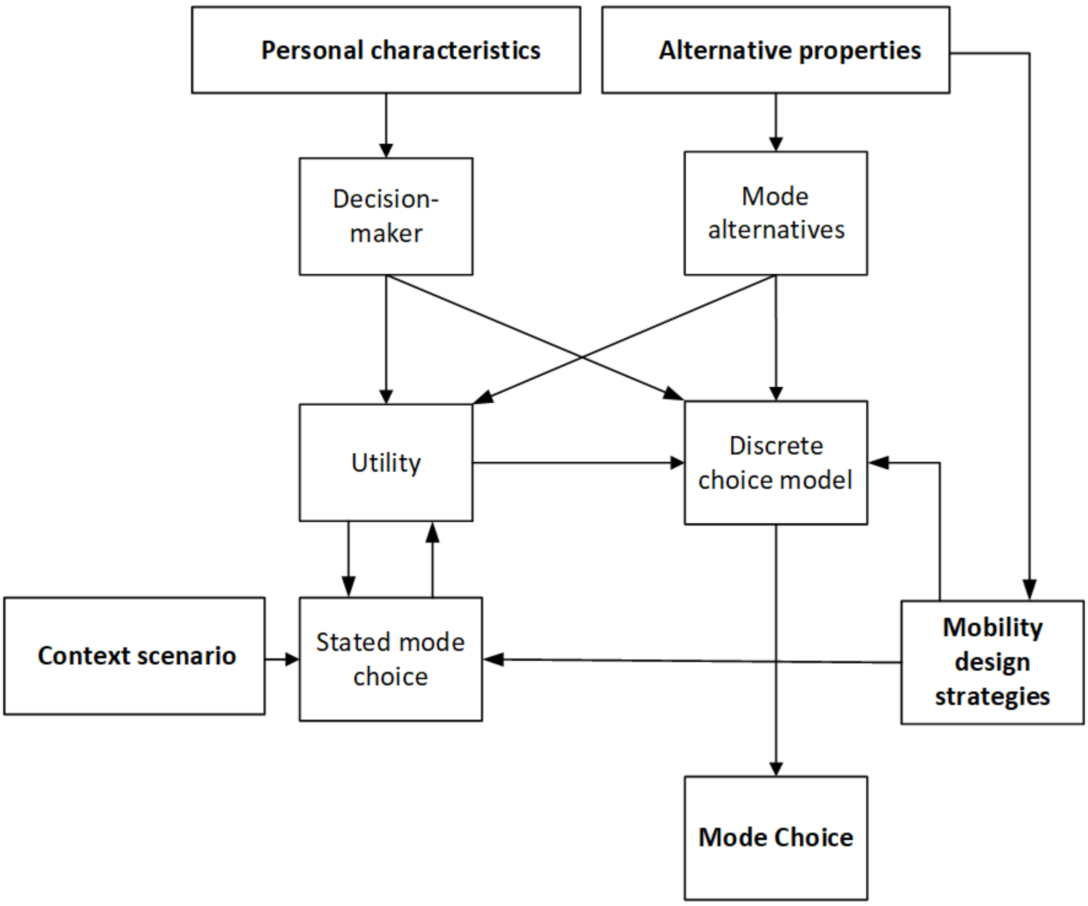


Figure 2.1: Theoretical framework of commuting mode choice behaviour modelling

This conceptual framework outlines how commuting mode choice behaviour in peripheral urban areas is shaped by the interaction between personal characteristics and the alternative properties of available transportation modes. These factors feed into the utility calculation, which the decision-maker uses to evaluate and choose the most suitable mode through a discrete choice model. Contextual scenarios, such as travel conditions and environmental factors, further influence the stated mode choice. The framework also integrates mobility design strategies, such as improved cycling infrastructure or public transport enhancements, which aim to shift mode choices towards more sustainable options. Ultimately, this model helps predict actual mode choices and assesses the impact of different strategies on promoting sustainable commuting practices.

3 Influencing Factors of Commuting

After a general review of the current literature in chapter 2, the influencing factors of commuting travel behaviour are studied in depth, to define the attributes for the mode choice modelling later on. The attributes consist of the characteristics of the travellers and the properties of the mode alternatives. As a result, a detailed overview of potential determinants for commuting mode choice behaviour is used as the basis for the development of the strategies and the Stated Preference experiment in the later chapters.

3.1. Mode Alternatives

The first main component of the study is the mode alternatives, which are all available alternatives that the decision-maker is expected to make. Depending on availability and the geographic characteristics of the travel location, the mode of transportation for commuting travel can vary. According to the statistics of Centraal Bureau Voor de Statistiek (2023b), the most commonly used mode of commuting in The Netherlands is by car (driver) with 66.2%, followed by train at 9.7%, and then by bicycle at third place with 7.5%. The urban public transportation (bus, tram and metro) is in fifth place with 2.6%, which is lower than the share of the car-passenger, 3.5%. Furthermore, a small portion of workers walk to their work, however this is strongly dependent on the travel distance.

Additionally, there is a share of 9.9% for other transportation modes, this includes the new types of transportation, such as e-bike, e-step, and e-scooter. Most likely, this portion will further increase in the upcoming years. The Netherlands Institute for Transport Policy Analysis (KiM) expects the share of e-bike in total bike use to increase from 22% in 2019 to 43% in 2027 (Kennisinstituut voor Mobiliteitsbeleid, 2022).

To capture the alternatives for mid-distance commuting, the three conventional transportation modes have been included in this study. These are the public transportation, (personal) car and bike. Depending on the context of the case study, specific types of public transportation (train, tram, bus or metro) can be selected. Due to the similarity and the development trend, e-bike is also be included in this study as an attribute of the bike option. While the car-passenger has a certain share in total modal-share, this mode choice depends on the availability situation and is therefore irrelevant for inclusion in this study. Similarly, walking has been excluded due to its impracticality for travelling this distance.

3.2. Determinants for Mode Choice Modelling

The influencing factors or the hypothetical determinants of travel choice behaviour have been extensively studied. Many determinants have been found and proven to have certain effects on the choice behaviour of the traveller. This section identifies the expected determinants of travel choice behaviour for commuting and travelling. These are also the hypothesised attributes of the discrete choice model that has been constructed later in this study. Ton et al. (2019) divided the determinants of commuting travel choice behaviour into six main categories related to different characteristics of a decision maker in different scales. Individual, household characteristics, trip characteristics, urban environment, work conditions and other external impacts.

3.2.1. Individual characteristics

The individual characteristics concern all determinants related to the decision-maker, namely socio-demographic factors and mobility factors. Understanding the influence of individual characteristics in their travel preferences and decisions is crucial for predicting travel choice behaviour in the new development area of Rijnenburg.

Socio-demographics

The socio-demographics include age, gender, income, educational level, and employment status. Age has been proven to be a crucial determinant in travel behaviour, influencing physical capability and preferences. Older commuters prefer cars due to comfort and safety, whereas younger commuters lean towards cycling and walking, driven by health and generational perceptions (Simons et al., 2014, Ahmed et al., 2020, Muñoz et al., 2016 and Hjorthol et al., 2010). Therefore, age is included in this study.

Although gender has been widely examined in mode choice studies, its impact on commuting is less significant in the Netherlands, a high-cycling country with well-developed infrastructure for all genders (Ton et al., 2019, Urmi et al., 2022). While some studies show men cycle more, women cycle more in countries like the Netherlands and Denmark (Heinen et al., 2010). Gender's limited influence on commuting leads to its exclusion from this study.

Educational level correlates with income and environmental awareness, influencing preferences for sustainable transport modes like public transport and cycling (Scheepers et al., 2013, J. Li et al., 2018). Employment status significantly affects daily travel patterns; full-time employees have different needs compared to part-time or unemployed individuals since they used to make more commuting trips. Income impacts the affordability and accessibility of transport modes; higher-income individuals prefer cars for convenience, while lower-income groups may dependent on public transport (Simons et al., 2014).

Personal mobility

Besides the socio-demographic factors, the personal mobility factor is also expected to have a significant influence on the mode choice of commuters. This factor captures all the mobility disability, availability and travel habits of the decision-maker.

Travel disabilities can limit the choice of transport modes available to individuals, making accessibility a crucial consideration in travel behaviour models. Including travel disabilities ensures that mobility strategies are inclusive and cater to the needs of all residents (Ton et al., 2019). Possessing a driving license is a precondition to owning a car: without a driving license, one cannot legally drive (Zijlstra et al., 2022). Car ownership typically leads to higher car usage due to convenience and flexibility, while having access to a bicycle promotes cycling (Simons et al., 2014, Ton et al., 2019). High car ownership in the Netherlands indicates significant private vehicle dependency (Zijlstra et al., 2022). The common travel mode also plays a significant role in travel mode choice; it takes more effort for travellers to shift from mode (Ton et al., 2020).

3.2.2. Household characteristics

The second determinant category regards the household characteristics of the decision-maker. Household composition and income are often found to be related to the car-ownership. Household composition concerns the size and the presence of young children in the family. This factor is included in this study because travelling with young children or in large groups often necessitates convenient transport options, like cars, due to ease and flexibility (Ton et al., 2019). Understanding how family dynamics influence mode choice helps tailor transport solutions to meet the needs of households with children or frequent group travel.

Household income is excluded as its impact is already captured under individual income, which provides sufficient insight into economic barriers to sustainable transport (Zijlstra et al., 2022). This focus ensures a more streamlined analysis, emphasising the practical aspects of household structure in travel behaviour. Analysing these household characteristics helps develop targeted strategies that cater to family needs, promoting sustainable travel while ensuring convenience for families in new resident area.

3.2.3. Trip characteristics

Besides the characteristics related to the decision maker, trip characteristics also greatly influence travel behaviour. These determinants such as in-vehicle time, access time and distance, waiting time, transfer time and distance, egress time and distance, travel fare, vehicle type, trip purpose, and departure time play an important role in shaping travel behaviour. This analysis mainly focuses on how these factors impact the choices among the defined mode alternatives: public transit, car, and bicycle users.

In-vehicle time is a crucial determinant for all travel modes. The duration spent in transit significantly influences mode choice. Longer in-vehicle times can deter individuals from using public transport, especially if alternatives like cars or bicycles offer shorter travel durations (Vos, Le, and Kroesen, 2022). Car users, in particular, benefit from typically shorter in-vehicle times, which makes driving a preferred option when speed and convenience are prioritized (Dubey et al., 2024).

Access time and distance are crucial in determining the attractiveness of a transport mode. Convenient access to public transit stops encourages its use, while longer access times and distances can be a significant deterrent. Similarly, the ease of accessing a car or bicycle storage can influence the preference for these modes. The closer and more accessible the starting point, the more likely individuals are to choose that mode of transport(Chen et al., 2021 and Ton et al., 2019).

Waiting time is also relevant to be included because it is directly related to the frequency of service, especially for public transit users. High-frequency services reduce perceived waiting times, improving overall satisfaction and making public transport more competitive. Studies on public transport reliability found that reducing waiting time is paramount for user satisfaction, as frequent and predictable services enhance user experience (Z. Li et al., 2010).

Waiting time at transit stops or in traffic congestion is another critical factor. Longer waiting times can discourage the use of public transport and cars in congested areas, making bicycles a more attractive option due to their predictability and independence from traffic delays (Esfeh et al., 2021). The less time spent waiting, the more appealing the mode becomes(Dubey et al., 2024).

Transfer time and distance add to the total travel time and inconvenience, making direct routes more attractive. Reducing the number of transfers and minimizing the distance between them can enhance the appeal of public transport and multimodal journeys. People generally prefer routes with fewer transfers and shorter transfer times (Ye et al., 2018).

Egress time and distance, or the time and distance required to reach the final destination from the transit stop or parking area, are equally important. Shorter egress times and distances improve the overall convenience of using public transport and cars. The less distance travelled on foot after exiting the primary mode, the more favourable the mode becomes (Ton et al., 2019).

Travel fare plays a significant role, especially for cost-sensitive individuals. Affordable fares can encourage the use of public transit over more expensive alternatives like private cars. Bicycles, known for their minimal operating costs, often attract users looking for cost-effective travel solutions. Financial considerations can significantly sway mode choice decisions (Chen et al., 2021).

Trip purpose significantly affects mode selection. Commuting for work or school typically requires reliable and timely transportation, making public transport or cars preferred options. Conversely, recreational trips often favour active modes like cycling or walking due to their flexible nature and associated health benefits. Understanding the purpose of trips helps in predicting mode choice accurately (Ton et al., 2019).

3.2.4. Urban environment

The urban environment significantly influences travel behaviour and mode choice. High-urbanisation areas typically have higher population densities, leading to increased traffic congestion. Consequently, residents in these areas are more likely to prefer public transit as an alternative to avoid the delays and stress associated with heavy traffic. Public transit systems in highly urbanised areas are often better developed and more efficient, making them a viable and attractive option for daily commuting (Schwanen and Mokhtarian, 2005.

Parking availability is also a crucial factor; limited or expensive parking can discourage car use, encouraging people to opt for public transport or cycling instead (Xue et al., 2024). Being located in a metropolitan area usually means better access to diverse transportation modes and services, which influences the likelihood of using public transport or bicycles over cars. Future plans for car-free areas are also important, as they can promote walking, cycling, and public transport use by reducing reliance on cars and enhancing urban livability (Schwanen and Mokhtarian, 2005).

3.2.5. Work condition

Work conditions, particularly those related to commuting, have a direct impact on mode choice. Commuting reimbursement programs that offer financial incentives for using specific travel modes, such as public transport subsidies or car allowances, can significantly influence commuting choices (Zijlstra et al., 2022, Shen et al., 2016). Additionally, the number of working hours per week affects travel behaviour; longer working hours might encourage faster travel modes like cars, while flexible working hours can favour public transport or cycling (Ton et al., 2020).

3.2.6. External impacts

External factors such as environmental conditions and policies also influence travel behavior. Season and weather characteristics play a significant role; adverse weather conditions can deter cycling and walking, while increasing reliance on cars and public transport (J. Li et al., 2018). Conversely, pleasant weather encourages active travel modes. Noise and air pollution also impact mode choice; high pollution levels can deter walking and cycling, pushing people towards using cars or public transport (Ton et al., 2020). However, in the planning stage of the new urban area, these external factors can not be considered with the mobility design. Therefore, external factors are be excluded in this study.

3.3. Overview of Determinants for Mode Choice Modelling

Understanding the factors influencing commuting mode choice behaviour is critical for developing effective urban mobility strategies. All the analysed determinants in the previous sections are summarised in the table 3.1. These determinants are categorised into individual characteristics, household characteristics, trip characteristics, built environment, and work conditions. Each category includes specific attributes that could impact travellers' decision-making process regarding commuting mode choice behaviour. Because of this, this comprehensive overview serves as a foundation for analysing and constructing the mode choice model.

Table 3.1: Overview of included determinants and its operationalisation

Determinant	Attributes
Individual characteristic	
	Age
	Educational level
	Employment status
	Income
	Travel disability
	Mobility availability
	Habit
Household characteristics	
	Household composition (size and children)
Trip characteristics	
	In-vehicle time
	Travel distance
	Access time and distance
	Waiting time
	Transfer time and distance
	Egress time
	Travel fare
	Vehicle type
	Travel purpose
Build environment	
	Urbanisation level
	Parking availability
Work condition	
	Commuting reimbursement

4

Case Study Rijnenburg

In line with the research objective to enhance sustainable mobility development in the peripheral urban areas, Rijnenburg has been chosen as the study area for this research. This section first provides an overview of Rijnenburg and analyses the current state of the infrastructure and the mobility concepts formed by the local government for the development of Rijnenburg. On top of this, in section 4.2, the hypothetical mobility strategies for Rijnenburg are developed to form the basis for determining the attribute levels of the Stated Preference design in Chapter 5.

4.1. Rijnenburg

Rijnenburg (including Reijerscop) is currently a polder and uninhabited area of around 1238 hectares, located on the southwestern outskirts of the municipality of Utrecht (PosadMaxwan and Goudappel Coffeng, 2020. In the new planning of Utrecht's municipality, the new urban area of Rijnenburg (Kleine Rijnenburg) will have a built-up area of 520 hectares where 20,000 to 25,000 housing units and 10,000 to 15,000 workplaces will be realised (Posad Maxwan and Goudappel Coffeng, 2022 and Gemeente Utrecht, n.d-a). Rijnenburg is expected to have a high urban density, with building densities ranging from forty to one hundred housing units per hectare. The realisation of Rijnenburg should contribute to the housing shortage and stimulate urban development in the metropolitan area (Gemeente Utrecht, n.d-a).

4.1.1. Rijnenburg accessibility

Rijnenburg is located immediately southwest of the Oudenrijn interchange 4.1. It borders the newly constructed neighbourhood Leidsche Rijn to the north by the A12 motorway, the municipality of Nieuwegein to the east by the A2, the municipality of IJsselstein to the south and Meerndijk to the west by the N228 (Wikipedia, 2024). At a distance, Rijnenburg is also connected with the green area Groene Hart and the Amsterdam Rijnskanaal (PosadMaxwan and Goudappel Coffeng, 2020). The location of Rijnenburg is shown in figure 4.1.

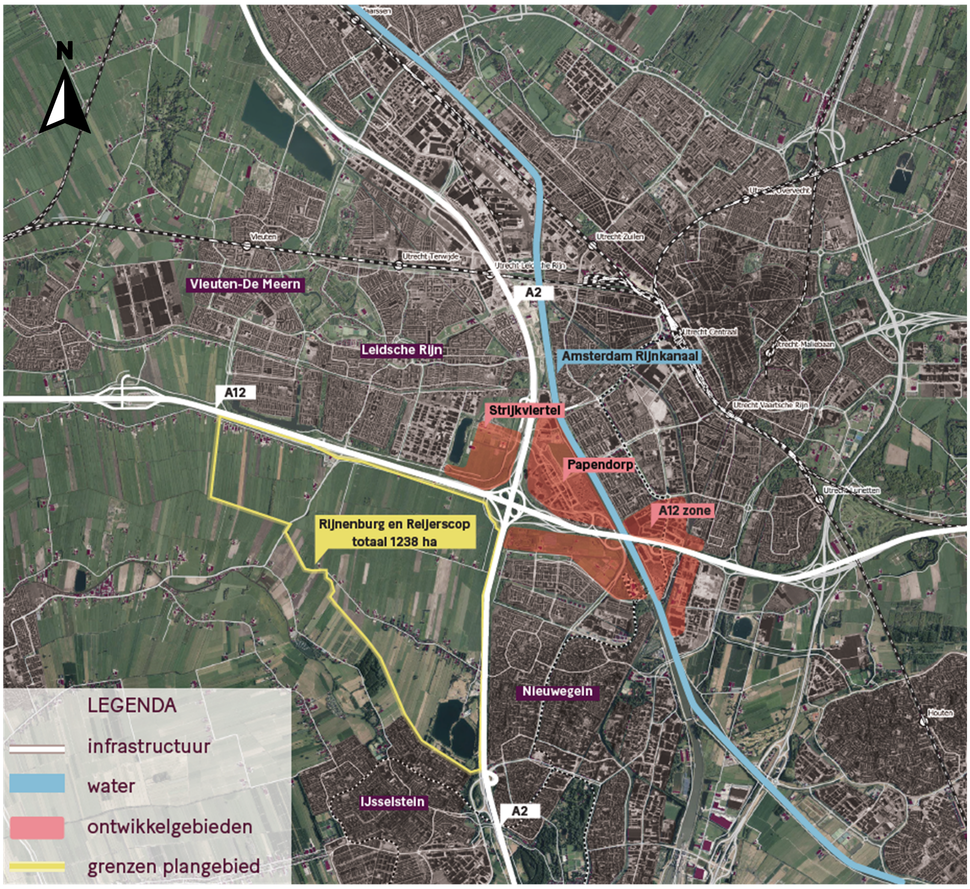


Figure 4.1: Location of Rijnenburg, Utrecht (Gemeente Utrecht, n.d-a)

Public transit

As stated earlier, Rijnenburg is currently an undeveloped and uninhabited area, hence public transportation infrastructure is almost absent. The existing public transportation facilities do not directly serve Rijnenburg but primarily consist of peripheral connections that serve nearby developed areas.

The nearest railway stations are in Utrecht and Leidsche Rijn, see figure 4.1. Utrecht Central Station (Utrecht CS) is the major node of the railway network, offering extensive regional and intercity rail connections. However, all the train stations are located at least five kilometres from Rijnenburg, making it less accessible without a dedicated public transit connection between Rijnenburg and the surrounding.

For the urban public transportation, no tram or light rail lines extend into Rijnenburg. The closest light rail line is the SUNIJ line, which mainly connects the area of Nieuwegein and IJselstein with the city centre of Utrecht via Utrecht CS and with Utrecht Science Park (USP) as the final destination. Currently, this is the only tram line in Utrecht; the local government is investigating the realisation of two more light rail lines, Merwedenlijn and Papendorplijn (Ned, 2024). The realisation of the two new light rail lines creates the opportunity to integrate Rijnenburg into the public transit network of the Utrecht region. Depending on the network planning of these new lines, Rijnenburg can connect with the rest of Utrecht in different ways.

Rijnenburg can be radially integrated into the network, resulting in a spoke-shaped network with Utrecht CS as the bundling node of the network. Secondly, Rijenbrug can also be tangentially connected to the network, forming a wheel-shaped network for the whole of Utrecht's light rail network. And finally, Rijnenburg can also be a branching of the network, with a bundling at a local hub on the edge of the area. The visualisations of the three potential types of light rail connections for Rijnenburg can be found in Appendix A.

Like the light rail services, there are currently no dedicated bus routes that directly serve the Rijnenburg area. The closest bus services operate only in neighbouring areas like Nieuwegein and Leidsche Rijn. These services provide regional connectivity but do not specifically cater to potential future residents of Rijnenburg. However, because buses can also run on usual car infrastructures, Rijnenburg can be better accessed by busses than other modes of public transit. The current public transportation network of Utrecht can be found in Appendix A.

Car

In contrast to public transit, Rijnenburg can be excellently accessed by car at the high network level. The area is surrounded by two important motorways of The Netherlands, A2 and A12, see figure 4.1. The A2 runs close to the eastern boundary of Rijnenburg, providing a direct north-south route that connects Amsterdam and Utrecht with southern parts of the Netherlands. It is one of the busiest highways in the country, facilitating significant passenger and freight traffic (Wegenwiki, 2024b). The A12 forms the northern boundary of Rijnenburg, connecting The Hague, Utrecht, and Arnhem. The A12 is a crucial east-west corridor, contributing to the area's high traffic volumes (Wegenwiki, 2024a). These two motors, together with the N228, provide an outstanding car connection from Rijnenburg to the other areas in the region, to the major cities in the Randstad metropolitan and also to other regions of the country.

Bike

The current cycling infrastructure in Rijnenburg is limited due to its undeveloped status and the presence of significant physical barriers. Within Rijnenburg itself, there are currently few dedicated cycling paths. The existing paths are primarily informal and mainly used for agricultural activities (Fietsknooppunt, n.d). For external connections, the surrounding motorways (mainly A2 and A12) act as terrain barriers for cyclists. Crossing these highways requires dedicated infrastructure, such as overpasses or underpasses, which are currently limited in number and quality. Furthermore, Amsterdam Rijnskanaal also poses another physical barrier for cyclists. Bridges are needed to facilitate safe and direct crossings.

However, the municipality has introduced development plans which emphasise creating a comprehensive internal network of cycling paths and improving external connectivity to the broader Utrecht cycling network and neighbouring areas. Overcoming the challenges of highways and canals through new crossings and integrating cycling into the urban design will be crucial for promoting active transportation and achieving sustainable mobility goals in Rijnenburg (Posad Maxwan and Goudappel Coffeng, 2022).

Conclusion on accessibility of Rijnenburg

Because of the location's characteristics, Rijnenburg is a complex location for sustainable urban development. The close proximity to the motorway leads to increasing use of the motorway for many short and mid-distance trips, resulting in a higher load per housing unit compared to other areas (Studio Bereikbaar, 2023). Compared to the rest of Utrecht, Rijnenburg is situated quite a distance away from the railway network, which will also lead to a reduction in the attractiveness of travelling by train (Studio Bereikbaar, 2022). Moreover, the terrain structure of the area, including the surrounding motorways, forms barriers for inter-region bicycle traffic. All of these result in a threat to sustainable urban development in Rijnenburg, making the area become more car-oriented.

4.1.2. Rijnenburg's Mobility Concepts

In an in-depth study, the municipality of Utrecht, together with Studio Bereikbaar (2023), presents three mobility concepts that could be applied to the development of Rijnenburg. These concepts incorporate various influencing, policy, and design measures that the municipality of Utrecht aims to implement to ensure sustainable urban development.

Conventional

In this mobility concept, transportation options are provided throughout the neighbourhood, aligning with conventional practices seen in recent urban developments. Accessible transportation modes include a light rail to Utrecht CS, a comprehensive network of bicycle paths, and convenient car access, with parking arrangements tailored to anticipated demand.

Neighbourhood planning draws inspiration from established urban areas such as Leidsche Rijn (Utrecht) and Nesselande (Rotterdam), aiming to understand the expected transportation mix without significant intervention. Streets are designed to accommodate all modes of transport, while parking policies comply with established standards within the municipality, considering factors like housing types and density.

Sustainable and urban
The second concept is about sustainability and urbanisation. This mobility concept in Rijnenburg will follow sustainable urban principles, drawing from past projects like Klein Rijnenburg and the Merwedekanaalzone (Posad Maxwan and Goudappel Coffeng, 2022). It prioritises high-quality public transportation HOV (in Dutch hoogwaardig openbaar vervoer), cycling infrastructures, and shared mobility options, aiming to reduce car dependency through neighbourhood parking hubs and lower parking standards.

Following the STOMP principle, the design ensures walking, cycling, and public transit are more convenient than driving (CROW-KpVV, 2021). The neighbourhood integrates elements from various urban developments and is divided into high-density and moderate-density zones. Both zones feature clustered parking within walking distance, with the high-density area having hubs further away and residential streets inaccessible to cars.

In the high-density zone, a 40% replacement of private cars with shared cars is anticipated. At the same time, in the moderate-density area, it's 25%, resulting in a parking standard of 0.42 spaces per household for residents and 0.2 for visitors, alongside 1,900 shared cars. The parking cost for whole Rijnenburg will follow the paid parking standards for urban areas of Utrecht, with discounts for shared mobility and cycling (Gemeente Utrecht, n.d-b).

Ambitious experimental
In the last mobility concept, Rijnenburg is proposed as a car-free urban area, offering facilities at short distances. This concept emphasises the attractiveness of living, walking, and residing, the design provides a smooth integration of urban and natural environments.

As stated in the name, the ambitious experimental concept will put everything in place to reduce car use and car ownership among the residents of Rijnenburg. Walking, cycling, and public transit are prioritised with cars relegated to large hubs at the outskirts, predominantly featuring shared cars. Higher parking fees with the majority of cars replaced by shared vehicles, totalling 2,400 shared cars and 1,500 visitor parking spaces, alongside a minimal 0.05 private cars per household. Although the neighbourhood aims for car-free living, exceptions are made for essential services like waste collection and emergency access.

4.1.3. Expected future population of Rijnenburg
Due to the significant housing shortage in The Netherlands and the Randstad, Rijnenburg can be an attractive residential place for everyone. However, according to Bongers et al. (2023) and Centraal Bureau Voor de Statistiek (2022), young people are affected by housing shortages the most in The Netherlands. This is because of the shortage of affordable, suitable housing and the economic uncertainty among this group.

In addition, the municipality of Utrecht has set a standard for the housing distribution in a new development area. This standard states that 40% of the total upcoming housing units has to be reserved as social housing intended for people with low income, 35% has to be reserved for middle-class incomes and 25% for the free rental sector (Gemeente Utrecht, n.d-c).

Due to the significant housing shortage among young people, the future composition of Rijnenburg's population is expected to consist mainly of young people aged 25 to 34 years old. Within this group, the composition can be different. They can be students, young professionals or young families with diverse income classes due to the set housing distribution standard of the municipality. Other demographic characteristics, such as educational level and employment status, are expected to be the same as the population composition in the province.

4.2. Mobility Strategies

Comprehensive mobility design strategies are essential to achieving sustainable urban development in Rijnenburg. Based on the reviewed area development principles (car-free, TOD and STOMP) in section 2.2, a set of policy strategies for the mobility design in Rijnenburg is developed. These strategies aim to enhance the attractiveness of PT use, holding back on car use and car ownership, preventing the increase of car traffic on the surrounding roads. Therefore, they mainly encompass public transit planning, parking policies in the Rijnenburg area, and other relevant policy measures. This chapter outlines these strategies, focusing on key variables design and how they can address the diverse needs of future residents while promoting sustainability.

4.2.1. Strategy for PT planning

For PT planning strategy, two main design aspects regarding the PT network design and the PT operation's level of service will be considered in the development of the PT design strategy for Rijnenburg. The strategy includes the decision variables of the optimisation problem for the Transit Network Design (TNDP) and taking into account the customer wishes of van Hagen and Peek (2006). Based on the study on Transit Network Design Problem (TNDP) study of van Nes et al. (n.d.) and review of Kepaptsoglou and Karlaftis (2009), the design variables for the PT planning strategy for Rijnenburg are defined as follows:

- **Stop density:**
  The stop density implies the distance between the PT stops and the number of stops that will be realised in Rijneneburg. The stop density directly relates to the access distance to the transit network. Higher stop density means shorter distances to the stop, and the attractiveness of the transit network will increase.
- **Route planning:**
  This variable implies the travel desire lines based on the origin-destination (OD) patterns of the demand and the existing PT network. For Rijnenburg, the route planning also depends on the existing PT network of Utrecht, creating a spoke or a wheel network form for the entire network (see section 4.1.2). The route planning is directly related to the in-vehicle time and the transfer aspects (including transfer time or distance) of the PT services.
- **Line frequency:**
  This variable will be based on the willingness to wait for the traveller. Higher frequency means less waiting time for the user, thus more convenience for the travelers.
- **Type of vehicle:**
  Depending on several aspects (technical, economical and comfort etc), bus (BRT) or tram (lightrail) will be considered in as PT mode for the area.

4.2.2. Strategy for car planning

The second objective of the research is to prevent the increase in car use and car ownership in the area, thereby discouraging residents from developing car-preferred behaviour already from the beginning of their residency in Rijnenburg. Effective car planning is therefore a critical component of urban mobility strategies, especially during the developing stage of the area. Limiting car usage and ownership can be influenced by several measures, such as increasing car taxes, reducing travel speeds, restricting parking availability and more. In the context of urban development, car parking is an efficient policy instrument to influence the choice of using cars (Action, 2005). The limitation of parking availability in this research can be managed through two design variables: the strategic location of parking facilities and the imposition of parking fees.

- **Parking location** The location of parking facilities directly impacts the accessibility of vehicles and influences how dependent residents are on cars. Parking at home offers the shortest access distance to the vehicle, providing the highest convenience for car usage. Another option is having communal parking for the whole neighbourhood; the access distance is therefore longer, but it offers car-free street. Lastly, the most restrictive approach is to make the whole area of Rijnenburg car-free by locating parking facilities at the edge of the residential area. This results in the longest access distance for the users.

- **Parking fee** Setting high parking fees is another crucial measure to discourage car ownership and usage. High parking costs can effectively reduce car dependency by making car ownership less financially attractive. By implementing significant fees for parking, especially in residential and central areas, residents are encouraged to use alternative modes of transportation such as public transit, cycling, and walking. The parking cost can be determined based on the government standard and the desired parking availability level.

4.2.3. Other aspects

Besides public transportation and car-related strategies, promoting the use of e-bikes can significantly encourage active transportation. While the government may not directly provide e-bikes, it can stimulate companies to offer them to employees through reimbursement programs or subsidies. By supporting infrastructure such as bike lanes, secure parking, and charging stations, and promoting e-bike benefits through public awareness campaigns, the government can create a conducive environment for e-bike adoption. This approach reduces car dependency and promotes a healthier, more sustainable lifestyle for Rijnenburg residents.

4.3. Conclusion on Mobility Strategies

On top of the strategies proposed previously, strategic design scenarios for Rijnenburg have been developed. These scenarios are directly retrieved from the framed mobility concepts in section 4.1.2. The overview of the strategies for each scenario is given in table 4.2.

1. Conventional

The conventional design principles will be applied to Rijnenburg in the Conventional scenario. Cars can be parked at home, providing cheap and convenient parking. This setup leads to lower demand for public transit, resulting in a lower PT level of service with low stop density and operation frequency. Also, due to the low demand, the Utrecht CS is the best location for passenger bundling, resulting in requiring transfers for trips with further destinations. Roads are designed mainly for cars, for this reason buses are the best choice for serving the area of Rijnenburg.

2. Sustainable

The Sustainable scenario offers a balanced approach for all traffic modes. By having cars parked outside of the streets, which provides a car-free environment for the neighbourhoods and reduces dependency on cars compared to the Conventional scenario. The quality of the PT services will also be improved. More stops should be placed to reduce access, and the frequency of the stops has to be increased for a shorter waiting time. This setup is expected to create a certain amount of modal shift from car to PT, leading to increasing PT demand. The wheel network should be applied to provide the robustness of the services. However, this demand may not be enough to have direct connections to everywhere in the network. Most of the detours can be avoided, but the transfers on the links with low demand could enhance the overall efficiency of the PT system.

3. Ambitious

The Ambitious scenario adopts Transit-Oriented Development (TOD) principles, aiming for a car-free environment within the whole area of Rijnenburg. In this scenario, there are no car parking facilities at home or close to home. The few parking facilities will be located outside the residential area, at the area entries. Public transit should have a robust network to reduce the detours and transfers during the trips. Furthermore, the cycling infrastructure should be optimal to promote the use of active mode, and for longer distance trips, commuting reimbursement for e-bikes can be experimented with. This setup significantly aims to minimise car dependency and promotes a high reliance on public transit and active transportation.

Table 4.1: Overview of design strategies and applicable levels for each development scenario

	Variable	**Conventional**	**Sustainable**	**Ambitious**
PT	Stop density	Low	Moderate	High
	Frequencies	Low	Moderate	High
	Route planning	Spoke	Wheel	Wheel
	Route planning	Transfer at CS	Transfer at normal stop	Direct
	Type of vehicle	Bus	Bus/Tram	Tram
Car	Parking location	Home parking	Outside the neighbourhood	Outside the area
	Parking cost	Standard	Standard	High
Bike	Vehicle type	Standard	Standard	E-bike

5 Research Methodology

Building upon the conceptual framework discussed in section 2.4, the methodology adopted for this research is presented in this chapter. First, the survey design for the Stated Preference (SP) experiment is developed using the developed mobility design strategies for Rijnenburg in Chapter 4. In section 5.2, the data analysis method for the collected data will be given.

5.1. Survey Design

From the theoretical framework in Chapter 3 and the case study of Rijnenburg, a survey is designed in order to collect the data for the construction of the discrete choice model. The survey consists of two parts: a Stated Preference experiment and a questionnaire regarding the personal characteristics of the decision-maker.

5.1.1. Stated Preference experiment

As mentioned in the literature review, a Stated Preference experiment is performed to collect the mode preference for the input of the discrete choice model. This choice experiment consists of three main parts. The context scenario, the mode alternatives and the alternative's attributes with their levels. The context scenario describes in which context the choice should be made. The mode alternatives consist of three main alternatives: public transit (PT), car and bike, according to section 3.1. And the attributes are the properties of the mode alternatives with their specific levels.

Context scenario

The context scenario reflects a mode choice for daily commuting with Rijnenburg as the resident location (origin). In this context, Rijnenburg is a high-density building area located on the outskirts of Utrecht and next to two busy motorways, A2 and A12. Based on the target travel distance and the target group, the Utrecht Science Park (USP) is chosen as the commuting destination for the experiment. This is the biggest science park in The Netherlands, and therefore it is expected to be the destination of the most commuting trips in Utrecht (Utrecht University, 2018). USP is located around ten kilometres as the crow flies distance from Rijnenburg. The shortest travel route from Rijnenburg and the USP will be shown for the car and bike alternative; for the PT route, it is indicated by a straight dashed line. Figure 5.1 visualises the context scenario for the experiment. This context scenario includes the fixed determinants: trip purpose, egress time, fixed costs for car ownership, travel distance, travel fare, high urban density, and location of Rijnenburg.

The egress distance is assumed to be five minutes for all choice situations, and the travel fare (gas or PT tickets) is assumed to be covered by the employer. In addition, a fixed cost of €300 per month will be applied for the car alternative; this amount includes the maintenance costs and taxes (Nibud, 2023). In real life, those egress distances and costs are always different for everyone. However, this study mainly focuses on the mobility strategies for resident area Rijnenburg; it is chosen to simplify the costs and the destination characteristics to ensure the effectiveness of the experiment (Arentze and Molin, 2013).

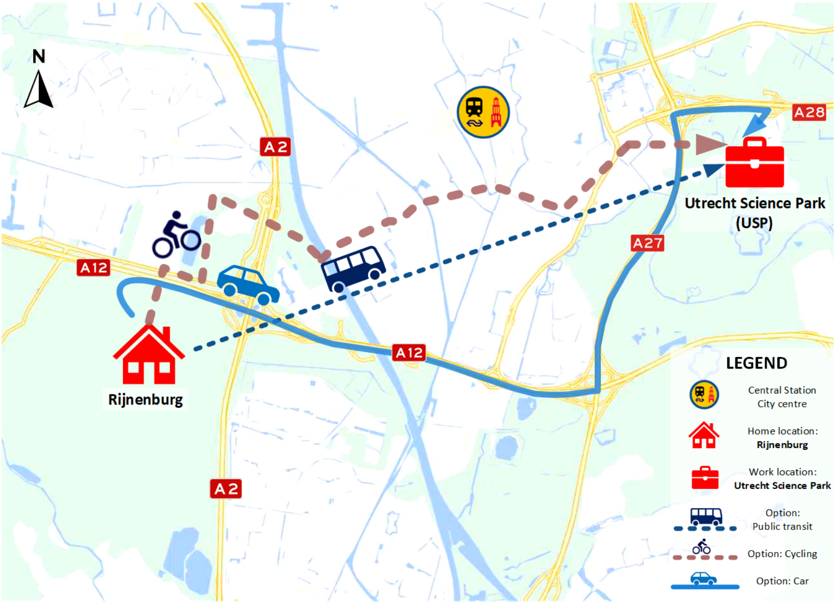


Figure 5.1: Situation maps for the context situation in the experiment

5.1.2. Attributes and levels

For the design of the SP experiment, hypothetical attributes and attributes are needed. The developed design variables for different development scenarios for Rijnenburg in table 4.2 are used to define the levels of the corresponding attributes. For the efficiency of the experiment, some of these variables are converted to the equivalent attributes from the literature in section 3.2. So, the access time is the attribute of the stop density for the PT planning and the parking location for the car planning. The waiting time is the attribute of the frequency. The route planning variable is determined by the in-vehicle time, the transfer time and distance. The vehicle type for the bike alternative will be applied as in-vehicle time. PT vehicle type and car parking cost remain the same. Figure 5.2 provides an overview for converting from the design variables to the model attributes.

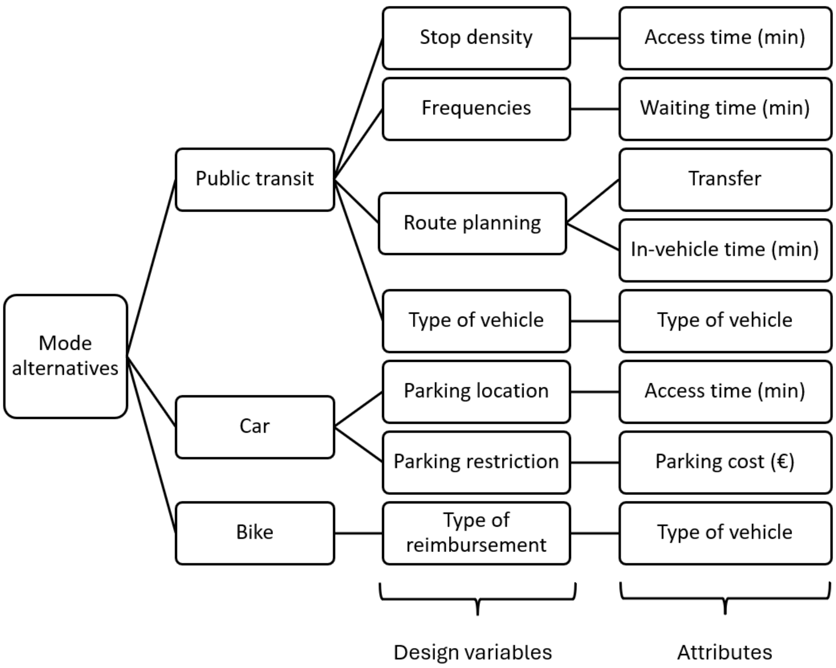


Figure 5.2: Relationship between design variables and attributes

Access time
The first attribute is the access time, which is defined as the walking time from the front door to the vehicle, with a walking speed of 1.14 m/s (Wirtz and Ries, 1992). The different levels for each mode alternative are determined based on the three mobility concepts for Rijnenburg; see section 4.1.2.

- **PT**: There are three variables for access time to the PT service, e.g. to the stop, based on the considered number of the stops that will be realised in Rijnenburg. According to De Ridder (2023), the optimal average stop spacing on suburban PT lines is more than 800 metres. When combining this optimal stop spacing with the considering number of stops in the whole area as 2, 3 or 4, the expected stop spacing for Rijnenburg is equivalent to 800, 1000 and 1300 metres. The access distance is assumed to be half of the stop spacing. This results in three levels of access time for PT, namely 5, 7 and 10 minutes.

- **Car**: For the car alternative, three parking scenarios define the access distance variables. In the conventional scenario, people will park their car in front of the house. For the second scenario, the parking standards will be reduced, and residents can only park their cars in parking spots that will be located on the edge of the neighbourhood. In the last scenario, the access distance is even longer, cars can only be parked in the hubs. These hubs will be located on the edge of the area, close to the access roads of the area. The access time variables are respectively 0, 5 and 15 minutes.

- **Bike**: For the bicycle alternative, the access time is set to 0. This is based on the cycling culture in The Netherlands, which means bicycles can always be parked close to the house. In front of the house, in the garden or bike parking in the apartment buildings. Therefore, the access time for bike alternative will not be varied.

Waiting time
The second attribute is the waiting time. This attribute is only applied to the PT alternative; since car drivers and cyclists are independent, they do not have to wait to embark on the vehicle. Most trams and busses in The Netherlands, particularly in the Randstad, run every 10 or 15 minutes depending on the passenger demand in the area and the operational time, during the day or late in the evening (U-OV, n.d-a, U-OV, n.d-b). The tram operation in Utrecht, U-tram, sometimes operates a frequency of ten times per hour per direction. The waiting time is generally defined as half of the headway between two vehicles, and therefore three levels of the waiting attributes will be 3, 5 and 7.5 minutes (Esfeh et al., 2021).

In-vehicle time
The in-vehicle time includes the time the traveller actually is in the vehicle. This includes driving time for the car, cycling time for the bike alternative and in-vehicle time when the traveller chooses PT.

- **PT**: The in-vehicle time levels of the PT alternative are determined based on the three design alternatives from the in-depth study of the Studio Bereikbaar (2023) and timetable of the current bus and tramlines between USP and the reference area IJselstein. For the shortest route (direct PT line to USP), the in-vehicle time will be 31 minutes. The longest in-vehicle time is 41 minutes, this is the case when all PT lines have to pass the Central Station.

- **Car**: For the car alternative, Google Maps routing has been used for defining the driving time from Rijnenburg to the USP. The driving time attribute for the car alternative is set for two scenarios, with and without congestion. In the congestion (rush hours Tuesday morning), the shortest driving time for car is around 43 minutes (40-45 min), and when there is no congestion, the shortest driving is reduced to 13 minutes(10-15min) (Google Maps, n.d).

- **Bike**: For the bike alternative, two types of cycling time are used, 50 and 38 minutes, which are respectively with the cycling time for a normal bike and for an electrical bike (e-bike). Hereby, the average cycling speed is assumed to be 17 km/h for a normal bike and 25 km/h for an e-bike (Google Maps, n.d, Rijksoverheid, n.d.).

Transfer
The transfer attribute is only applicable for the PT alternative, and its levels are based on the three connecting alternatives for Rijnenburg, see Appendix A. Three types of transfer are included in the design. For a direct PT route, no transfer is applied. For alternative C, there will be a transfer at the same stop. And for alternative A, a transfer at the central station with a walking distance of 400 metre will be applied.

For the calculation of the transfer time, the walking transfer speed is assumed to be 1.1 m/s (De Ridder, 2023). Also taking a possible waiting time for the next vehicle into account. The transfer time levels are therefore 0, 5 and 8 minutes. To avoid confusion, only the transfer distance will be visualised, the transfer time will be shown together with the in-vehicle time.

Vehicle type
Different vehicle type is applied for PT and bike alternatives; see section 3.1. For PT, tram and bus are two levels of the attribute. For the bike alternative, bike and e-bike are mainly different in travel speed (or cycling time). Because of this, bike and e-bike are only visualised in the choice set and for choice set generating, only the cycling time is used in the utility function.

Parking cost
Rijnenburg will be an area with high-density urbanisation according to section 4.1, free (street)parking seems unlikely, so this option will not be given. The parking cost attributes, therefore, have only two levels, the maximum and the minimum parking cost, which could be applied in the Rijnenburg area. The minimum level is based on the parking fare for residents in zone B1 of the municipality of Utrecht and Nieuwegein (Gemeente Utrecht, n.d-b, Gemeente Nieuwegein, n.d). The higher scenario will have a parking cost of €60 per month. This is based on the parking fare for residents in zone A1, the city centre of Utrecht (Gemeente Nieuwegein, n.d).

All attributes and corresponding levels above are summarised in the table 5.1.

Table 5.1: Overview of attributes and levels for SP experiment

	Mode alternatives		
Attributes	**PT**	**Car**	**Bike**
Access time (min)	5, 7, 10	0, 5, 15	-
Waiting time (min)	3, 5, 7.5	-	-
In-vehicle time (min)	31, 41	13, 43	38, 50
Transfer	No, Yes	-	-
(Transfer distance (m))	(0, 400)		
Type of vehicle	Tram or bus	-	(Bike or e-bike)
Parking cost (€)	-	10, 60	-

5.1.3. Decision maker characteristics
After the choice experiment part, each respondent has been asked to provide some information related to the personal characteristics from section 3.2. These variables are measured in different scales with the proposed categories. Table 5.2 gives the overview of asked personal attributes and their categories.

Table 5.2: Overview of individual characteristics

Attributes	Measurement scale	Variables
Age group	Nominal	Younger than 18 yr
		18 - 24 yr
		25 - 34 yr
		35 - 44 yr
		45 - 54 yr
		55 - 64 yr
		Older than 64 yr
Household composition	Nominal	Single without children
		Single with children
		Couple without children
		Couple with children
Educational levels	Nominal	High school
		Vocational Education and Trainning (VET)
		Bachelor’s degree
		Master/PhD
Employment status	Nominal	Student
		Job seeker
		Full time employed
		Part time employed
		Self employed
		Retired
Individual income	Ordinal	Lower than €833
		€833 - €2,310
		€2,310 - €3,975
		€3,975 - €4,798
		Higher than €4,798
Mobility disability	Nominal	Yes
		No
Mobility availability	Nominal	Driving license
		Private car
		Leased car
		Bike
		E-bike | Scooter | Moped
		PT subscription
Commuting reimbursement	Nominal	No reimbursement
		Mileage reimbursement
		Leased car
		PT subsciption
Commuting frequent mode	Nominal	PT
		Car
		Bike

5.1.4. Choice set generation and design

After defining the attributes and attributes levels, the choice sets for the Stated Preference experiment will be constructed using Ngene software (Cranenburgh and Collins, 2019). Since this experiment has never been done before, no priors can be applied; therefore, a fractional factorial orthogonal design is chosen for generating the choice set. This type of design has the advantage of a smaller number of required choice sets, but no interaction effects between the attributes are covered, therefore, it only allows the estimate of the main effects. Therefore, the correlations between attributes are assumed to be zero. This leads to low standard errors and thus reliable parameters (Molin, 2024).

Furthermore, the design will have simultaneous construction. This type of construction applies to labelled alternatives (PT, car and bike). For simultaneous construction, it also holds that there are no correlations within and between the alternatives (Cranenburgh and Collins, 2019). For a simultaneous construction, it is also required to have alternative specific attributes. The disadvantage of this type of construction is more choice sets are needed.

To determine the number of choice tasks that will be given in the survey, the minimum required number of choice sets for the model should be given in the syntax for Ngene. This required minimum is based on the number of to-estimate parameters, or β's, and the number of choice alternatives. According to table 5.1 and table 5.2, there will be 18 estimation parameters, including the alternative's specific constants (ASCs). Therefore, the minimum required number of choice sets can be calculated with equation 5.1, according to Cranenburgh and Collins (2019).

$$\textit{Number of choice sets} = \frac{\textit{Number of parameters}}{\textit{Number of alternatives - 1}} = \frac{17}{2} = 8.5 \tag{5.1}$$

Using the defined input syntax, Ngene provided an experiment design of 36 choice sets, which means each respondent will face 36 choice situations. This large number may reduce the experiment's effectiveness (Arentze and Molin, 2013). To reduce the number of choice tasks for each respondent, the blocking procedure of Ngene is applied. As a result, the total number of 36 choice tasks is divided into four blocks with nine choice situations per block, which is an acceptable number of choice tasks for participants. These four blocks will be randomised and equally distributed. The complete input code and choice set output (final design) from Ngene can be found in Appendix B.

5.1.5. Survey distribution

The survey is constructed with the only survey platform Qualtrics, in Dutch and English at the B1 language level; this allows for a large group to understand the survey and the experiment correctly. According to the calculation of Qualtrics, it will take around 8 minutes to finish the survey. The completed survey can be found in Appendix C. An example of a choice task provided to the respondents in the survey is shown in figure 5.3.

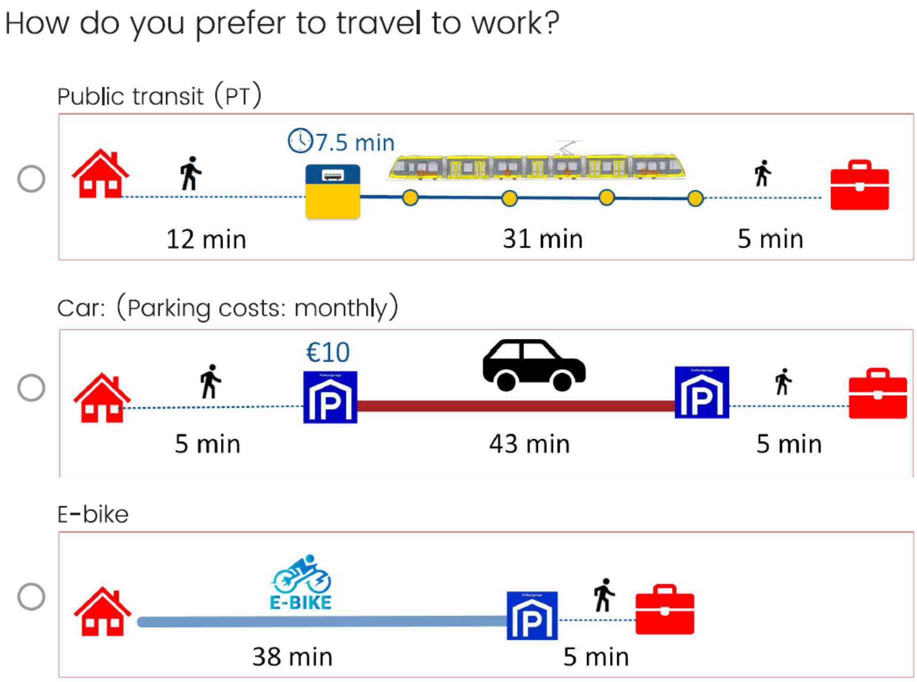


Figure 5.3: Example of choice situation

This research aims to predict the travel behaviour of Rijnenburg's residents; therefore, they are also the target group for the survey. However, they have not been defined yet, since Rijnenburg is still an empty area. Therefore, the survey's target group includes people who are most likely to live there or are expected to have the same characteristics as future residents of Rijnenburg. Hence, the survey is distributed via communication channels that can approach residents of the province of Utrecht, such as information sessions and social media groups of different communities of Utrecht. Furthermore, students and young professionals in the Randstad metropolitan are been approached for the survey since they have the potential to live in Rijnenburg in the future.

Required sample size

To have a reliable result for the discrete choice model, a minimum sample size is required. There are different methods for the determination of the minimum required sample size, the most commonly used rule of thumb to determine the required sample size for a discrete choice experiment was proposed by Johnson and Orme (2003). The rule indicates that the number of respondents should satisfy the inequality of Orme (1998), given in equation 5.2:

$$N \geq 500 \cdot \frac{L^{max}}{J \cdot S} \tag{5.2}$$

In this equation, N is the minimum acceptable sample size, L^{max} = the largest number of levels for any of the attributes, J is the number of alternatives, and S is the number of choice sets each respondent received. Applying in this study:

$$N \geq 500 \cdot \frac{3}{3 \cdot 9} = 56 \tag{5.3}$$

Since the total number of generated choice sets 36 is divided into four blocks, this means four times N from equation 5.3 is required. This results in a minimum acceptable sample size, N ≥224 responses for this research.

5.2. Data Analysis Methodology

After all data are collected, the collected dataset is analysed. This data analysis section includes the methodology for data preparation, the specification of the utilities, the estimation of the discrete model and the validation of the estimated the model.

5.2.1. Data preparation methodology

The condition of the dataset has a very important role in every study; the data quality directly affects the reliability of the results. Before analysing the dataset, the collected dataset needs to be prepared to enhance the quality of the dataset. The data preparation consists of two steps: data cleaning and data integration.

Data cleaning should ensure the completeness and correctness of data since not all of the collected surveys are completely and properly filled. First, all answers that have not completed the first part of the survey (SP experiment) are removed. This is because the choice preference analysis can not be carried out with an uncompleted choice set. This part accounts for 50% of the survey progress. Second, the other uncompleted answers with a duration shorter than 2.5 minutes are also removed from the dataset; this is the minimum required time to read all the information on the survey. A shorter survey duration will provide unreliable data. Besides, all missing data and options 'other/ rather not answer' in the personal characteristics questions will be set as 9999 since these values cannot be used for model estimation.

After data filtering, some of the categories with similar characteristics will be merged if the number of data in a category is too low (< 30). Although, exceptions can be made for the minor categories with irreplaceable characteristics. This data integration handling should enhance the individual categories to have a sufficient number of values and improve the statistical power of the analysis.

Statistical test

To examine the differences between the collected data sample and the intended population, as well as the differences in mode preferences across various socio-demographic groups, a Chi-square test was performed. The Chi-square value is calculated using the following formula:

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} \tag{5.4}$$

In this equation, $O_i =$ is the observed frequency and $E_i =$ is the expected frequency for category i. This will be calculated based on the population. Based on the calculated Chi-square values, the probability values (p-value) for all data variables will be computed to assess the statistical significance of the difference.

The results of a Chi-square test indicate whether there is a statistically significant difference between observed and expected frequencies, with a confidence level of 95%. A high Chi-square value with a low p-value (< 0.05) suggests a significant difference, meaning the intended variable strongly influences the outcome (like mode choice). A low Chi-square value with a high p-value (> 0.05) indicates no significant difference, meaning the variable likely has little to no effect. Extremely high Chi-square values with very low p-values confirm a major impact of the variable on the outcome. These results help determine which variables are important for further analysis or model construction.

5.2.2. Utility specification

The final dataset will be analysed under the conventional random utility maximisation theory (RUM) (Ben-Akiva and Bierlaire, 2003). According to this theory, the utility that individual n associates with alternative i within the choice set C_n is expressed as:

$$U_{in} = V_{in} + \varepsilon_{in} \tag{5.5}$$

In this equation, U_{in} represents the utility for person n of alternative i, which is composed between a systematic V and a random ε component, the so-called error term (Ben-Akiva and Bierlaire, 2003). In

RUM models, the systematic part V is a linear-additive function of observed attributes of alternative i and characteristics of decision maker n

$$V_{in} = \sum_{m} \beta_m \cdot Z_{inm} + \sum_{k} \beta_k \cdot Z_{nk} \tag{5.6}$$

Herein, Z_{in} is a vector of m experimented attributes of alternative i for decision maker n and S_n is a vector of k characteristics of decision maker n. The vector β consists of coefficients for these m and k observed variables, which need to be estimated from the collected dataset. Besides, an alternative specific constant (ASC) should be added to the utility function for all but one reference alternative. This constant captures the preferences that are not fully explained by the given attributes. This adjustment improves the model's explanatory power by removing unobserved attributes from the error term (Ben-Akiva and Bierlaire, 2003).

In this research, the estimated model parameters β's for mobility design strategies will describe their influence on the mode choice of travellers and identify effective mobility measures. Together with other estimated parameters for model attributes and ASCs, these values will help estimate the probability of selecting a particular mode. Furthermore, considering interaction terms between alternative attributes and personal characteristics allows for assessing whether the effect of certain attributes varies across specific socio-demographic groups. This analysis will offer insights into how different user groups influence the effectiveness of mobility strategies. The parameters β's of the model will be estimated using Python-Biogeme (Bierlaire, 2020).

Because only the systematic part of the utility function can be observed, the probability of an alternative being chosen will be estimated as the following equation.

$$P_i = Prob(U_i > U_j, \forall i \neq j) \tag{5.7}$$

5.2.3. Model estimation

Choice probabilities depend on the error term's distribution. If errors follow a standard normal distribution, it results in probit models. Alternatively, assuming errors are independent and identically distributed (i.i.d.) and following the Extreme Value Type I (Gumbel) distribution leads to logit models (Ben-Akiva and Bierlaire, 2003). By far, the easiest and most widely used discrete choice model is the Multinomial Logit model (MNL).

Multinomial Logit model (MNL)

MNL models offer computational efficiency due to closed-form probabilities by assuming i.i.d. error terms, implying alternatives do not share unobserved characteristics. (Ben-Akiva and Bierlaire, 2003). The probability of choosing a particular alternative i according to the MNL model can be calculated as follows:

$$P(i \mid C_n) = \frac{e^{V_i}}{\sum_{i=1}^{n} e^{V_i}} \tag{5.8}$$

Despite the user-friendliness of the MNL model, it can only be used as a base model for the parameter estimation due to various limitations. It does not account for random taste variation, unrestricted substitution patterns, or correlation in unobserved factors over time (Train, 2002). Such correlations may occur when multiple choices of the same individual are observed, which is the case in this research. Therefore, a more advanced model is required to address these issues.

Panel Mixed Logit model (ML)

As described in the previous section, a more advanced model is required for the model estimation to capture the correlation structure in panel datasets. The Panel Mixed Logit model is, therefore a suitable model to apply in this research. This model can accommodate taste heterogeneity by treating the estimated parameters of the utility function as random parameters. Instead of having fixed values, these parameters are drawn from a probability distribution, allowing for variation in individual preferences (Train, 2002).

Panel ML models are defined by their choice probabilities' functional form and can be derived from various behavioural specifications. These models integrate standard logit probabilities over a parameter density (Train, 2002). According to Train (2002), a Panel ML model is any model whose choice probabilities that can be expressed as:

$$P_{ni} = \int L_{ni}(\beta) \cdot f(\beta) d\beta \tag{5.9}$$

where:

$$L_{ni}(\beta) = \frac{e^{V_{ni}(\beta)}}{\sum_{j=1}^{J} e^{V_{ni}(\beta)}} \tag{5.10}$$

Herein, $L_{ni}(\beta)$ is the logit probability evaluated at model parameter β and $f(\beta)$ is a density function. $V_n i(\beta)$ is a portion of utility, also depends on parameters β. Assuming that all utilities are linear in β, $V_{ni}(\beta)$ is equal to $\beta' x_{ni}$, where x_{in} are observed attributes of alternative i for individual n and β_n is the vector of random coefficients that varies across individuals (the panel effect). The expression of the probability function for the Panel ML model therefore becomes:

$$P_{ni} = \int (\frac{e^{\beta' x_{ni}}}{\sum_{j=1}^{J} e^{\beta' x_{ni}}}) \cdot f(\beta) d\beta \tag{5.11}$$

Herein, the panel effect for person n can be expressed as:

$$\beta_n' = \beta_0 + \sigma_{panel} \cdot v_i \tag{5.12}$$

With β_0 (equal to 0) is a fixed parameter set to zero the mean of the distribution, σ_{panel} represents the standard deviation of the random component and $v_i \sim N(0,1)$ is a random draw from a standard normal distribution.

Estimation strategy and goodness of fit
The discrete choice model is estimated based on the Maximum likelihood principle. This is the most widely used technique and based on the set of parameters chosen to make the data most likely (Ben-Akiva and Bierlaire, 2003). The model fit or the model performance can be determined by the model performance indicators, which can be found in the estimation output file. The indtended model performance indicators are: rho-square (ρ^2), adjusted rho-square ($\bar{\rho}^2$), Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) and the Likelihood ratio test (Train, 2002).

The rho-square compares the log-likelihood of the estimated model (LL_final) to the log-likelihood of the initial model with only constants (LL_initial), using $\rho^2 = 1 - \frac{LL\_final}{LL\_initial}$. The adjusted rho-square corrects for the number of estimated parameters; the higher $\bar{\rho}^2$ value, the better the explanatory power of the model, and thus the better the model fit. Both AIC and BIC aim to balance model fit and complexity; for comparing the models during the estimation, the lower AIC and BIC values, the better the model. Lastly, in order to assess the restrictiveness of the parameters the likelihood ratio test was also conducted. In this test, the loglikelihood of a restrict and a unrestricted model is compared with $D = -2 \cdot (LL_{final_R} - LL_{final_U})$. The higher the value of the likelihood ratio test, the better the model fit.

For all models, model estimation is conducted using the forward stepwise estimation strategy to identify significant attributes influencing discrete choices among alternatives. Firstly, an initial model (LL_initial) is defined as one that only contains the Alternative Specific Constants (ASCs), capturing the preferences for a specific mode alternative. Then, the attributes were added one at a time to the initial model, and the model performance was calculated. The inclusion of an attribute has been considered based on the assessment of the model performance indicators. When the model performance has been improved, the included attribute will be kept, and the estimation process will continue with the next. Otherwise, the concerning attributes will be neglected.

Additionally, a parameter will also be excluded from the model if it is not significant. The significance of a parameter is indicated by the t-test with the corresponding probability value (p-value). For a confidence level of 95%, a parameter with a p-value higher than 0.05 means it is not statistically significant and may not have a meaningful impact on the dependent variable. This means all model parameters p-value higher than 0.05 will also be neglected.

This process continues until all potential attributes have been tested. The final model includes the ASCs and only those attributes that significantly improve performance. This strategy allows the evaluation of the significance of each attribute. By excluding attributes that do not significantly improve the model, the forward stepwise approach helps prevent overfitting, enhancing the model's generalisability to new data.

5.2.4. Model simulation

In the context of both validating and applying the models, the stochastic nature of the commuter's choice process necessitates a robust simulation approach. The simulation begins with sample enumeration. This method is highly flexible for predicting how changes in the choice set or conditions affect various segments of the population. It uses a sample to represent the entire population. The sample can be drawn randomly, and the predicted share of the sample choosing alternative i is used as an estimate for W(i):

$$W(i) = \frac{1}{N_s} \sum_{n=1}^{N_s} P(i|x_n) \tag{5.13}$$

Once an individual is selected, the model computes the utilities and probabilities for each of the three mode alternatives, using equation 5.8 and 5.9. Following the probability calculation, a mode choice is drawn for the individual. This draw is random, yet based on the calculated probabilities, meaning that the mode with the highest probability is more likely to be selected. Still, all alternatives have a chance of being chosen due to the properties of discrete choice modelling. By repeating the process with the population size, the modal spit can be determined.

However, to ensure that the results are not merely a product of random variation, the entire process, from drawing an individual to calculating the modal split is repeated multiple times. This repetition allows the simulation to produce results that are stable and representative of the broader population's commuting behaviour. The required minimum number of repetitions or number of runs (N) is determined by:

$$N \geq \frac{Z_{\alpha/2}^2}{(\alpha \cdot \mu)^2} \cdot \sigma^2 \tag{5.14}$$

In this equation, α gives the area of the tails outside the desired confidence interval on a normal distribution. So, with a desired confidence interval of 95%, α is 0.05. $Z_{\alpha/2}$ is the corresponding Z-score of α, which is 1.96 in this case μ and σ are the mean and standard deviation of the modal splits. Because μ and σ are unknown initially, the simulation needs to start with a small number of runs. After that, the required minimum number of runs can be computed based on the results of these runs.

Finally, after determining the necessary number of runs, the full simulation is conducted. The results from this simulation are used to predict the modal shares, providing valuable insights that are both stable and applicable for validating and applying the models.

6

Results of Data Analysis

Applying the described methodology in the chapter 5, the result of the data analysis steps is obtained and described in this chapter. The result concerns the summary of the data collection as Descriptive Statistics in section 6.1 and the construction of the Discrete Choice Model in section 6.2.

6.1. Descriptive Statistics

The survey was distributed from April 8^{th} to 1^{st} of June 2024. In this section, the descriptive statistics of the collected dataset will be given. The collected dataset is first prepared by cleaning the inconsistencies and missing values and integrating the categories in the dataset. After the preparation of the collected data, the final dataset is analysed statistically and presented under personal characteristics. Finally, the commuting mode choice of the SP experiment has been analysed. In total, the survey was filled by 291 respondents; however, not all of them completed the survey properly. Following the data cleaning steps in section 5.2.1, a data sample of 200 valid responses is obtained.

The survey was initially designed with detailed categories to capture a broad range of respondent characteristics, allowing for a nuanced understanding of different demographic and behavioural factors. This level of specificity aimed to assess whether specific subgroups exhibited distinct preferences that might be overlooked in broader categories. However, during data analysis, it became evident that some categories had too few responses (under 30), potentially weakening the statistical power. To address this, categories with similar characteristics were combined, ensuring sufficient sample sizes while enhancing the statistical power of the analysis.

First of all, the number of respondents in the all age categories, except two group 18 to 24 and 25 to 34, is lower 30. Consider expected similar characteristics of students, junior and senior, three new age categories are formed, 15 to 24, 25 to 44 and 45 to 64 year. The categories of respondents older than 64 years is removed in consideration of very low number of responses (only 7) and retired characteristic. For the household composition, both single with children and couples with children categories will be merged into one category: the presence of children. This assumes there is no difference in choice preference for single or couples when there is the presence of children in the household. This helps to increase the sample size for the group presence of children.

For the educational level, the high school and Vocational Education and Training (VET) educated categories will be merged into low-educated according to the definition of CBS (2024). For the employment status, the student category will be kept; the new employment category will include both full-time and part-time employment, assuming the number of working hours will not greatly affect their mode preference. Other categories will be marked as other, since they are significantly much smaller and have to consider different factors in their mode preference. Lastly, the income will be reduced from five to three categories based on the income divisions for one-person housing, according to Woonnet Haaglanden (2024). Lastly, the availability of e-bike, scooter and moped will also be merged into one category, since these vehicles are mostly similar in characteristics.

Although the number of respondents in the category of people with disabilities is lower than 30, this group is retained in the analysis due to its critical importance in understanding mobility and choice preferences, as their experiences may significantly differ from the general population, with no other category able to represent their unique perspectives. Similarly, the ”leased car as reimbursement” category is retained despite its small sample size because it represents a distinct group whose commuting behaviour is influenced by employer-provided vehicle benefits. Excluding these categories would miss vital insights, while retaining them ensures their unique perspectives are included in the analysis, despite smaller sample sizes.

The data preparation stage results in a sample size of 200 answers with only nine answers with progress lower than 100% but still higher than 75% . For the SP experiment, each answer included nine choices, which means 1800 choice situations were observed in the survey. All categories contain more than 30 respondents, except for the group with disability and availeble of leased car.

6.1.1. Sample characteristics

The sample characteristics from the collected dataset are given in table 6.1 with the count and percentage of each aggregated category. The collected dataset is then compared to the population of the province of Utrecht, these statistics are gathered by the Statistics Netherlands (Staat van Utrecht, 2023). Since the statistics for 2024 are not available yet, the statistics from 2023 is used. Table 6.1 shows the overview of the collected dataset and the corresponding statistics of Utrecht’s population

For the assessment of statistical differences between the collected sample and the population, the Chi-square test is applied to calculate the p-values for the categories where it is possible. The results of the test is given in the table 6.2. Overall, it can be observed that all p-values are smaller than 0.05. For a confidence level of 95%, the low p-values indicate that there are significant differences between the collected sample and the population of Utrecht province.

Table 6.1: Sample characteristics

Variables	Survey respondents		Population Utrecht	
	Count	**Percentage(%)**	**Count**	**Percentage(%)**
Age group				
15 - 24 yr	93	46.5	174,907	12.6
25 - 44 yr	61	30.5	375,107	27.0
45 - 64 yr	38	19.0	358,953	25.9
Other/Rather not to answer	8	4		
Total	**200**	**100**		
Household composition				
Single without children	74	37.0	247,011	39.2
Couple without children	54	27.0	173,441	27.5
Presence of children	33	16.5	252,598	33.3
Other/Rather not to answer	39	19.5		
Total	**200**	**100**		
Educational level				
Low educated	43	21.5	381,000	48.0
Bachelor's degree	91	45.5	415,000	52.0
Master/PhD	59	29.5		
Other/Rather not to answer	7	3.5		
Total	**200**	**100**		
Employment status				
Student	95	47.5		20.9
Employment (parttime & fulltime)	83	41.5		53.0
Other/Rather not to answer	22	11.0		
Total	**200**	**100**		
Individual income				
Lower than €2,310	107	53.5		
€2,310 - €3,975	30	15.0		
Higher than €3,975	46	23.0		
Other/Rather not to answer	17	8.5		
Total	**200**	**100**		
Mobility disability				
No	185	92.5		88
Yes	7	3.5		12
Other/Rather not to answer	8	4.0		
Total	**200**	**100.0**		

Variables	Survey respondents		Population Utrecht	
	Count	Percentage (%)	Count	Percentage(%)
Mobility availability				
Driving license	156	26.5	908,679	61.0
PT subscription	127	21.6		
Car	79	13.6	752,271	54.0
Bike	168	28.6		1.18
E-bike | Scooter | Moped	45	7.7		0.18
Other/Rather not to answer	7	1.2		
Total	**588**	**100.0**		
Commuting reimbursement				
No reimbursement	79	39.5		
PT subscription	32	16		
Mileage reimbursement	46	23		
Leased car	10	4.9		
Other/Rather not to answer	39	19.5		
Total	**200**	**100.0**		
Commuting frequent mode				
PT	71	35.5		
Car	38	19.0		
Bike	69	34.5		
Other/Rather not to answer	22	11.0		
Total	**200**	**100.0**		

Table 6.2: Chi-square statistical test

Variable	Chi square	DF	Critical Value	P value
Age group	213.300	3	7.815	5.64E-11
Household composition	17.216	2	5.991	0.00018
Educational levels	49.607	2	5.991	1.69E-11
Employment status	72.700	1	3.841	1.51E-17
Mobilily disability	12.500	1	3.841	0.00193

The age distribution in the collected dataset is notably skewed towards younger individuals, particularly those aged 15-24, who make up 46.5% of the sample compared to just 12.6% in the general population of Utrecht. The Chi-square test confirms a highly significant difference, indicating that the model might be biased towards the commuting preferences typical of younger demographics, potentially limiting its applicability to older populations.

The dataset also shows an overrepresentation of students, who constitute 47.5% of respondents compared to 20.9% in the general population. This discrepancy, confirmed by a highly significant Chi-square test, indicates that the commuting preferences captured may lean heavily towards those of students, such as a higher reliance on public transport or non-motorised modes, which might not be as prevalent in a more employment-diverse population.

Educational level is another area where the dataset diverges from the population, with a higher percentage of highly educated respondents (Bachelor's degree and higher) than is typical in Utrecht. This difference is largely attributed to the survey's focus on students and professionals in the Randstad area. The significant Chi-square result reflects this disparity, suggesting that the model may be biased towards the preferences of highly educated individuals, potentially overlooking the commuting behaviours of those with lower educational levels.

In contrast, the household composition in the dataset closely mirrors that of the general population, with similar proportions of singles, couples without children, and households with children. Despite the Chi-square test indicating a significant difference, the alignment is much closer than for other factors. This suggests that household composition is relatively well-represented, providing a strong foundation for predicting commuting behaviours related to different household types.

Finally, the representation of individuals with mobility disabilities is lower in the dataset (3.5%) compared to the general population (12%). The significant Chi-square result highlights this discrepancy, which could limit the model's effectiveness in predicting the transportation needs and mode choices of individuals with mobility disabilities, a group that often requires specific considerations in transportation planning.

The Chi-square tests reveal significant discrepancies between the dataset and the population of the province for all the variables tested, with a notable overrepresentation of highly educated young students. This discrepancy is largely due to the survey's distribution primarily within universities in the Randstad region. However, this discrepancy may not be very problematic, as Section 3 highlights the expected growth in the young adult population in the future. While this does not mean the dataset is fully representative of the current population, it may still be valuable for predicting future trends. Also, it may introduce biases in the model's predictions when applied to the current general population. Therefore, these differences need to be carefully considered in the interpretation of the results to ensure accurate and contextually relevant conclusions.

6.1.2. Commuting mode choice

Looking at the total collected mode preferences, it can be observed that the share of each mode is quite equal to each other, 30.3% for PT, 33.3% for car and 36.4% for bike option. This indicates that the SP experiment was developed properly; the attributes and levels did not give an advantage to a specific mode. When looking at the mode preferences of each socio-demographic group specifically, differences in mode choice can be observed. To evaluate the differences in mode preferences of different socio-demographic groups quantitatively, the Chi-square test statistic is performed, and the results are presented in table 6.3.

Table 6.3: Chi-square test results for mode preferences

Variables	Chi-square	DF	Critical Value	P-value
Age group	33.54	4	9.488	9.22E−7
Household composition	16.88	4	9.488	0.002
Education level	3.09	4	9.488	0.545
Income	18.97	4	9.488	0.0008
Employment status	13.49	2	5.991	0.0012
Mobility availability	26.22	10	18.31	0.0035

Age group
Among different age categories, younger individuals (15-24 years) tend to prefer cars (40%), with public transport (PT) and biking being less popular choices (29% and 31% respectively). As age increases, the preference for biking grows, peaking at 46% for the 45-64 age group, while car and PT usage have decreased. The oldest group shows a significant shift towards biking (46%) compared to younger groups, with car and PT usage dropping to 27% each, see figure 6.1. The Chi-square test confirms that the differences in mode choice across age groups are statistically significant. The shift towards biking with increasing age may reflect lifestyle changes, such as greater health consciousness or the preference for more leisurely modes of transport among older individuals. The expected convenience of car ownership might influence the high preference for cars among younger people at a younger age.

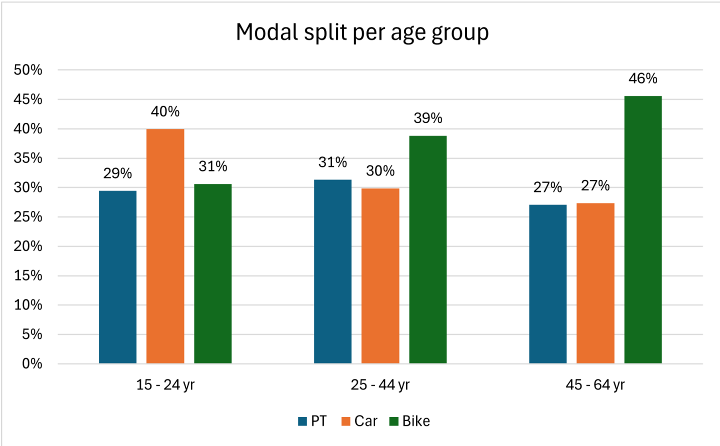


Figure 6.1: Mode preferences among age groups

Household Composition
Regarding mode preference among different household types, single individuals and couples show similar patterns in mode choice, with a relatively even split among PT, car, and bike usage (figure 6.2). Households with children show a strong preference for biking (45%), with car usage being lower (30%) and PT being the least preferred (25%). This high preference for biking among households with children could have a direct link with the mode preference of older people, as shown in the previous paragraph. This difference in mode preferences between different household types is also confirmed by the Chi-square test results, as the high Chi-square value is caused by the large difference of bike's preference in the households with children.

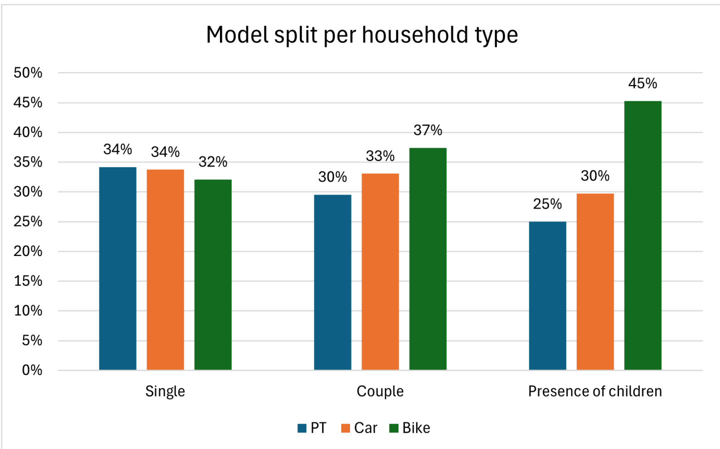


Figure 6.2: Mode preferences among households

Educational level

Across all education levels, it can be observed from 6.3 that mode choices are relatively balanced. The lowest education level shows a slightly lower preference for PT (27%), while car and bike are nearly equal (36% and 37%). Individuals with a master's degree show a slight preference for biking (38%) compared to car (32%) and PT (30%). The Chi-square test reveals that educational level does not significantly influence mode choice in this dataset. The balanced mode preferences across educational levels suggest that factors other than education may be more influential in determining mode choice. This could also imply that preference for different transportation modes is relatively equitable across the education levels of the survey's respondents.

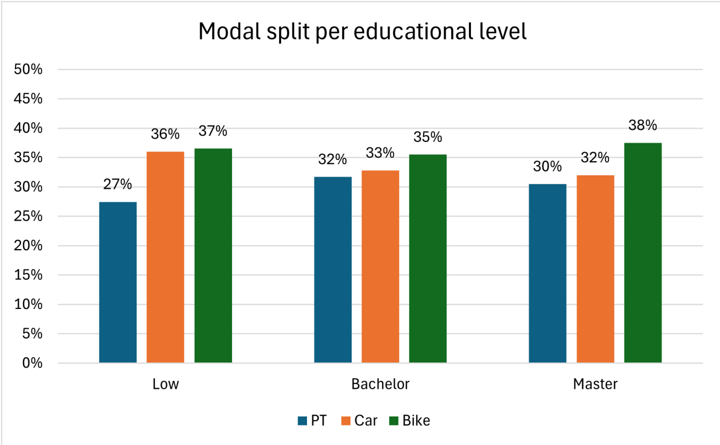


Figure 6.3: Mode preferences for different educational level

Income

Lower-income individuals show a fairly even split among the three modes, with car and bike usage at 32% each and PT slightly lower (36%), see figure 6.4. Middle-income individuals also show a balanced mode choice pattern, while higher-income individuals are likely to have a strong preference for biking (48%). The Chi-square test shows that income significantly impacts mode choice due to the strong preference for biking among higher-income individuals. This strong preference might be created by a greater awareness of health and environmental benefits and (negative) experiences of car usage. The more balanced mode choices among lower-income groups may reflect a need to choose the most cost-effective transportation option, which can vary based on accessibility and necessity.

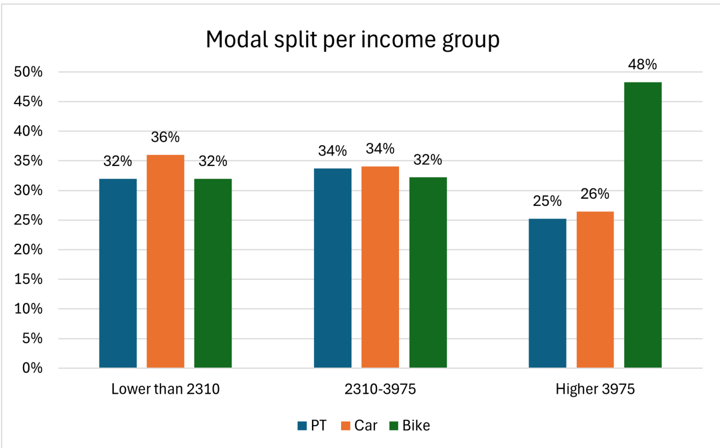


Figure 6.4: Mode preferences for different income groups

Employment

From the model split in figure 6.5, the relatively even distribution across all three modes under students can be seen. It indicates that students are flexible in their mode choices. For employed individuals, biking is the most preferred mode (41%), followed by cars and PT, both at 29%. The significant preference for biking among employed individuals could be attributed to the availability of better biking infrastructure, health considerations, or environmental awareness. The high Chi-square and P-value of the statistical test indicated significant differences between students and employed individuals, suggesting that these factors are likely influencing commuting mode choices in a significant way.

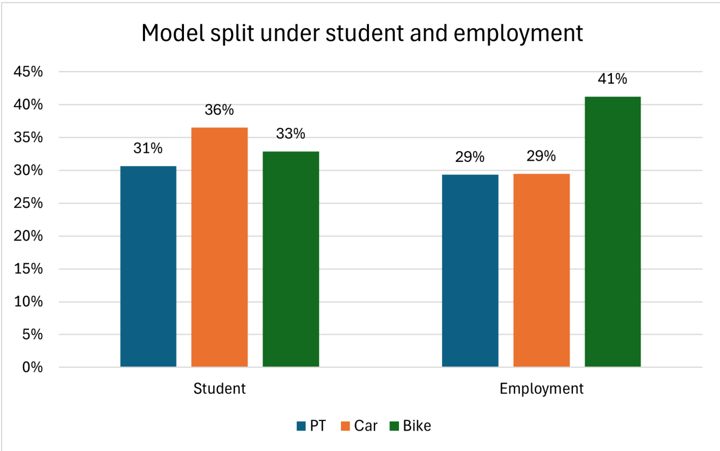


Figure 6.5: Mode preferences for different employment status

Mobility availability

The mode choice is also analysed for different mobility disability and availability groups with the modal split presented in figure 6.6. Obviously, the differences between the modes for other mobility availability groups are slightly small. It seems that the availability of a mode does not affect choosing other modes. Only the PT is less chosen compared to the car and the bike among the group with driving license. The car option has been chosen more than other modes, but the difference with the bike is also minimal. The significant Chi-square result suggests that there are significant differences in the distribution of mode choices across the different categories regarding the availability of personal mobility. This implies that these factors are likely influencing commuting mode choices in a significant way.

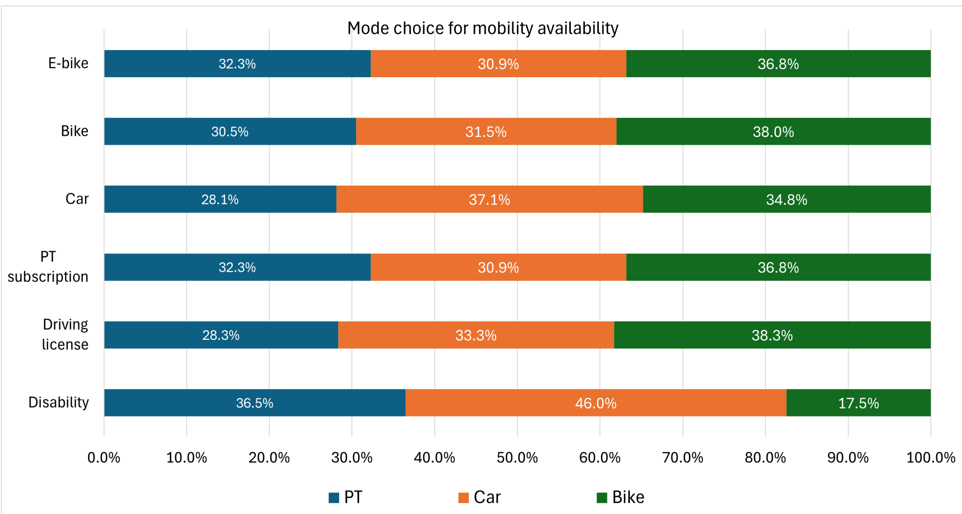


Figure 6.6: Modal preferences for different mobility disability and availability groups

6.1.3. Conclusion on descriptive statistics

This chapter has provided a comprehensive analysis of the dataset collected through the survey, highlighting significant socio-demographic factors that influence commuting mode choices. The data preparation steps, including the aggregation of categories with low response rates, have ensured the robustness of the dataset, allowing for more reliable statistical analysis. The Chi-square tests confirmed notable discrepancies between the sample and the general population, particularly in terms of age, education level, and employment status, reflecting the specific demographic characteristics of the respondents, most of whom are younger and highly educated.

The analysis of mode preferences among the survey respondents reveals an equal distribution across all three modes when considering the entire dataset. However, the Chi-square test results indicate statistically significant differences in commuting mode choices across various socio-demographic groups, with the exception of educational levels. Younger individuals and students show a balanced preference across cars, public transport, and biking, reflecting their need for flexibility and affordability. In contrast, older individuals and employed persons favour biking, likely due to health benefits and supportive infrastructure. Households with children also prefer biking, driven by practical and health considerations. Income further differentiates choices, with higher-income individuals leaning towards biking, possibly due to greater health and environmental awareness. Educational level, however, does not significantly impact mode choice, suggesting that other factors like income or location are more influential. Mobility availability strongly dictates preferences, emphasising the importance of access to resources like driving licenses and car ownership.

The significant differences identified in commuting behaviours across these groups will guide the model's structure, ensuring that it accurately reflects the preferences and behaviours of different demographic segments. Moreover, the overrepresentation of younger, highly educated individuals in the dataset suggests that the model may initially be more reflective of these groups' preferences. However, this focus could be beneficial given the expected growth in the young adult population in the future. For further applications, it will be essential to account for the potential biases to enhance the model's applicability to a broader population, particularly in peripheral areas where the demographic composition may differ from that of the surveyed sample.

6.2. Discrete Choice Model

The second part of the data analysis is constructing a discrete choice model to predict the commuting mode choice behaviour in peripheral areas. As described in section 5.2.3, Python library Biogeme is used for the specification of model parameter $\beta's$. Applying the estimation strategy in section 5.2.3, MNL base model, MNL final model and Panel ML model have been constructed sequentially and the estimation results are given in section 6.2.1, 6.2.2, 6.2.3. After estimating, the models have been assessed by the performance indicators in section 6.2.4 and validated in section 6.2.5.

6.2.1. MNL base model

The MNL base model is the most basic model with only alternative property attributes (Z_{inm}). Based on the design variables from section 4.3, the attributes Z_{imn} and the model parameters $\beta's$ inputs for the model estimation are identified in the table 6.4. Because of the differences in the mode properties, all model parameters $\beta's$ are alternative-specific. This means each parameter and the corresponding attribute only apply to a particular transportation mode. This table also provides the description and the range of the input values for each parameter.

Table 6.4: Overview of alternatives properties attributes

Attibutes Z_{imn}	Parameter β_m	Description	Range
PT$_{access}$	β_{PT_access}	PT access time in minutes	[0, 10]
Waiting	$\beta_{waiting}$	PT waiting time in minutes	[0, 7.5]
PT$_{IVtime}$	β_{PT_IVtime}	PT in-vehicle time in minutes	[0, 41]
Transfer	$\beta_{transfer}$	Categorical (effect coded): transfer, without transfer	{0, 1}
Transfer$_{time}$	$\beta_{transfer\_time}$	PT transfer time in minutes	-
Transfer$_{distance}$	$\beta_{transfer\_distance}$	PT transfer distance in metres	[0, 400]
PT$_{type}$	β_{PT_type}	Type of PT vehicle: tram, bus	{0, 1}
ASC_{car}	-	Alternative specific constant for car	-
Car$_{access}$	β_{car_access}	Car access time in minutes	[0, 15]
Car$_{Ivtime}$	β_{car_IVtime}	Car in-vehicle time in minutes	[13, 43]
Car_$parking$	$\beta_{parking}$	Monthly parking cost in euro	[10, 60]
ASC$_{bike}$	-	Alternative specific constant for bike	-
Bike$_{Ivtime}$	β_{bike_IVtime}	Bike in-vehicle time in minutes	[38, 50]

Applying the forward stepwise estimation strategy from section 5.2.3, alternative specific constants ASC's are defined as the initial model. After that, each attribute and the corresponding β are added iteratively to the initial model to define the best base model. The estimated model with the highest model performance is considered the MNL base model. The model performance is determined based on the model performance indicators from the estimation report, this will be discussed further in section 6.2.4.

The estimation results of the base MNL model are obtained and shown in the table 6.5. The complete

Python code for the model estimation and output file of the estimation results of the base model can be found in Appendix E.

Table 6.5: Estimation result of the base model

Alternative specific	β	Value	Rob. Std err	Rob. t-test	Rob. p-value
PT	$\beta_{transfer}$	-0.764	1.27E-01	-6.03	1.64E-09
	β_{PT_access}	-0.133	3.17E-02	-4.18	2.87E-05
	β_{PT_IVtime}	-0.0679	1.26E-02	-5.39	7.04E-08
Car	ASC_{Car}	-1.41	5.38E-01	-2.63	8.57E-03
	β_{car_access}	-0.105	1.09E-02	-9.61	0
	β_{car_IVtime}	-0.0389	4.22E-03	-9.22	0
	$\beta_{car\_parking}$	-0.0166	2.49E-03	-6.66	2.65E-11
Bike	ASC_{Bike}	2.01	6.61E-01	3.05	2.32E-03
	β_{bike_IVtime}	-0.128	1.03E-02	-12.40	0

From the obtained results for the base model above, it can be seen that the total model parameters are reduced from thirteen to nine parameters. For the PT alternative, the attributes waiting time, type of vehicle, transfer time and transfer distance are not included in the base model. These parameters are eliminated from the model because they did not contribute to the explanatory power of the model and/or they are also not statistically significant. In the base model, all of the model parameters have a robust probability value (p-value) lower than 5%, meaning that the effect observed in the dataset is considered statistically significant.

Concerning the estimated values of the model parameters, it can be observed that the alternative specific constant for the bike alternative has the largest positive value; this means that preference for the bike is expected to have the largest (positive) impact on the decision-making process of the traveller. This is also the only positive value in the base model. On the contrary, according to this result, the negative alternative-specific constant for car use suggests travellers are less likely to choose the car as their mode of transport compared to other alternatives in the given choice condition.

For the PT alternative, the presence of transfer during the trip has the largest negative value, indicating that people seem not to appreciate the inconvenience of transferring during their journey. However, this large value does not directly mean the transfer has the largest negative impact on the decision-making process, because of the linear relation with the binary value of the transfer attribute. The PT in-vehicle time has quite a small value, so the PT in-vehicle time does not affect too much in this model.

For the car alternative, the access time has the largest negative value (after the ASC for the car) compared to the in-vehicle time or the parking cost. So, people do not appreciate long distances to access their car. This means the location of the parking facilities is expected to have the largest impact on the choice of car use. However, the access time to the PT stops attribute still has a higher value, which means that for the same access distance, the utility for PT is lower than that for cars. Remarkably, the parking cost attribute has the smallest value, so the impact of the parking cost is small in the utility for car alternative. For the bike alternative, the parameter for in-vehicle time has a larger (negative) value than for PT or car. Longer cycling time leads to a higher decrease in utility for bike alternatives.

Overall, besides the preferences of the alternatives (ASCs), the parameter for PT transfer is ecpected to have the largest impact on the mode choice behaviour of travellers. The smallest parameter concerns the parking cost attribute for car alternative. And the PT access time should be smaller than the access time to the car to enhance the higher utility for PT usage.

6.2.2. MNL final model

Once the MNL base model is estimated, the MNL final model can be constructed by including the interaction effects of the personal characteristics attributes (Z_{nk}) into the base model. Table 6.6 presents an overview of all attributes for personal characteristics and model parameters those are used in the estimation of the MNL final model. Because of the categorical characteristics of these attributes, dummy variables are used for the model estimation.

Table 6.6: Overview of personal characteristics attributes

Attribute	Dummy variable Z_{nk}	Parameter β_k	Dummy codes
Age group			
15 - 24 yr			0
25 - 44 yr	Age_{25_44}	β_{25_44}	1
45 - 64 yr	Age_{45_64}	β_{45_64}	1
Household composition			
Single without children	Singel	β_{single}	1
Couple without children	Couple	β_{couple}	1
Presence of children			0
Education level			
Low educated			0
Bachelor's degree	$Edu_{bachelor}$	$\beta_{edu\_bachelor}$	1
Master/PhD	Edu_{high}	β_{edu_high}	1
Employment status			
Student			0
Employment	Employment	$\beta_{employment}$	1
Individual income			
Lower than €2,310			0
€2,310 - €3,975	$Income_{mid}$	β_{income_mid}	1
Higher than €3,975	$Income_{high}$	β_{income_high}	1
Mobility disability			
No			0
Yes	Disability	$\beta_{disability}$	1
Mobility availability			
Driving license	$Driving_{l}icense$	$\beta_{driving_{l}icense}$	1
PT subscription	$Availability_{PT}$	$\beta_{Availability\_PT}$	1
Car	$Availability_{car}$	$\beta_{Availability\_car}$	1
Bike	$Availability_{bike}$	$\beta_{Availability\_bike}$	1
E-bike | Scooter | Moped			0
Commuting reimbursement			
No reimbursement			0
PT subsciption	$Reimbursement_{PT}$	$\beta_{Reimbursement\_PT}$	1
Mileage reimbursement	$Reimbursement_{mileage}$	$\beta_{Reimbursement\_mileage}$	1
Leased car	$Reimbursement_{leased\_car}$	$\beta_{Reimbursement\_leased\_car}$	1
Commuting frequent mode			
PT			0
Car	$Frequent_{car}$	$\beta_{Frequent\_car}$	1
Bike	$Frequent_{bike}$	$\beta_{Frequent\_bike}$	1

Using the same forward stepwise estimation strategy, each input dummy variable is first generically added to the utility function of all alternatives. This means the model parameter for each attribute is the same for all alternatives. However, the explanatory power did not increase for any added attributes. Therefore, it is chosen to have the attribute specific for each alternative, which indicates that for each alternative, the model parameter β can be different. Fortunately, the explanatory power of the model has increased after this differentiation, see later in section 6.2.4. Table 6.7 shows the final estimation results for the model with the highest explanatory power; this is also the final MNL model.

Table 6.7: Estimation results final with interaction

Alt. specific	β	Value	Rob. Std err	Rob. t-test	Rob. p-value
PT	β_{PT_access}	-0.14	3.20E-02	-4.37	1.24E-05
	β_{PT_IVtime}	-0.0675	1.29E-02	-5.25	1.50E-07
	$\beta_{transfer}$	-0.803	1.30E-01	-6.19	6.16E-10
	$\beta_{PT\_Availability\_bike}$	-0.626	1.19E-01	-5.26	1.42E-07
	$\beta_{PT\_Mid\_income}$	0.299	1.07E-01	2.81	5.02E-03
	$\beta_{PT\_High\_income}$	-0.299	1.07E-01	-2.81	5.01E-03
	$\beta_{PT\_Reimbursement\_PT}$	0.364	1.66E-01	2.19	2.84E-02
	$\beta_{PT\_Reimbursement\_leased}$	-0.364	1.66E-01	-2.19	2.84E-02
Car	ASC_{Car}	-1.35	5.51E-01	-2.45	1.43E-02
	β_{car_access}	-0.11	1.13E-02	-9.75	0
	β_{car_IVtime}	-0.0412	4.31E-03	-9.58	0
	$\beta_{car\_parking}$	-0.0174	2.54E-03	-6.85	7.40E-12
	$\beta_{Car\_Age\_25\_44}$	0.000294	6.91E-05	-4.26	2.08E-05
	$\beta_{Car\_Availability\_bike}$	-0.626	1.19E-01	-5.26	1.43E-07
	$\beta_{Car\_Disability}$	0.000174	7.04E-05	-2.48	1.32E-02
	$\beta_{Car\_Driving\_license}$	0.000195	6.91E-05	2.82	4.78E-03
Bike	ASC_{Bike}	2.19	6.80E-01	3.22	1.26E-03
	β_{bike_IVtime}	-0.134	1.07E-02	-12.60	0
	$\beta_{Bike\_Availability\_PT}$	-0.419	1.22E-01	-3.43	5.93E-04
	$\beta_{Bike\_Availability\_car}$	-0.207	7.73E-02	-2.68	7.43E-03

From the obtained results, it can be observed that, in general, parameters of personal characteristics are significantly lower than the alternative properties in absolute terms. The values of the alternative properties parameters (β_m) are higher than in the base model, but the difference is minimal.

Obviously, the income and the commuting reimbursement only affect the choice for PT. Remarkably, the values for high-income and middle-income parameters are exactly the same but with different signs. If the decision-maker has a high income, the utility for the PT alternative will decrease; when they belong to the middle-income group, the utility for PT will increase at the same rate. In addition, when the decision-maker belongs to the low-income category, which is not included in the final model, the income attribute will no longer contribute to the utility of the PT alternative. This phenomenon also occurs similarly for commuting reimbursement attributes, with utility increasing for PT subscriptions and decreasing when they are provided with a leased car. When mileage reimbursement is applied, reimbursement aspects no longer contribute to the mode choice behaviour of the commuter.

When considering the car alternative, it can be seen that age, mobility disability, and the availability of a driving license only affect the utility of the car option. They all positively affect the utility of the car alternative, this means commuters who belong to these categories are more likely to opt for car use. However, the model parameters regarding these attributes are significantly smaller in comparison to other parameters, so they do affect, but the impact is minimal. The parameter for the availability of bikes becomes the largest parameter in absolute terms, so having a bike will directly lead to a decrease in the car's utility.

For the bike alternative, the availability of a PT and a car will have an aversion effect to the mode utility. These parameters are around twice as large as the bike in-vehicle time. That means that when having a PT subscription or a car, the decision-maker will not choose the bike even if the cycling time becomes shorter.

Conversely, the availability of bikes reduces the utility of both public transportation (PT) and car options. The alternative-specific constants remain the most significant factors. And among the parameters for alternative properties, the PT transfer parameter still has the highest value. Access time to the vehicle continues to have a greater impact on the utility of PT compared to the utility of the car. Additionally, parking costs have the least effect on overall choice behaviour.

In conclusion, the final MNL model, which consists of attributes and model parameters regarding the alternative properties and personal characteristics, has been developed for modelling the commuting mode choice behaviour. The utility functions of the final model are specified as follows:

$$\begin{aligned} V_{PT} = {} & \beta_{PT\_access} \cdot PT_{access} + \beta_{PT\_IVtime} \cdot PT_{IVtime} + \beta_{transfer} \cdot Transfer \\ & + \beta_{PT\_Availability\_bike} \cdot Availability_{bike} + \beta_{PT\_Mid\_income} \cdot Income_{Mid} \\ & + \beta_{PT\_High\_income} \cdot Income_{High} + \beta_{PT\_Reimbursement\_PT} \cdot Reimbursement_{PT} \\ & + \beta_{PT\_Reimbursement\_leased} \cdot Reimbursement_{leased} \end{aligned} \tag{6.1}$$

$$\begin{aligned} V_{Car} = {} & ACS_{car} + \beta_{car\_access} \cdot Car_{access} + \beta_{car\_IVtime} \cdot Car_{IVtime} \\ & + \beta_{parking} \cdot Car_{Parking} + \beta_{car\_Age\_25\_44} \cdot Age_{25\_44} \\ & + \beta_{car\_Availability\_bike} \cdot Availability_{Bike} + \beta_{car\_disability} \cdot Disability \\ & + \beta_{car\_driving\_license} \cdot Driving\_license \end{aligned} \tag{6.2}$$

$$\begin{aligned} V_{Bike} = {} & ASC_{bike} + \beta_{bike\_IVTime} \cdot Bike_{IVtime} \\ & + \beta_{bike\_Availability\_PT} \cdot Availability_{PT} + \beta_{bike\_Availability\_car} \cdot Availability_{Car} \end{aligned} \tag{6.3}$$

6.2.3. Panel Mixed Logit model

To capture the panel effect in the choice behaviour of the decision-maker, the random parameter β' is added to the constructed final MNL model. Following the estimation strategy in section 5.2.3, the random parameter β' is included in the utility of the mode alternatives iteratively. As described in section 5.2.3, β' is a random normal distributed parameter with a mean of zero and a standard deviation σ_{panel}, this σ_{panel} is the new estimation parameter and included in the estimated results output.

Due to the introduction of the random parameter β', seven model parameters related to personal characteristics have become insignificant and resulting in parameter eliminations from the utilities. The parameters for incomes (both middle and high) and reimbursement of PT subscriptions are subtracted from the utility of PT. The age and driving license parameters are removed from the utility of the car alternative, and the availability of PT and car are also removed from the bike's utility.

The decrease in significance of these parameters may caused by the effective capturing of individual-specific variations of the random parameter in the Panel ML model. If the panel parameter effectively captures the variability or unobserved factors that were previously partially explained by observable attributes, these observable attributes may become statistically insignificant in the Panel ML model. Essentially, the model "shifts" the explanatory power from these observable attributes to the random parameter, which represents a broader, more flexible form of heterogeneity.

The utility functions of the final Panel ML model can be specified in equations 6.4, 6.5 and 6.6. In this model, the random parameter β' is added only to the utility of bike alternative. This inclusion provides a higher explanatory power to the model than also including the β' in the utility of PT and/or car.

$$\begin{aligned} V_{PT} = {} & \beta_{PT\_access} \cdot PT_{access} + \beta_{PT\_IVtime} \cdot PT_{IVtime} \\ & + \beta_{transfer} \cdot Transfer + \beta_{PT\_Availability\_bike} \cdot Availability_{bike} \\ & + \beta_{PT\_Reimbursement\_leased} \cdot Reimbursement_{leased} \end{aligned} \tag{6.4}$$

$$\begin{aligned} V_{Car} = {} & ACS_{car} + \beta_{car\_access} \cdot Car_{access} + \beta_{car\_IVtime} \cdot Car_{IVtime} \\ & + \beta_{parking} \cdot Car_{Parking} + \beta_{car\_Availability\_bike} \cdot Availability_{Bike} \end{aligned} \tag{6.5}$$

$$V_{Bike} = ASC_{bike} + \beta_{bike\_IVTime} \cdot Bike_{IVtime} + \beta' \tag{6.6}$$

The final estimation results of the panel ML model with the best fit is shown in table 6.8. From the estimated result, it can be observed that there are significant changes in the values of the parameters. All parameters related to the personal characteristics have become significant larger than in the final MNL model and also larger than the parameters of alternative property. The ASC of for car and bike are also become larger, indicating stronger preference for bike and rejection for car among the respondents in the dataset in the given SP context.

Besides, the standard deviation for panel parameter (σ_{panel}) has an estimated value of -2.24, quite a large value in comparison to the rest of the parameters. This value of 2.24 indicates a substantial level of variability in these unobserved factors across the individuals in the dataset. This means there is considerable heterogeneity in how individuals' mode choices are influenced by factors not directly measured by the model's observed variables.

Table 6.8: Estimation results of the Panel Mixed Logit model

Alternative	β	**Value**	**Rob. Std err**	**Rob. t-test**	**Rob. p-value**
PT	β_{PT_access}	-0.168	0.023	-7.38	1.56E-13
	β_{PT_IVtime}	-0.0937	0.016	-5.76	8.20E-09
	$\beta_{transfer}$	-0.891	0.150	-5.92	3.15E-09
	$\beta_{PT\_Availability\_bike}$	-1.07	0.396	-2.71	6.65E-03
	$\beta_{PT\_Reimbursement\_leased}$	-1.04	0.533	-1.95	0.051
Car	ASC_{Car}	-2.10	0.641	-3.27	1.07E-03
	β_{car_access}	-0.112	0.010	-10.7	0
	β_{car_IVtime}	-0.0443	0.006	-7.80	6.22E-15
	$\beta_{car\_parking}$	-0.02	0.003	-7.32	2.47E-13
	$\beta_{Car\_Availability\_bike}$	-1.57	0.381	-4.13	3.55E-05
Bike	ASC_{Bike}	2.88	1.050	2.74	6.10E-03
	β_{bike_IVtime}	-0.20	0.021	-9.51	0
	σ_{panel}	-2.24	0.207	-10.8	0

6.2.4. Model fit

As described in section 5.2.3, model performance indicators from the estimation results have been used for the assessment of the goodness of fit for the model estimation. Table 6.9 shows the final performance indicators for the base MNL, final MNL and Panel ML model.

From the obtained results, it can be observed that the base MNL model has the worst explanatory power, as evidenced by the lowest likelihood ratio, rho-square, and rho-square-bar values. This is logical since it does not include the personal characteristics attributes as the final and Panel ML model. The Panel ML model clearly has the best explanatory power because it takes the randomness of repeated choice decisions for the same individual across different choice situations of panel data, while all observations are independent in the MNL models. The Panel ML model also contains fewer model parameters than the final MNL model, reflecting higher effectiveness in capturing individual-level heterogeneity of the panel parameter.

However, both AIC and BIC are lower for the Final MNL model, suggesting that it is more selective when balancing model fit and complexity. The Panel MNL is, therefore, more complex and potentially overfitted. Also, a higher number of model attributes of the final MNL can contribute to better interpretations of model attributes. Consequently, the base MNL is clearly the worst of the three constructed models and should not be used for the application. The final decision between the final MNL and Panel ML can not be made only based on the goodness of fit of these models due to the trade-off between the explanatory power and the generalisation of the models.

Table 6.9: Performance indicators for base and final MNL models

Performance indicators	**Base MNL model**	**Final MNL model**	**Panel ML model**
Number of estimated parameters	9	20	13
Sample size	1440	1440	200
Observations	-	-	1800
Likelihood ratio test	480.708	571.6227	1073.731
Rho-square	0.152	0.181	0.271
Rho-square-bar	0.146	0.168	0.265
Akaike Information Criterion	2701.295	2632.381	2907.273
Bayesian Information Criterion	2748.747	2737.829	2950.151

6.2.5. Model validation

Besides the model fit, the quality of the model can also be assessed by model validation. The final MNL model and the Panel ML model have been validated using a validation dataset. This dataset contains the remaining 20% of the dataset that is not used for the estimating of the MNL models. For consistency, this validating dataset has also been used to validate the Panel ML. Using the model validation method described in section 5.2.4, a simulation of 20 runs is performed to calculate the modal splits for both models using the attribute values from the validation dataset. Figure 6.7 shows calculated modal splits of the actual choices and the predicted choices of the final MNL and Panel ML model. The explicit differences between the modal splits can be found in table 6.10.

From the calculated modal splits, it can be observed that the final MNL model has better accuracy than the Panel ML model, and the share of all three modes is closer to the actual choices. However, both models still have pretty good quality. The share of PT has been worst predicted but with small differences, 2.75% and 5.53%. On the other hand, the share of cars is best predicted with 0.86% and 2.06%. Based on the validation results, it can be concluded that the final MNL model has a higher accuracy than the Panel ML model, indicating a better predictive power of the final MNL.

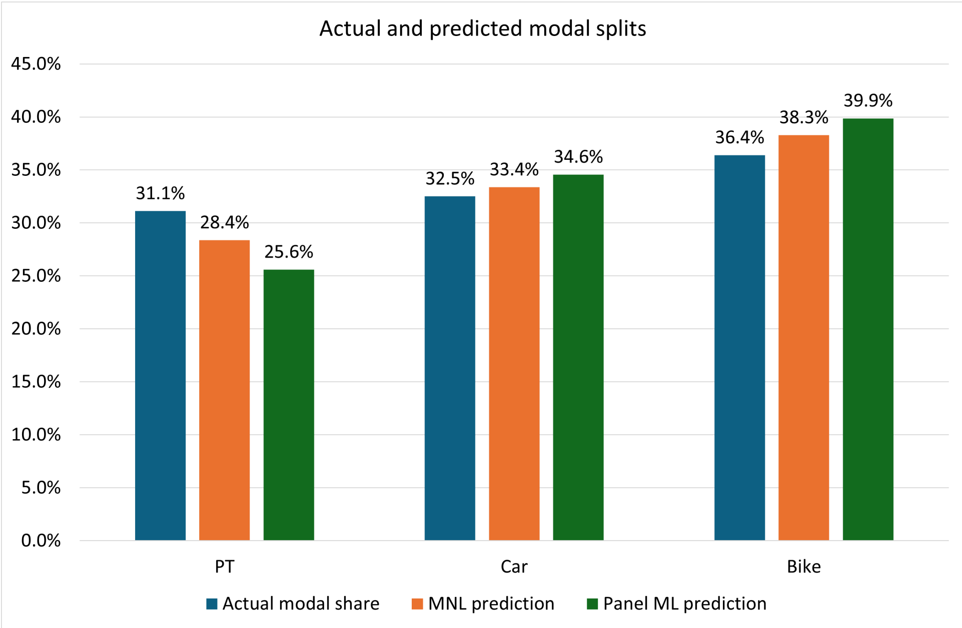


Figure 6.7: Modal split of model validation

Table 6.10: Differences between actual modal share and predictions

Alternative	Actual modal share	MNL difference	Panel ML difference
PT	31.1%	-2.75%	-5.53%
Car	32.5%	0.86%	2.06%
Bike	36.4%	1.89%	3.47%

6.3. Conclusion on Data Analysis Results

In this chapter, a comprehensive analysis of commuting mode choice behaviour was conducted using both descriptive statistics and discrete choice modelling. Data preparation steps, including category aggregation, ensured the dataset's robustness. Chi-square tests revealed significant discrepancies between the sample and the broader population of Utrecht, particularly in age, education, and employment, with respondents being younger and more educated. While mode preferences were evenly distributed across cars, public transport, and biking, significant differences emerged across socio-demographic groups, except for educational levels. These findings emphasise the importance of considering the sample's demographic characteristics when interpreting the results and the need to account for potential biases, particularly due to the overrepresentation of younger, educated individuals.

After that, the dataset is used for the conctructions of three discrete choice models, Base MNL, Final MNL, and Panel Mixed Logit (Panel ML). The Base MNL model, which only includes alternative-specific attributes, serves as a foundational model, but its explanatory power is limited. The Final MNL model improved upon this by incorporating personal characteristics, leading to better model fit and more nuanced insights into the factors influencing commuting choices. The Panel ML model further enhanced explanatory power by accounting for individual heterogeneity through random parameters. However, while the Panel ML model shows the highest fit to the training data, as indicated by larger likelihood ratio and adjusted-rho-square values, it also has higher AIC and BIC values, suggesting greater complexity and a risk of overfitting.

Validation results indicated that, despite its higher explanatory power, the Panel ML model was outperformed by the Final MNL model in terms of predictive accuracy. This suggests that the Final MNL model, with its balance of explanatory power and model simplicity, is better suited for practical applications in predicting commuting mode choices. Ultimately, while the Panel ML model provides a deeper understanding of individual behaviours, the Final MNL model offers a more reliable and interpretable tool for forecasting mode choice in peripheral areas.

7

Model Application for Rijnenburg

To assess the extent of the impact of the proposed mobility strategies from section 4.2 on the mode choice behaviour of the commuters in Rijnenburg, the constructed discrete choice model is applied for Rijnenburg in this chapter. As concluded in chapter 6, the MNL final model has been chosen for the model application. First, the population and scenarios set-up is described in section 7.1, after that the modal spits as results of the application is presented and evaluated in section 7.2. After that, sensitivity analyses of the relevant design variables are shown in section 7.3.

7.1. Application Set Up

For the application of the model to the case study are Rijnenburg, a simulation is performed as described in section 5.2.4 to determine the modal split for different design scenarios. In this section, the first two steps of the application concerning the population set-up and defining the values of the model attributes are described.

7.1.1. Population set up

The population set-up contains the defining of the size and the share of characteristics for the sample population. The population size for the sample enumeration is assumed to be equal to the total of commuters in Rijnenburg. In section 4.1, the plan is to have 25,000 dwellings in Rijnenburg. With an average of 2.1 residents per house, the expected total population of Rijnenburg is 52.500 people. According to Centraal Bureau Voor de Statistiek (2024), the labour participation of Utrecht's province in the first quarter of 2024 is 75.4%; it follows the sample size of 39.585 for the application.

The availability of PT subscription, bike and reimbursements are assumed as 100%. Assuming that everyone is able to own a bike and a PT prescription. And commuters are free to choose the reimbursement form that suits their need. The shares of the middle- and high-income groups are determined as 35% and 25% based on the expected housing distribution, see section 4.1.3. The remaining attributes are assumed to be the same as the population from the province Utrecht; see section 6.1. An overview of the characteristics of the applied population is given in the table 7.1.

Table 7.1: Overview of characteristics of population

Attributes	Share in population
Age 25 - 44	27%
Middle income	35%
High income	25%
Mobility disability	12%
Driving license	61%
Availability PT subscription	100%
Availability car	54%
Availability bike	100%
Reimbursement PT subscription	100%
Reimbursement leased car	100%

7.1.2. Scenarios attributes

The model attributes for the application are retrieved directly from proposed design scenarios in section 4.2 and the corresponding levels from section 5.1.2. Besides the established design variables, the effect of the congestion is taken into account by different in-vehicle times (43 and 13 minutes) for car alternative. The overview of three design scenarios (Conventional, Sustainable and Ambitious) with final model attributes and input values are shown in table 7.2.

Table 7.2: Overview of input attributes input value per design scenario

	Design variable	Attribute	Conventional	Sustainable	Ambitious
PT	Stop density	PT Access	10	7	5
	Route planning	PT In-vehicle time	41	31	31
	Route planning	PT Transfer	1	1	0
Car	Parking location	Car Access	0	5	15
	Parking cost	Car Parking cost	10	10	60
	Congestion	Car In-vehicle time	43, 13	43, 13	43, 13
Bike	Vehicle type	Bike In-vehicle time	50	50	38

7.2. Model Splits Rijnenburg

Once the population and the values of the attributes are defined, a simulation of ten runs is performed to calculate the modal splits for the commuting of Rijnenburg during rush hours and off-peak hours. This number of runs satisfies the minimum required sample size according to the equation 5.14.

Firstly, the modal split for three design scenarios for rush hours is presented in figure 7.1). It can be observed that modal splits are significantly different in three scenarios. In the Conventional scenario, the car alternative is dominant as expected, with a share of 66%. The choice for the bike is three smaller than the car but still almost twice as high as the choice for PT. This means in this scenario, the PT is very unattractive and even better is disregarding the long cycling time.

The Sustainable scenario leads to a more balanced share between the modes, though the share of bike is still smaller than PT and car. The share of PT and car became almost equal (37.6% for PT and 39.8% for car), and the share of bike stayed almost the same as in the Conventional scenario. This is because there is no change in the utility of bike alternative. In contrast, the share of choosing PT increased strongly, almost twice as high as the share of bike. The reduction of in-vehicle time and access distance have made the PT more attractive to commuters. The share of car did increase significantly, but the difference with the PT is still minimal. Thus, there is still room for improvement in this scenario to ensure a significant difference and convince commuters to more often opt for the PT option.

Lastly, in the Ambitious scenario, the attractiveness of the car is very low. This is because of the improvement of the PT infrastructure and the availability of the e-bike. The access time and the in-vehicle time are shorter than in the sustainable scenario, and travellers do not even need to transfer in their trips. on the other hand, the decreased attractiveness of the car (with a share of 2.4%) is caused by the longer access time to the parking facility, since the parking is now located outside Rijnenburg and the parking fee has been strongly increased.

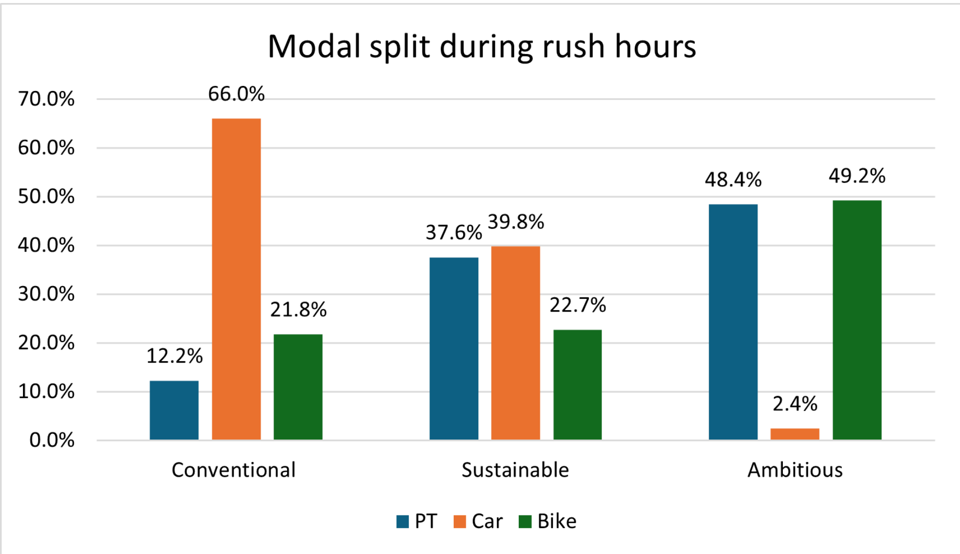


Figure 7.1: Predicted modal split during rush hours for Rijnenburg

However, commuting (and travelling) does not only happen during the rush hours but also during the off-peak hours. Therefore, the modal share of Rijnenburg during the off-peak hours also has to be taken into account for the assessment of the design scenarios. Outside rush hours, no congestion is expected on the roads and therefore the in-vehicle time for car use decreases from 43 to 13 minutes. In figure 7.2, the modal split of commuting during off-peak hours is presented.

In comparison to the modal split during rush hours, it can be observed that the share of car use became more dominant in the Conventional scenario, with larger difference with the share of PT and bike. For the Sustainable scenario, the share of PT is now much lower than car's (19.2% for PT and 69.3% for car). While during the rush hours the difference between the share of these two modes is much smaller. Shorter in-vehicle time clearly makes the car alternative more attractive in comparison to PT and bike. The ratio between the share of PT and bike are still the same since there is no change in the utility of these modes. In the Ambitious scenario, the share of car use also increased, but with smaller portion (5.5%). Therefore, the share of car use is still much lower than PT and bike.

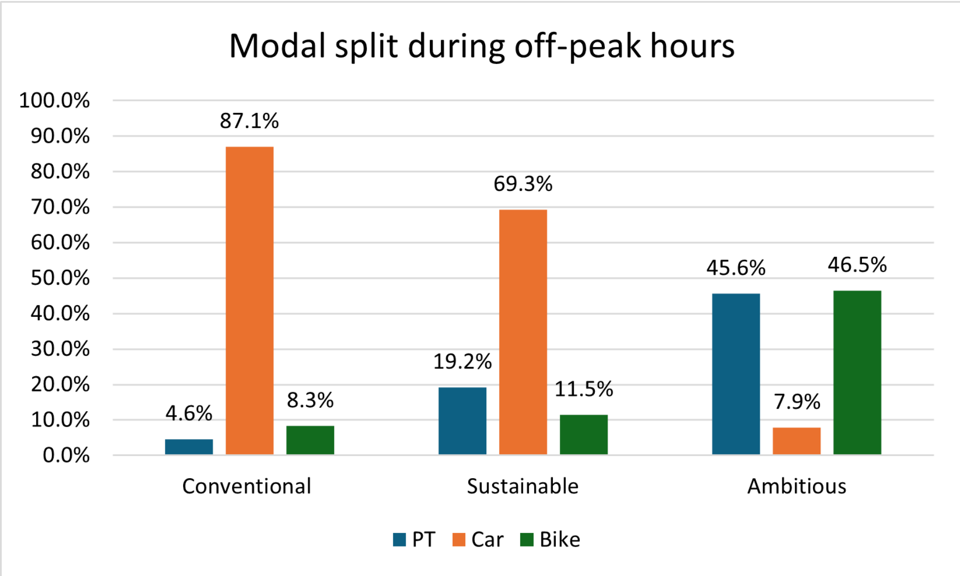


Figure 7.2: Predicted modal split during off-peak hours for Rijnenburg

From the obtained commuting modal splits for Rijnenburg, it can be seen that in the Conventional scenario the car use is dominant during rush hours and off-peak hours. In the Sustainable scenario, the share of PT and car became almost equal during the rush hours. However, when the effect of congestion is absent, car use becomes more attractive for travellers and the share of the car increases immediately. This leads to a drop of 50% for both PT and bike. Only in the Ambitious scenario, the share of car is minority. This is because of the increasing of convenience of the PT, shorter cycling time of the e-bike and more difficulty of using car (longer access time and more expensive for parking).

In comparison to the conventional situation, the Sustainable scenario does reduce the choice of using the car for commuting trips. However, it seems to be not effective enough, the share of cars is still dominant during the off-peak hours. To enhance the prevalence of sustainable transportation modes (PT and bike), the Ambitious scenario should be implemented for Rijnenburg.

7.3. Sensitivity Analysis

From the modal splits as results of the model application, it can be concluded that the Ambitious scenario should be implemented in Rijnenburg. However, this scenario requires very ambitious policy measures in the area's mobility design. In this scenario, two variables are very ambitious and may not be realistic for realisation, namely the direct travelling of the PT and the provision of an e-bike as commuting reimbursement. The direct PT lines are very desirable for all new residential areas, but it is mostly impossible due to financial and technical reasons. For the reimbursement of E-bike, there are companies that have already implemented their reimbursement system, but the decision is not in the hands of the government. Therefore, it is difficult to ensure the availability of every commuter from the position of the government. Because of the mentioned reasons, the sensitivity of other design variables should be analysed to explore the relationship between the design variables and the outcome.

Four continuous design variables, access time for PT and car, PT in-vehicle time, and the car parking cost have been analysed for their effectiveness. For an optimal analysis for overall outcome, the Sustainable scenario during off-peak hours is used as basic conditions for the sensitivity analyses in this section.

PT access

The first analysed variable is PT access time, which is directly related to the stop density in the area. The analysis examines how modal shares change in response to variations in PT access time, as illustrated in Figure 7.3. The plot reveals a clear linear relationship between PT access time and the modal shares for PT, car, and bike, though each mode exhibits different development trends (slopes). The PT share shows a negative trend, indicating that as access time to the PT system increases, the share of PT decreases steadily. This suggests that people are less likely to use public transport when it requires more time to reach a stop. Conversely, the shares of car and bike increase as PT access time increases. While the bike share shows a slight positive trend, this increase is minimal compared to the more pronounced changes in PT and car shares, indicating that biking is less sensitive to changes in PT access time. Among the modes, the PT share has the steepest slope, suggesting that PT access time has the most significant impact on the share of PT compared to car and bike for the same change in PT access time.

Overall, car share consistently remains higher than PT share. Even with a PT access time of zero, hypothetically placing a PT stop directly in front of every home, the preference for cars still dominates. Therefore, it may be reasonable to maintain the PT access time at 7 minutes, as proposed in the Sustainable scenario, to balance feasibility with efforts to encourage PT usage.

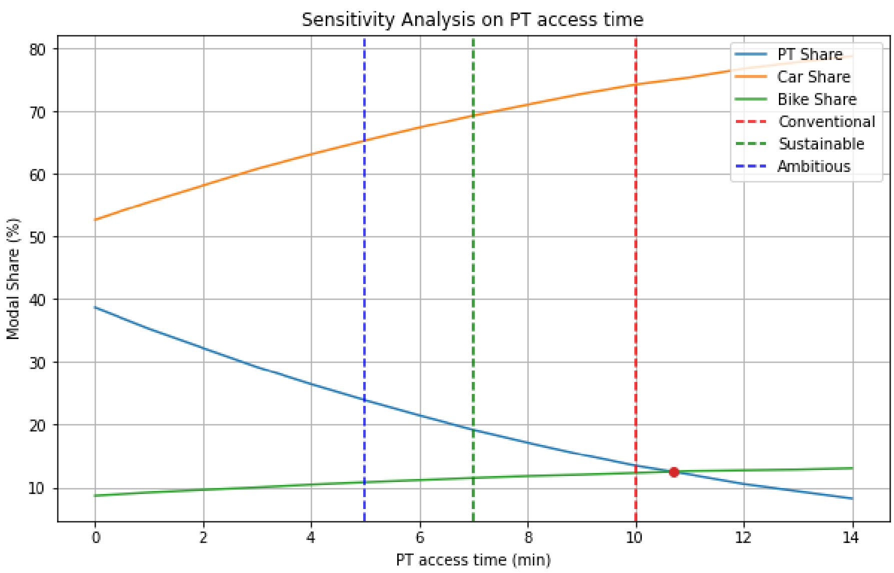


Figure 7.3: Effect PT access time on different mode share

PT in-vehicle time

When looking at the sensitivity of the PT in-vehicle time, almost a similar development behaviour as with the PT access is applied. The share PT will decrease when the in-vehicle time becomes longer while the share of car and bike increases. When the shortest (possible) IV time is applied, the share of car is still around 50% higher than the share of PT. The only difference is the presence of an intercept point at a in-vehicle time of 11.85 minutes. This intercept indicates that if the PT in-vehicle time is shorter than 11.85 minutes, the share of PT becomes higher than the car's. Unfortunately, this short in-vehicle time is impossible for the given travel distance. Therefore, PT in-vehicle time should be kept at 31 minutes.

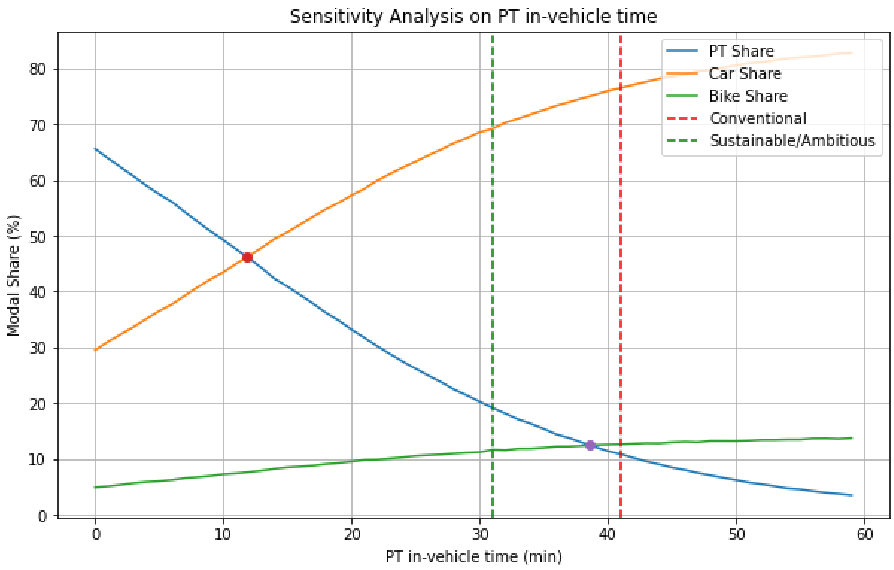


Figure 7.4: Effect of PT in-vehicle time on different mode share

Car access time

Completely different than the previous two PT variables, figure 7.5 shows a clear inverse relationship between car share and car access time. As car access time increases, the share of commuters using cars decreases sharply. The car share demonstrates high sensitivity to access time, with a sharp decline as access time increases. The PT share shows a strong increase in response to longer car access times, indicating a moderate to high sensitivity. This reflects the idea that as cars become less accessible, public transport becomes a more attractive option. The bike share, while positively affected by increased car access time, shows relatively low sensitivity compared to PT and car shares. This suggests that while biking may gain some users as car access becomes more cumbersome, its growth is less responsive to changes in car access time.

The plot features an intercept point where the PT share surpasses the car share at a car access time of 16.75 minutes. This indicates that more people may opt for PT over driving if car access time is extended beyond this point. This is higher than the defined value in the Ambitious scenario, but the difference is not very large; therefore, an access time to the car parking of 16.75 minutes could be considered in the design.

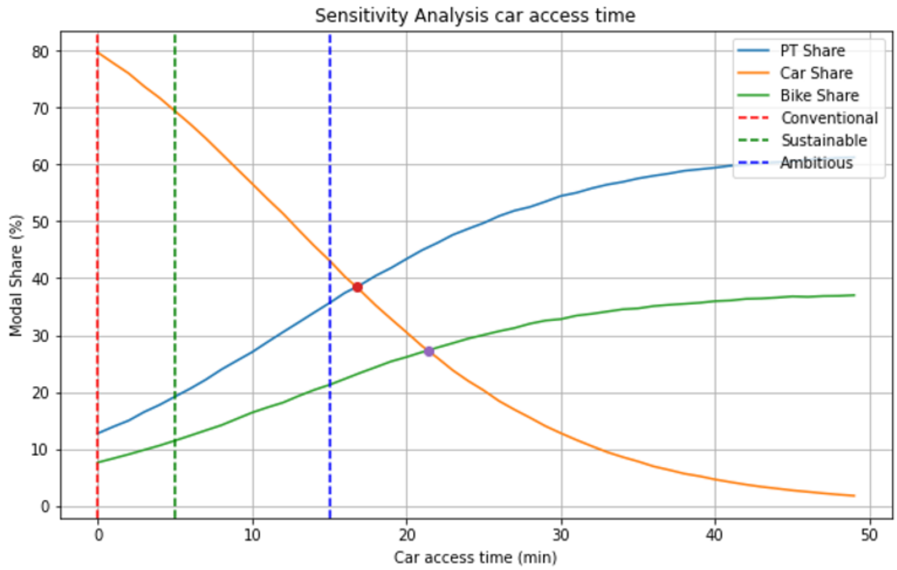


Figure 7.5: Effect of car in-vehicle time on different mode share

Car parking cost
Lastly, the mode shares are also plotted against the monthly parking cost variable in figure 7.6. Overall, this plot follows the almost similar pattern of the car in-vehicle time. Higher parking cost also leads to lower share for car use and higher share of PT and bike. The car share shows a steep decline as parking costs rise, indicating that car usage is highly sensitive to financial disincentives. commuters are likely to consider public transport as a more cost-effective alternative. The intercept between the car and PT shares occurs at a parking cost of €84.36, which is notably higher than the parking cost indicated in the Ambitious scenario. This suggests that to achieve a car share lower than PT share requires parking costs that exceed the levels proposed in even Ambitious scenarios.

The car parking cost can be an effective strategy to maximise the discouragement of car use and ownership in the area. This sensitivity analysis suggests that substantial increases in parking costs could be a key strategy for reducing car dependency and promoting more sustainable transportation options due the easy implementation compared to other technical design strategies.

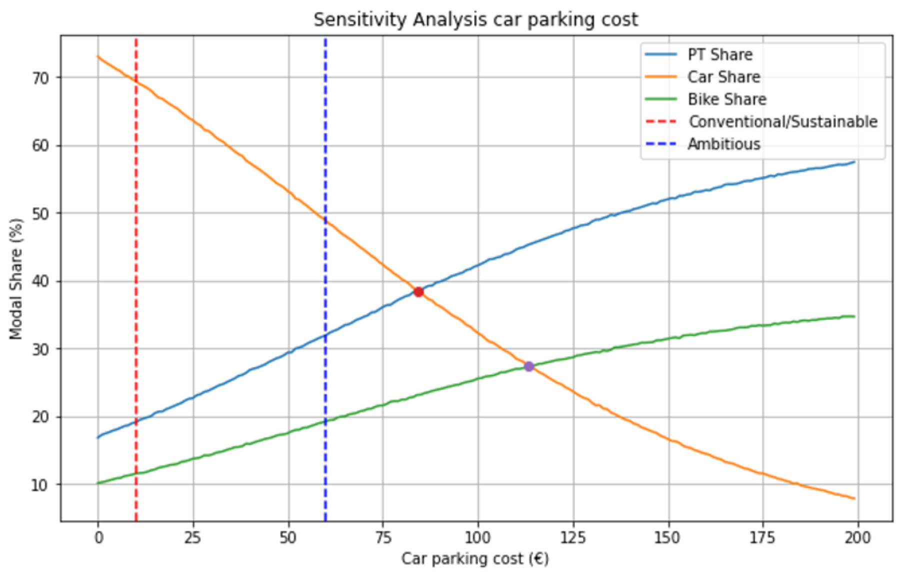


Figure 7.6: Effect of car parking cost on different mode share

Impact on the modal split for Rijnenburg
The analysis of PT access time, PT in-vehicle time, car access time, and car parking costs reveals distinct relationships between these variables and commuting mode shares. PT access time and in-vehicle time both show that as the convenience of public transport decreases (through longer access or travel times), car usage becomes more dominant. Conversely, increasing car access time and parking costs have a strong, inverse effect on car share, making public transport and biking more attractive options.

Besides, the shares of all mode alternatives exhibit the almost same sensitivity. Car share is most sensitive to the changes in the design variables, and PT follows with a less steep slope. Biking behaviour least reacts to the change of the variables, it may be influenced more by other factors such as infrastructure, safety, and personal preference rather than just car access convenience.

Notably, two interesting intercept points identified in the analyses of car access time and parking costs suggest specific thresholds at which public transport becomes more competitive than car use. These values could be considered for adjustments in the Sustainable scenario to enhance a low share of car. By applying two new values, 16.75 minutes for the car access time and €84.36 for the monthly car parking cost in the Sustainable scenario, a new modal split has been determined in figure 7.7. The new values indeed contribute to a more desirable outcome. In this scenario, the car has been least chosen also during the off-peak hours when there is no congestion on the road. The PT becomes dominant and the bike is also much more preferable than the car. Therefore, the found thresholds provide valuable guidance for setting realistic targets in the design of mobility strategies.

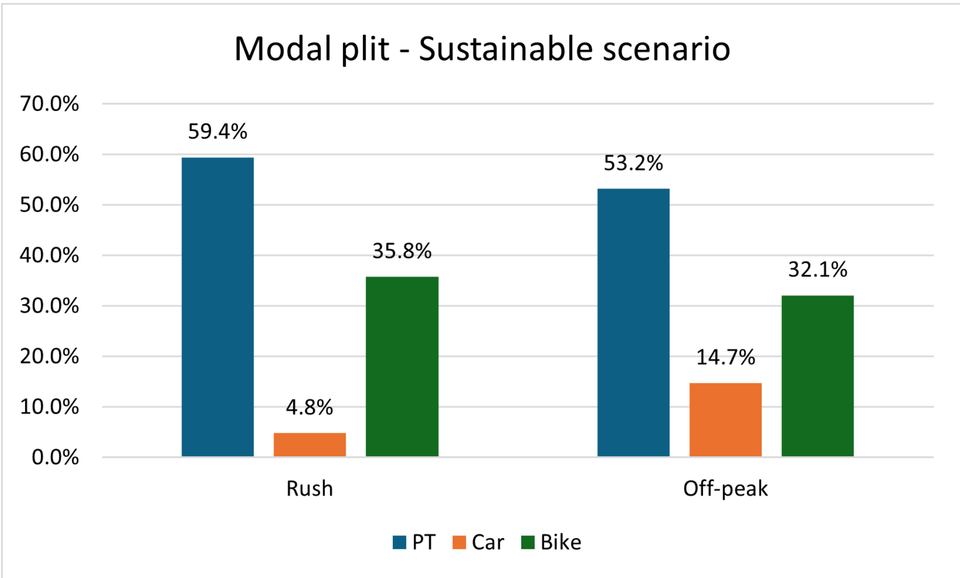


Figure 7.7: Modal splits with modified values

7.4. Conclusion on Model Application

In conclusion, when no intervention is applied in the development of Rijnenburg, the area will develop in a conventional way, likely to be the usual suburban area in The Netherlands. This automatically leads to a demand for car use and car ownership among the new residents. The more sustainability-oriented approach, Sustainable scenario, does improve the infrastructure and the operational service level for the PT system and making the local street car-free, resulting in a more balanced share between the car and the PT. However, the difference between the share of PT and car is not very clear during the rush hours and the car remains dominant option during the off-peak hours. In the most ambitious scenario, car use is most restricted by having quite a long access to the parking facility and expensive parking fees. At the same time, the use of more sustainable and healthier modes is stimulated, resulting in the most attractive options to discourage the car use in Rijnenburg.

However, the Ambitious scenario contains two unfeasible strategies, namely the direct PT lines for commuting and the availability of e-bikes for all commuters. Therefore, a sensitivity analysis of the remaining design variables with the Sustainable scenario as the initial condition could provide more insight into the impact of individual variables on the choice behaviour of travellers. On top of that, more optimal thresholds for the design strategies have been found for balancing ambitious policies with realistic and feasible interventions.

The model application results, including sensitivity analysis, suggest the Sustainable scenario (with higher car access time and car parking cost) for the mobility development in Rijnenburg to ensure a higher share of PT compared to car's. This scenario includes a maximum PT access time of 7 minutes and an in-vehicle time of 31 minutes with a transfer during the trip. These strategies lead to a maximum stop spacing of 1000 metres (see section 5.1.2) in Rijnenburg and the wheel form for the global PT network. Regarding the car parking planning, the parking facilities should be located outside the area with a minimum access (walking) time of 16.75 minutes. This is equivalent to an access distance of 1146 metres, using the walking speed of 1.14 m/s (Wirtz and Ries, 1992). Additionally, a high monthly parking cost of minimal €84.36 should be charged to enhance the low preferences for the car alternative. There is no specific strategy for the bike alternative due to the complexity of implementation. Table 7.3 provides an overview of the found attribute levels and the strategies with corresponding design variables.

Table 7.3: Overview of design strategies for Rijnenburg

	Attribute	Attribute levels	Design strategies	Applicable levels
PT	Access	7 (min)	Stop spacing	1000 m
	In-vehicle time	31 (min)	Route planning	Wheel network
	Transfer	1	Route planning	Transfering at normal stop
Car	Access	16.75 (min)	Parking location	Outside the area
	Parking cost	84.36 (€)	Parking cost	High cost
Bike	In-vehicle time	50 (min)	Vehicle type	Standard

8 Discussions & Limitations

After obtaining the results, the research findings are discussed in section 8.1, where key insights and implications are explored. This is followed by section 8.2, which offers a reflection on the study's limitations, acknowledging areas for potential improvement and considerations for future research.

8.1. Discussions

The first point of the discussion focuses on the quality of the collected survey data and whether it accurately reflects the choice behaviour of the target population. As noted in section 5.1.5, the study aimed for a minimum sample size of 224 respondents to ensure reliable and robust model estimation. However, due to the duration of this research, only 200 usable responses were collected, resulting in a shortfall of approximately 10%. This shortage could potentially impact the reliability and certainty of the model estimation results, as a smaller sample size may not capture the full range of variability in travel behaviour.

Furthermore, the survey did not have a wide diversity of respondents. Since the survey was distributed mainly among urban residents in big cities, they may have similar experiences with more robust PT facilities, heavy congestion, and parking issues in big cities. Furthermore, personal networking was also used for survey distribution. Undoubtedly, a great part of the respondents are highly educated students and professionals in the field of (sustainable) mobility and transportation. They are expected to have a wider knowledge and awareness of sustainable transportation than the average traveller. Possibly, from the accountability of their position, they could have the force avoid the less sustainable mode alternative. This leads to a large negative value for the ASC of the car alternative.

Regarding the choice modelling, the decision to use the MNL model for the model application can also raise a discussion. The MNL model is chosen because of its higher accuracy and better interpretations of the observable model attributes. However, the Panel ML still has higher explanatory power with higher model performance indicators, see section 6.2.4. Furthermore, this model also considers the panel effect in repeated answers, which is more appropriate for panel data. Using the MNL model, the randomness in individual responses may be neglected. Therefore, additional assessment criteria may result in different decisions for the choice of the model.

From the model estimation results, it appeared that the type of public transit vehicle, waiting time, and transfer distance are insignificant and, therefore, do not contribute to the explanatory power of the models. However, this does not necessarily mean that the mentioned factors do not affect the behaviour of commuting mode choice. The small sample size of survey data may have led to uncertainty and insignificant model estimation. Besides, it is also possible that the complexity and volume of information in the survey caused respondents to unintentionally overlook or underestimate this factor in their decision-making process. Consequently, these factors are eliminated from the model and are not captured in the decision-making process of mode choice behaviour.

Lastly, the final modal splits of Rijnenburg in figure 7.7 are derived from the Sustainable scenario and two critical intercept points of car access time and parking cost, where the shares of public transit (PT) and car usage are equal. This approach was chosen to identify the critical thresholds between PT and car use. By incorporating both of these critical values, a higher prevalence of PT usage is ensured, leading to a low share of car usage during both peak and off-peak hours. However, it is important to note that these are not necessarily the optimal solutions for Rijnenburg's mobility design. Depending on the policy objectives of urban planners, more relaxed values could be considered, taking into account other socio-economic factors and priorities.

8.2. Limitations

While this research provides valuable insights into mid-distance commuting behaviour in peripheral areas, it is important to recognise several limitations that may impact the study's findings and conclusions.

First, the research primarily focused on the commuting purpose for mid-distance travelling, considering that the majority of road users belong to this group of travellers. However, this assumption may be too narrow, as it overlooks the importance of addressing various travel purposes and distances, which are also essential for creating a comprehensive and effective mobility strategy. Furthermore, for the SP experiment, the Utrecht Science Park (USP) was used as the destination for the commuting trip. However, when applying the model to other origin-destination pairs, it is crucial to account for the specific characteristics of different destinations. Factors such as the availability of public transit, the density of economic activities, and the overall attractiveness of the destination can significantly influence commuting behaviour. Therefore, the model's applicability might be limited if these destination-specific characteristics are not carefully considered.

Secondly, the design of the SP experiment has its own limitations. Although the research problem emphasised that peripheral areas have distinct location characteristics compared to other areas, these differences may not have been effectively communicated to or recognised by the respondents. This lack of clarity might have led respondents to overlook the unique aspects of peripheral areas during the experiment, resulting in similar decision choices between the commuting mode choice of the peripheral and other areas. Consequently, this could fade the distinctions between commuting mode choices in peripheral and non-peripheral regions, potentially impacting the accuracy of the findings.

Furthermore, the quality of the information and choice situations presented in the survey poses another limitation of this research. In an effort to capture as many influencing factors as possible, the survey included a high level of detail. This complexity may have overwhelmed respondents, potentially leading to a loss of crucial information in their decision-making process. Additionally, the number of choice situations each respondent had to face is also a concern. Although the number of choice tasks in this study was fewer than the maximum recommended by Arentze and Molin (2013), nine questions might still be experienced as excessive by respondents. This could increase the likelihood of inconsistent answers, which in turn may introduce greater randomness and error into the Panel Mixed Logit (ML) model, potentially affecting the reliability of the results.

Finally, this study primarily focuses on traveller's choice behaviour to develop design strategies for public transportation (PT) and car parking use. However, several other critical factors were not addressed, such as personal perceptions, costs, and the technical feasibility of implementation. Commuters' perceptions of reliability, safety, and attitude could play a significant role in their transportation choices, and neglecting these subjective factors might lead to strategies that fail to resonate with users. Additionally, the substantial financial investments required for ambitious PT infrastructure could make these strategies impractical without considering budget constraints. Furthermore, the technical challenges of implementing new infrastructure, such as new tram lines or extra stops, could pose significant hurdles. To create more viable and approachable mobility solutions, it is essential to explore these factors in the planning process.

9

Conclusion & Recommendations

In conclusion, this report presents a comprehensive study of commuting mode choice behaviour, aimed at developing effective mobility design strategies for peripheral areas. This chapter begins by answering the research questions in section 9.1, followed by practical recommendations for mobility design in section 9.2. Finally, section 9.3 offers suggestions for future research, highlighting areas that could further enhance the understanding and implementation of sustainable mobility solutions in peripheral regions.

9.1. Answers to The Research Questions

The main objective of this research is to define the mobility design strategies for the development of peripheral areas. Therefore, the main research question of this study is raised as:

"Which mobility design strategies, regarding public transit and car parking planning, should be implemented to reduce car usage and enhance the attractiveness of public transit in a peripheral urban area?"

To gain insight for answering the main question, four sub-questions have been developed. The answers to these sub-questions form the key components for defining the mobility design strategies for peripheral areas.

1. Which characteristics of travellers, public transit and car use could affect the mode choice behaviour of travellers?

For the first sub-question, a literature study on commuting travel behaviour is performed. On top of this, several key characteristics that could influence the mode choice behaviour of travellers, particularly for the commuting choice behaviour, are identified. The characteristics related to the traveller, car use and public transit form the answer to the first sub-question.

Regarding the traveller's characteristics, demographic and mobility factors play a significant role. Socio-demographic factors such as age, educational level, employment status, and individual income are expected to affect the mode choice behaviour of travellers. **Age** influences preferences, with older travellers favouring cars for comfort and safety and younger commuters leaning towards cycling and walking. Higher **educational levels** are linked to environmental awareness, promoting sustainable public transit and cycling modes. **Employment status** affects commuting needs; the employees often require regular and reliable transport, impacting their mode choice. **Income** also plays a crucial role; higher-income travellers prefer cars for convenience, while lower-income groups depend more on public transport. Besides, the personal mobility factors of disability and availability are expected to significantly impact the mode preferences. Due to the physical barriers, travellers with **mobility disabilities** are not free to choose all of the modes and often require more advanced accessibility of the transport mode. The **Mobility Availability** factor involves the ownership of different transportation modes. Since they already own these modes, their mode choices can be directly influenced by firsthand experiences with the advantages and disadvantages of each option. Lastly, for the **household characteristics**, larger travel group size and the presence of young children in the family often necessitate car use for convenience.

Regarding Public Transit, longer **in-vehicle times** can discourage public transit use, especially when cars offer shorter travel time. **Access time and distance** to transit stops are critical; longer access times discourage the choice of using PT. The high frequency of the PT service reduces **waiting times**, enhancing public transit's attractiveness. **Transfer time and distance** add to total travel time, making direct routes more appealing. Shorter **egress times** improve convenience, and affordable **fares** encourage public transit use. Modern, comfortable, and reliable **vehicles** attract more users, so the choice between different types of PT, such as light rail or bus, could affect the mode choice behaviour.

Same as the influencing factors for Public Transit, in-vehicle time, travel distance, costs, access time and distance could have a certain impact on the choice of car use. Shorter in-vehicle times make driving preferable due to the speed. Access time and distance to cars or parking facilities also play a role; convenient access makes car use more attractive.

Furthermore, high-density urban areas with better-developed public transit systems encourage its use. The urban type directly relates to the **parking policy** and could also be critical for car ownership. Limited or expensive parking at residential places could discourage car ownership and impact the decision to use a car. In addition, the **reimbursement form** from the employer also contributes to the choice of a specific mode for commuting and travelling.

2. Which mobility design strategies could be applied in the development of peripheral urban areas?

To be able to answer the second sub-question of the research, the case study of Rijnenburg is analysed due to its peripheral properties. To ensure sustainable mobility development in Rijnenburg, several comprehensive mobility design strategies can be implemented. These focus on enhancing public transit accessibility, reducing car dependency, and promoting active transportation modes.

The strategies for Public Transit Planning involve the planning of local and global networks, including the planning of **stop density**, **route planning**, operational **line frequency** and choosing the suitable **vehicle for the PT** operation in the area. Increasing stop density reduces access distances, making transit more appealing. Have a route planning, using spoke or wheel network forms, optimise travel time, transfers and enhances the robustness of the PT network. Improving line frequency reduces waiting times, making public transit more time-efficient. Selecting the appropriate type of operational vehicle, buses (BRT) or trams (light rail), ensures experience and comfort for the travellers.

For car planning, the strategies should contribute to decreasing car usage and ownership by making car travel less attractive. Strategically locating parking facilities influences car dependency. Different **parking locations**, i.e., at home, in communal neighbourhood lots, or on the outskirts of a residential area result in different access distances to the car. The most restrictive approach involves car-free urban with parking facilities outside the residential area. Setting **parking fees** is another strategy to reduce the car-ownership. High parking costs discourage car ownership, making alternative modes more financially attractive and encouraging residents to use public transit or cycling.

Next to directly in steering in the area development, the commuting mode choice behaviour can also be influenced by the **commuting reimbursement**. An e-bike support plan can also be an option for reimbursement for commuting instead of the conventional forms such as a leased car, mileage or a PT subscription. This relates directly to the cycling trips' in-vehicle time.

The defined mobility strategies can be applied to the development of Rijnenburg. These strategies have different applicable levels for each design scenario (Conventional, Sustainable and Ambitious), which are developed from the three shaped development directions by the local government. Conventional scenario follows the traditional design principles for outskirt areas with an imbalanced ratio between the high convenience for car use and basic quality of PT due to low demand. The Sustainable scenario offers a more balanced quality between infrastructure for PT and cars, aiming to create more sustainable urban mobility. In the Ambitious scenario, the most ambitious levels will be applied for the PT planning and the highest restrictions will be applied for car use by prohibiting car traffic in the whole area and high cost for parking facilities.

3. To what extent do these characteristics affect the mode choice behaviour of commuters in a peripheral urban area?

On top of the explored influencing factors from the literature and the potential mobility design strategies for the case study area Rijnenburg, a discrete choice model is constructed only involving the attributes with significant impact.

For public transportation (PT), the requirement to transfer had the largest negative value ($\beta_{transfer}$ = -0.803), indicating a significant impact of transfer into the utility of the PT. However, this does not necessarily mean that the transfer heaviest weights in the choice behaviour of the commuter, as it has a linear relationship with the binary value of the transfer attribute.Next, PT access time negatively impacted the PT utility (β_{PT_access} = -0.14), with longer access times reducing its attractiveness. Additionally, PT in-vehicle time had a moderate negative effect. Regarding personal characteristics, the dummy variables of income ($\beta_{PT\_Mid\_income}$ and $\beta_{PT\_High\_income}$) and commuting reimbursement form ($\beta_{PT\_Reimbursement\_PT}$ and $\beta_{PT\_Reimbursement\_PT}$) are in complete opposition to each other. Middle income increases PT utility, while high income decreases it with an absolute value of 0.299. Commuting reimbursement for PT subscriptions increases PT utility, while the availability of the leased car decreases it with an absolute value of 0.364.

For car alternatives, the alternative-specific constant was negative (ASC_{car} = -1.35), indicating a general aversion to car use among commuters. Car access time had a significant negative impact, with longer access times reducing utility. The smallest attribute regarding the alternative property is the parking cost; it has a value of $\beta_{parking}$ = -0.0174), indicating the little impact of increasing the parking cost. Regarding personal characteristics, age, mobility disability and the availability of a driving license positively affected car utility. However, these impacts are very small in comparison to the other attributes.

Finally, the bike's attributes showed a different trend. The alternative-specific constant for bikes was positive (ASC_{bike} = -2.19), indicating a strong preference for cycling. However, the in-vehicle time attribute had the smallest negative impact (β_{bike_IVtime} = -0.134), with longer biking times greatly reducing its utility. The availability of the PT subscription and car also reduces the utility of the bike alternative. Conversely, the availability of bikes reduced the utility of both PT and car significantly ($\beta_{PT\_Availability\_bike}$ = $\beta_{car\_Availability\_bike}$ = -0.626), indicating that if commuters have a bike, they are less likely to choose PT or car.

4. To what extent are these mobility design strategies effective in reducing car use among commuters in peripheral urban areas?

The mobility design strategies for Rijnenburg can be applied at different levels. These applicable levels are implemented in the three design scenarios from sub-question 2: Conventional, Sustainable and Ambitious. By calculating the modal splits for Rijnenburg using these three design scenarios, the effectiveness of the design strategies can be evaluated.

From the obtained modal split of the commuters in Rijnenburg, it can be observed that the car has the highest share in the Convention scenario. Although the share of cars does decrease during rush hours due to congestion on the roads, it is still 53.8% higher than the share of PT. The share of bikes is also almost twice as high as that of PTs, while no specific strategy has been applied to promote cycling. These outcomes indicate that the PT does not attract travellers at all; they prefer other modes regardless of the heavy congestion and long-distance cycling.

In the Sustainable scenario, strategies for better PT facilities are applied: shorter distance to the stop and shorter in-vehicle time due to another PT network form. At the same time, the car is prohibited from the local street, leading to longer access distance to the car. These changes have led to a more balanced distribution of the mode split during the rush hour, with the share of PT becoming nearly as large as car's. However, when the congestion effect is absent, travellers seems to prefer car use again, leading to a new increasing of cars during off-peak hours.

Finally, the Ambitious scenario has the largest impact on mode choice behaviour as expected. With the highest restrictions on car use and maximal investment in PT and bike facilities, this scenario is most likely to ensure the absolute prevalence of public transport and biking as the primary modes of commuting. During both rush and off-peak hours, the car remains the least attractive option with a maximum share of 7.9%. Therefore, the Ambitious scenario has the most impact on shifting commuter preferences toward PT and bike, and enhancing the absolute prevalence of sustainable transportation modes (PT and bike).

However, the Ambitious scenario includes very ambitious strategies that may be unfeasible for real-world implementation. Specifically, the provision of direct PT trips without transfer and an e-bike for all commuters, these strategies might not be achievable due to financial and technical constraints. The sensitivity analyses of the modal shares suggested a more feasible approach for Rijnenburg. This approach integrate higher restricted levels for the car parking-related strategies into the defined PT planning strategies from the Sustainable scenario. This balanced scenario should be considered for implementation in the mobility development of Rijnenburg, combining the achievable aspects of the Sustainable scenario with targeted adjustments to ensure practicality and effectiveness in promoting sustainable transportation modes.

Main research question

Based on the research outcomes and key findings from the answers to the sub-questions, it can be concluded that a combination approach of the Sustainable scenario and high restricted levels for the car parking-related strategies could ensure sustainable mobility development with prevalence preferences for public transit and bike for Rijnenburg. Therefore, the mobility design strategies regarding public transit network and car parking policy should be implemented to reduce car usage and enhance the attractiveness of public transit in a peripheral urban area are:

- For the planning of public transit, transferring, access time and in-vehicle time are the founded design variables that significantly impact the mode choice behaviours of commuters. The access time is directly related to the stop spacing in the local network. The access (walking) time to the PT stops should be considered at a maximum of 7 minutes, which leads to a stop spacing of a maximum of 1000 metres. At the higher network level, the transit network needs to be designed to ensure the lowest possible in-vehicle time to the big working places in the region, with a desirable maximum in-vehicle time of 31 minutes. Transferring during the trip will result in higher attractiveness for the use of public transit, but it is not required.
- For car parking planning in the area, the access distance to the car was found to be the most effective strategy to hold back car use. A minimal walking time of 16.75 minutes is found for the access to the car, which results in an access distance of 1146 metres. Additionally, a high monthly fee for the parking facility can be an effective policy measure to enhance the resistance to low number of car use and car ownership in the peripheral areas. Although the effort to access the car is found to weigh heavier than the fee that they have to pay for the parking facility, this economic strategy might be easy to adopt in practice.

In summary, the research findings underscore that a strategic combination of sustainable public transit and stringent car parking planning measures can significantly promote sustainable mobility in peripheral urban areas. By optimising public transit variables such as access time, in-vehicle time and reducing transfers, alongside implementing extended access distances and higher fees for car parking, these strategies can effectively shift commuter preferences towards public transit and cycling, thus reducing car dependency and supporting sustainable development of new urban areas.

9.2. Recommendation for Practice

Based on the conclusions of this research, recommendations for mobility design for Rijnenburg are discussed in this section. As obtained in the research conclusion, a stop spacing of a maximum of 1000 metres is needed, resulting in a minimum of three transit stops in Rijnenburg. It is also recommended that the extension of the public transit line of Rijnenburg should be integrated with a Wheel network form into the existing network of Utrecht. This should enhance a short in-vehicle time for commuting between Rijnenburg and Utrecht Science Park. When possible, consider a direct line to USP and other large workplaces in Utrecht, which will increase the attractiveness of public transit use.

Regarding the car parking strategies, the conclusion recommends a minimum walking distance of 1146 metres, resulting in a parking facility outside the residential area of Rijnenburg. This not only creates resistance to using cars but also results in a car-free, safe and healthy living environment for the residents of Rijnenburg. Furthermore, a cost of €84.36 per month also should be maintained for using the car parking facilities in Rijnenburg. This amount is slightly higher than the current parking cost for high urban areas of Utrecht, but this will create significant economic resistance for car ownership and shift the commuters from car use to public transit.

Additionally, as discussed in section 8.1, the impact of the type of public transit vehicle, waiting time, and transfer distance are neglected in the derived choice model due to the uncertainty in the model estimation. To address the possibility that these factors were overlooked or underestimated, it is recommended to further investigate the potential impact of these factors. This could be achieved by conducting a new, refined SP experiment with a more focused design to emphasise these variables or by continuing with the distribution of the current survey to increase the sample size. Both approaches could help clarify the significance of these factors in influencing commuting mode choice behaviour, leading to more accurate and robust mobility design strategies.

As reflected in section 8.2, the realisation costs and technical possibilities were not included in this research, even though they are critical factors in the mobility design for area development. Therefore, it is recommended to conduct an optimisation study that integrates the derived design strategies with cost and technical constraints. Solutions from this optimisation process could provide a more balanced and practical approach to mobility design for Rijnenburg, ensuring that the proposed strategies are not only theoretically effective but also feasible within the financial and technical limitations of the project.

Lastly, this research only evaluates the modal split of the commuting middle distance. However, it is important to also evaluate and address the needs and preferences of commuters travelling for shorter and longer distances. By incorporating strategies that cater to all distance ranges, a more comprehensive and effective transportation plan can be developed, ensuring that the mobility needs of the entire population are met.

9.3. Recommendation for Research

Besides the recommendations for the practice, there are also various recommendations for future research. First of all, it is recommended to simplify the design of Stated Preference experiments by reducing both the number of details included and the number of choice situations presented to respondents. By focusing on the most critical factors and streamlining the survey content, researchers can help prevent respondent fatigue and information overload, which often lead to random or inconsistent answers. This approach would likely result in more stable and reliable data, enhancing the quality of insights into commuter behaviour and reducing error randomness in models like the Panel ML model.

Given the limitation that the survey lacked diversity among respondents, it is recommended that future research includes a pilot study with a broader and more diverse group of participants. This pilot study should aim to capture a wider range of experiences and perspectives by including respondents from various geographic locations, including peripheral and rural areas, as well as individuals from different educational backgrounds and professions. By doing so, the study would better reflect the diversity of the general population and provide more accurate and generalizable insights. Additionally, involving respondents with varying levels of familiarity with sustainable transportation could help mitigate biases related to professional expertise and personal accountability, leading to more balanced and representative findings.

Bibliography

Action. (2005). Parking policies and the effect on economy and mobility. https://www.europeanparking.eu/media/1207/cost-action-342-final-report-1.pdf

Ahmed, F., Catchpole, J., & Edirisinghe, T. (2020). Understanding young commuters' mode choice decision to use private car or public transport from an extended theory of planned behavior. *Transportation Research Record*, *2675*. https://doi.org/10.1177/0361198120967936

Alonso-González, M. J., van Oort, N., Cats, O., Hoogendoorn-Lanser, S., & Hoogendoorn, S. (2020). Value of time and reliability for urban pooled on-demand services. *Transportation Research Part C: Emerging Technologies*, *115*. https://doi.org/https://doi.org/10.1016/j.trc.2020.102621

Arentze, T. A., & Molin, E. J. (2013). Travelers' preferences in multimodal networks: Design and results of a comprehensive series of choice experiments. *Transportation Research Part A: Policy and Practice*, *58*. https://doi.org/https://doi.org/10.1016/j.tra.2013.10.005

Ben-Akiva, M., & Bierlaire, M. (2003). *Handbook of transportation science - second edition*. Kluwer Academic Publishers.

Bierlaire, M. (2020). A short introduction to pandasbiogeme. *Transport and Mobility Laboratory*.

Bongers, R., Warnecke, J., & Faber, J. (2023). Woningkrapte - deel 1: Gevolgen van krapte op de woningmarkt (nl representatief – jongeren – senioren). https://nieuws.aegon.nl/download/b348b4db-4e30-4d6c-a455-39103d6ecb06/onderzoeksrapportpersbericht-woningkrapte-gevolgenvankrapteopdewoonmarkt.pdf

CBS. (2024). Opleidingsniveau. https://www.cbs.nl/nl-nl/nieuws/2019/33/verschil-levensverwachting-hoog-en-laagopgeleid-groeit/opleidingsniveau

Centraal Bureau Voor de Statistiek. (2022). Demografische ontwikkelingen en de woningvoorraad. https://www.cbs.nl/nl-nl/longread/de-nederlandse-economie/2022/achtergrond-bij-de-huizenprijs-stijgingen-vanaf-2013/3-demografische-ontwikkelingen-en-de-woningvoorraad

Centraal Bureau Voor de Statistiek. (2023a). Hoeveel reisden inwoners van nederland en hoe? https://www.cbs.nl/nl-nl/visualisaties/verkeer-en-vervoer/personen/hoeveel-reisden-inwoners-van-nederland-en-hoe-

Centraal Bureau Voor de Statistiek. (2023b). Meer kilometers van en naar het werk. https://www.cbs.nl/nl-nl/nieuws/2023/27/meer-kilometers-van-en-naar-het-werk

Centraal Bureau Voor de Statistiek. (2024). Arbeidsparticipatie naar regio. https://www.cbs.nl/nl-nl/visualisaties/dashboard-arbeidsmarkt/werkenden/arbeidsparticipatie-naar-regio

Chatman, D. G. (2009). Residential choice, the built environment, and nonwork travel: Evidence using new data and methods. *Environment and Planning A: Economy and Space*. https://doi.org/10.1068/a4114

Chen, P., Boyd, T., & Williams, K. (2021). *Multimodal transportation planning*. U.S. Department of Education.

Cranenburgh, S. V., & Collins, A. T. (2019). New software tools for creating stated choice experimental designs efficient for regret minimisation and utility maximisation decision rules. *Journal of Choice Modelling*.

CROW-KpVV. (2021). Toepassen stomp voor duurzame gebiedsontwikkeling.

De Ridder, T. (2023). Optimal stop location analysis for urban tram systems.

Delice, Y., Özen, H., & mirnazmiafshar, E. (2019). Analyzing modal shift based on critical travel time for different trip purposes in medium-sized city. *International Conference on Transportation and Development*.

Dubey, S., Cats, O., & Serge, S. H. (2024). Understanding preferences for mobility-on-demand services through a context-aware survey and non-compensatory strategy. *Transportation Research Part C: Emerging Technologies*. https://doi.org/10.1016/j.trc.2023.104455

Esfeh, M. A., Wirasinghe, S. C., Saidi, S., & Kattan, L. (2021). Waiting time and headway modelling for urban transit systems – a critical review and proposed approach. *Transport Reviews*. https://doi.org/10.1080/01441647.2020.1806942

European Commission. (2024). Sustainable urban mobility planning and monitoring. https://transport.ec.europa.eu/transport-themes/urban-transport/sustainable-urban-mobility-planning-and-monitoring_en#:~:text=A%20sustainable%20urban%20mobility%20plan%20is%20a%20strategic%20plan%20designed,%2C%20participation%2C%20and%20evaluation%20principles.

Fietsknooppunt. (n.d). Fiets routes in rijnenburg (utrecht). https://www.fietsknooppunt.app/nl-nl/fietsen/routes/nederland/utrecht/utrecht/rijnenburg/a5-122516

Gemeente Nieuwegein. (n.d). Parkeer service - tarieven. https://nieuwegein.parkeerservice.nl/tarieven/

Gemeente Utrecht. (n.d-a). *Ontwikkeling rijnenburg*. https://www.utrecht.nl/wonen-en-leven/wijken/wijk-vleuten-de-meern/ontwikkeling-rijnenburg/

Gemeente Utrecht. (n.d-b). Parkeervergunning bewoner aanvragen. https://pki.utrecht.nl/Loket/product/072c0c626eaa6749c2e66a169466bace

Gemeente Utrecht. (n.d-c). Wonen. https://www.utrecht.nl/wonen-en-leven/wonen/

Goudappel. (n.d). Mobiliteitstransitie: De aanpak van maatschappelijke vraagstukken via duurzaam en toegankelijk vervoer. https://www.goudappel.nl/nl/themas/gebiedsontwikkeling/mobiliteitstransitie

Haustein, S., & Kroesen, M. (2022). Shifting to more sustainable mobility styles: A latent transition approach. *Journal of Transport Geography*. https://doi.org/10.1016/j.jtrangeo.2022.103394

Heinen, E., van Wee, B., & Maat, K. (2010). Commuting by bicycle: An overview of the literature. *Transport Reviews*. https://doi.org/10.1080/01441640903187001

Hjorthol, R. J., Levin, L., & Sirén, A. (2010). Mobility in different generations of older persons. the development of daily travel in different cohorts in denmark, norway and sweden. *Journal of Transport Geography*. https://doi.org/10.1016/j.jtrangeo.2010.03.011

Ibraevaa, A., de Almeida Correiab, G. H., Silvac, C., & Antunesa, A. P. (2020). Transit-oriented development: A review of research achievements and challenges. *Transportation Research Part A*. https://doi.org/10.1016/j.tra.2019.10.018

Johnson, R., & Orme, B. (2003). Getting the most from cbc, sawtooth software research paper series. *Sawtooth Software, Sequim*.

Jonkeren, O., Wust, H., & de Haas, M. (2019). Mobiliteit in stedelijk nederland.

Kennisinstituut voor Mobiliteitsbeleid. (2022). *Kerncijfers mobiliteit 2022* (tech. rep.). Ministerie van Infrastructure en Waterstaat.

Kepaptsoglou, K., & Karlaftis, M. (2009). Transit route network design problem: Review. *JOURNAL OF TRANSPORTATION ENGINEERING © ASCE*.

Kroesen, M. (2024). *Lecture: Sen1721 travel behaviour research*, Delft University of Technology.

Li, J., Chen, D., Li, X., & Godding, L. (2018). Investigating the association between travelers' individual characteristics and their attitudes toward weather information. *Travel Behaviour and Society*. https://doi.org/10.1016/j.tbs.2017.11.001

Li, Z., Hensher, D. A., & Rose, J. M. (2010). Willingness to pay for travel time reliability in passenger transport: A review and some new empirical evidence. *Transportation Research Part E: Logistics and Transportation Review*. https://doi.org/10.1016/j.tre.2009.12.005

Loo, B. P. Y. (2018). Realising car-free developments within compact cities. *Proceedings of the Institution of Civil Engineers: Municipal Engineer*. https://doi.org/10.1680/jmuen.16.00060

Molin, E. (2024). *Introduction to experimental designs - lecture 1 (sen1221 – part ii)*, Delft University of Technology.

Moyano, A., Solís, E., Díaz-Burgos, E., Rodrigo, A., & Coronado, J. M. (2023). Typologies of stations' catchment areas in metropolitan urban peripheries: From car-oriented to sustainable urban strategies. *Land Use Policy*. https://doi.org/10.1016/j.landusepol.2023.106903

Muñoz, B., Monzon, A., & Daziano, R. A. (2016). The increasing role of latent variables in modelling bicycle mode choice. *Transport Reviews*, *36*. https://doi.org/https://doi.org/10.1080/01441647.2016.1162874

Nabielek, K. (2012). The compact city: Planning strategies, recent developments and future prospects in the netherlands. *AESOP 26th Annual Congress*. https://www.pbl.nl/sites/default/files/downloads/AESOP2012_PBL_nabielek.pdf

Ned, U. (2024). Mirt-verkenning ov en wonen regio utrecht - tussenbericht. https://www.programma-uned.nl/relevante+documenten/HandlerDownloadFiles.ashx?idnv=2683664

Nibud. (2023). Autokosten. https://www.nibud.nl/onderwerpen/uitgaven/autokosten/

Ola Adetokunboh MRICS. (2023). Embracing peripheral housing development over urban regeneration: A call for coherent smart thinking (part 1 of 2). https://www.linkedin.com/pulse/embracing-peripheral-housing-development-over-urban-1-ola/

Olsen, J. R., Nicholls, N., Whitley, E., & Mitchell, R. (2024). Association between local amenities, travel behaviours and urban planning: A spatial analysis of a nationwide uk household panel study. *Journal of Transport and Health*. https://doi.org/10.1016/j.jth.2024.101784

Orme, B. (1998). Sample size issues for conjoint analysis studies, sawtooth software technical paper, sequim. *Research Paper Series 98382*.

Plan bureau voor leefomgeving. (2020). Afstanden woon-werkverkeer nemen toe; hoogopgeleide full-time werkende man reist het verst. https://www.pbl.nl/actueel/nieuws/afstanden-woon-werkverkeer-nemen-toe-hoogopgeleide-fulltime-werkende-man-reist-het-verst#:~:text=Vooral%20de%20afstanden%20in%20het,in%201995%20was%20dat%2027%25.

Posad Maxwan and Goudappel Coffeng. (2022). *Ontwerpstudie rijnenburg* (tech. rep.). Gemeente Utrecht.

PosadMaxwan and Goudappel Coffeng. (2020). Rijnenburg - quickscan verstedelijking: Eindrapport.

Rijksoverheid. (n.d.). Welke regels gelden voor mijn elektrische fiets (e-bike, elektrische bakfiets en fatbike)? https://www.rijksoverheid.nl/onderwerpen/fiets/vraag-en-antwoord/welke-regels-gelden-voor-mijn-elektrische-fiets-e-bike#:~:text=De%20speedpedelec%20gaat%20sneller%20en,motorvermogen%20tot%20maximaal%20250%20watt.

Rijksoverheid. (2023). Meer grootschalige woningbouwlocaties nodig om toekomstig woningtekort op te lossen.

Ritchie, H., & Roser, M. (2021). Transport [https://ourworldindata.org/transport]. *Our World in Data*.

Rye, T., & Hrelja, R. (2020). Policies for reducing car traffic and their problematisation. lessons from the mobility strategies of british, dutch, german and swedish cities. *Sustainability*. https://doi.org/10.3390/su12198170

Sarwar, T. (2021). The human consequence of suburban planning. *Friends of the Ashley Schiff Preserve*.

Scheepers, E., Wendel-Vos, W., van Kempen, E., Panis, L. I., Maas, J., Stipdonk, H., Moerman, M., den Hertog, F., Staatsen, B., van Wesemael, P., & Schuit, J. (2013). Personal and environmental characteristics associated with choice of active transport modes versus car use for different trip purposes of trips up to 7.5 kilometers in the netherlands. *PLoS ONE*. https://www.scopus.com/inward/record.uri?eid=2-s2.0-84883508855&doi=10.1371%2fjournal.pone.0073105&partnerID=40&md5=b47cccc6b5e114afc63b2b17a34239d0

Schwanen, T., & Mokhtarian, P. L. (2005). What affects commute mode choice: Neighborhood physical structure or preferences toward neighborhoods? *Journal of Transport Geography*. https://doi.org/https://doi.org/10.1016/j.jtrangeo.2004.11.001

Shabani, A. (2023). Cars remain most used transportation mode in eu despite efforts for green travelling, report shows. https://schengen.news/cars-remain-most-used-transportation-mode-in-eu-despite-efforts-for-green-travelling-report-shows/#:~:text=Cars%20remain%20widely%20used%20throughout,by%20this%20means%20of%20transport.

Shen, Q., Chen, P., & Pan, H. (2016). Factors affecting car ownership and mode choice in rail transit-supported suburbs of a large chinese city. *Transportation Research*.

Simons, D., Clarys, P., De Bourdeaudhuij, I., de Geus, B., Vandelanotte, C., & Deforche, B. (2014). Why do young adults choose different transport modes? a focus group study. *Transport Policy*, *36*. https://doi.org/https://doi.org/10.1016/j.tranpol.2014.08.009

Snellen, D., Bastiaanssen, J., & Hoen, M. '. (2021). Brede welvaart en mobiliteit.

Staat van Utrecht. (2023). Swing viewer - staat van utrecht databank. https://staatvanutrecht.databank.nl/

Studio Bereikbaar. (2022). Mobiliteitsonderzoek rijnenburg: Synthese op basis van bestaand onderzoek.

Studio Bereikbaar. (2023). Verdiepend mobiliteitsonderzoek.

Terlien, E. (2018). How to fit the car in urban areas. https://medium.com/@ericterlien/how-to-fit-the-car-in-urban-areas-a264c26044c6

Ton, D., Bekhor, S., Cats, O., Duives, D. C., Hoogendoorn-Lanser, S., & Hoogendoorn, S. P. (2020). The experienced mode choice set and its determinants: Commuting trips in the netherlands. *Transportation Research Part A: Policy and Practice*, *132*, 744–758. https://doi.org/10.1016/j.tra.2019.12.027

Ton, D., Duives, D. C., Cats, O., Hoogendoorn-Lanser, S., & Hoogendoorn, S. P. (2019). Cycling or walking? determinants of mode choice in the netherlands. *Transportation Research Part A: Policy and Practice*, *123*, 7–23. https://doi.org/https://doi.org/10.1016/j.tra.2018.08.023

Train, K. (2002). *Discrete choice methods with simulation*. Cambridge University Press.

U-OV. (2023). Lijnennetkaart. https://uov-prod.cdn.prismic.io/uov-prod/4cef4515-1d3a-4bd0-b20d-5a9642bee9c6_109+-+WEB+-+Netplan+BRU.PDF

U-OV. (n.d-a). Dienstregeling. https://www.u-ov.info/

U-OV. (n.d-b). Tram 22 gaat vanaf 2 maart vaker rijden. https://www.u-ov.info/meer-qbuzz/nieuws/XlexbRMAACIAH0Pe/tram-22-gaat-vanaf-2-maart-vaker-rijden

Urmi, U. F., Rahman, K., Uddin, M. J., & Hasan, M. N. (2022). The prevalence of active commuting to school and the factors influencing mode choice: A study of university students in a secondary city of bangladesh. *Sustainability (Switzerland)*. https://doi.org/10.3390/su142416949

Utrecht University. (2018). Concentratie van kennis heeft magneetwerking - utrecht science park grootste van nederland. https://www.uu.nl/nieuws/utrecht-science-park-grootste-van-nederland#:~:text=Het%20Utrecht%20Science%20Park%20is,Prinses%20M%C3%A1xima%20Centrum%20voor%20kinderoncologie.

van Exel, N. J. A., de Graaf, G., & Rietveld, P. (2010). "i can do perfectly well without a car!" an exploration of stated preferences for middle-distance travel. https://link.springer.com/article/10.1007/s11116-010-9315-8

van Hagen, M., & Peek, G. (2006). Prettiger reizen met de trein. *Verkeerskunde: vaktijdschrift over verkeer en vervoer, 57(5), 50-55.*

van Nes, R., Hamerslag, R., & Immers, B. H. (n.d.). Design of public transport networks. *Transportation research record*.

van Wee, B., & Handy, S. (2016). Key research themes on urban space, scale, and sustainable urban mobility. *International journal of sustainable transportation*.

Vos, J. D., Le, H. T., & Kroesen, M. (2022). Does commute duration attenuate the effect of travel mode choice on commute satisfaction? *Travel Behaviour and Society*. https://doi.org/10.1016/j.tbs.2022.02.004

Vos, J. D., Waygood, E. O. D., Letarte, L., & Cao, M. (2022). Do frequent satisfying trips by public transport impact its intended use in later life? *Transportation*. https://doi.org/10.1007/s11116-021-10209-0

Wegenwiki. (2024a). A12 (nederland). https://www.wegenwiki.nl/A12_(Nederland)

Wegenwiki. (2024b). Nederland. https://www.wegenwiki.nl/Nederland#:~:text=De%20drukste%20snelweg%20in%20Nederland,met%2010.600%20voertuigen%20per%20dag.

Wikipedia. (2024). Rijnenburg (polder). https://nl.wikipedia.org/wiki/Rijnenburg_(polder)

Wirtz, P., & Ries, G. (1992). The pace of life - reanalysed: Why does walking speed of pedestrians correlate with city size? *Behaviour*.

Woltjer, J. (2014). A global review on peri-urban development and planning. jurnal perencanaan wilayah. *Perencanaan Wilayah*. https://doi.org/10.5614%2Fjpwk.2014.25.1.1

Woninck, S. (2015). Rotterdam nesselande - aantrekkingskracht met een houdbaarheidsdatum. https://studenttheses.uu.nl/bitstream/handle/20.500.12932/20576/S.R.Woninck%20%283616169%29%2c%20Rotterdam%20Nesselande%2c%20aantrekkingskracht%20met%20een%20%20houdbaarheidsdatum.pdf?sequence=2&isAllowed=y

Woonnet Haaglanden. (2024). In welke prijsklasse mag ik reageren? https://www.woonnet-haaglanden.nl/aanbod/informatie-over/huurinkomenstabel#/

Wynia, S. (2021). Https://www.wyniasweek.nl/stop-waarom-wil-iedereen-nederland-volbouwen/. *Wynia's week*.

Xue, F., Yao, E., Cherchi, E., & de Almeida Correia, G. H. (2024). Modeling the joint choice behavior of commuters' travel mode and parking options for private autonomous vehicles. *Transportation Research Part C: Emerging Technologies*. https://doi.org/https://doi.org/10.1016/j.trc.2023.104471

Ye, J., Chen, J., Bai, H., & Yue, Y. (2018). Analyzing transfer commuting attitudes using a market segmentation approach. *Sustainability (Switzerland)*, *10*. https://doi.org/10.3390/su10072194

Zijlstra, T., Bakker, S., & Witte, J.-J. (2022). The widespread car ownership in the netherlands.

A

Possible Lightrail Connection of Rijnenburg

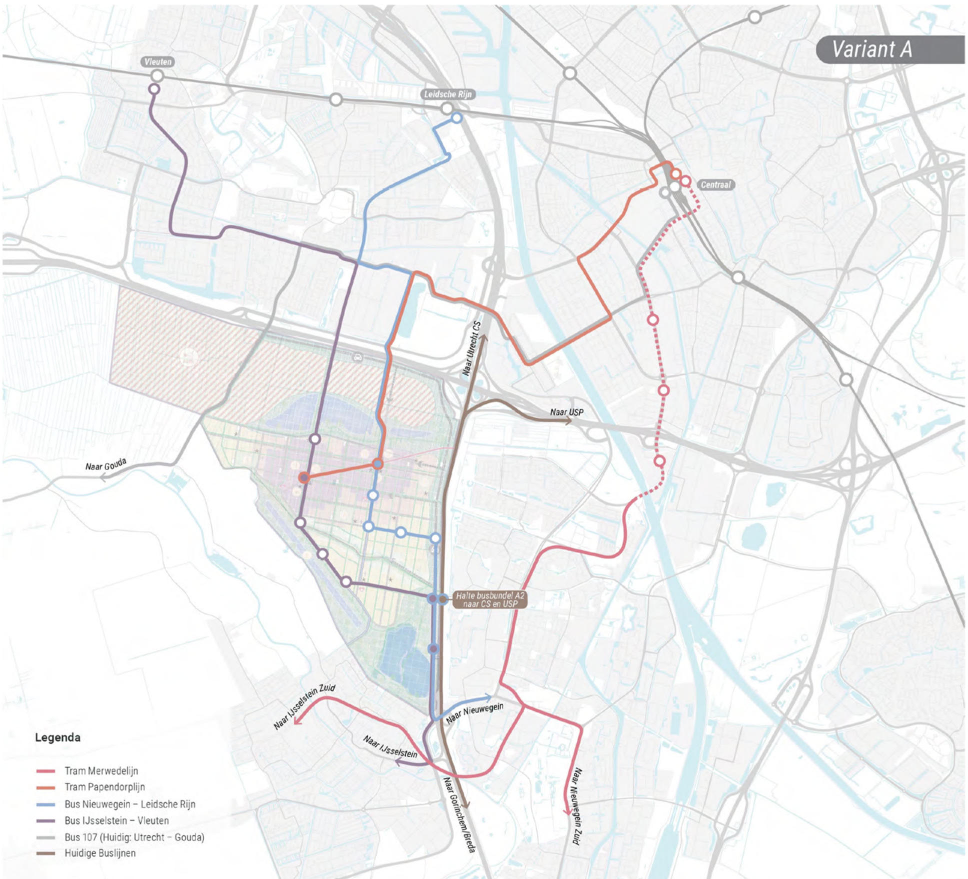


Figure A.1: Potential light rail connection for Rijnenburg (Spoke) (Studio Bereikbaar, 2023)

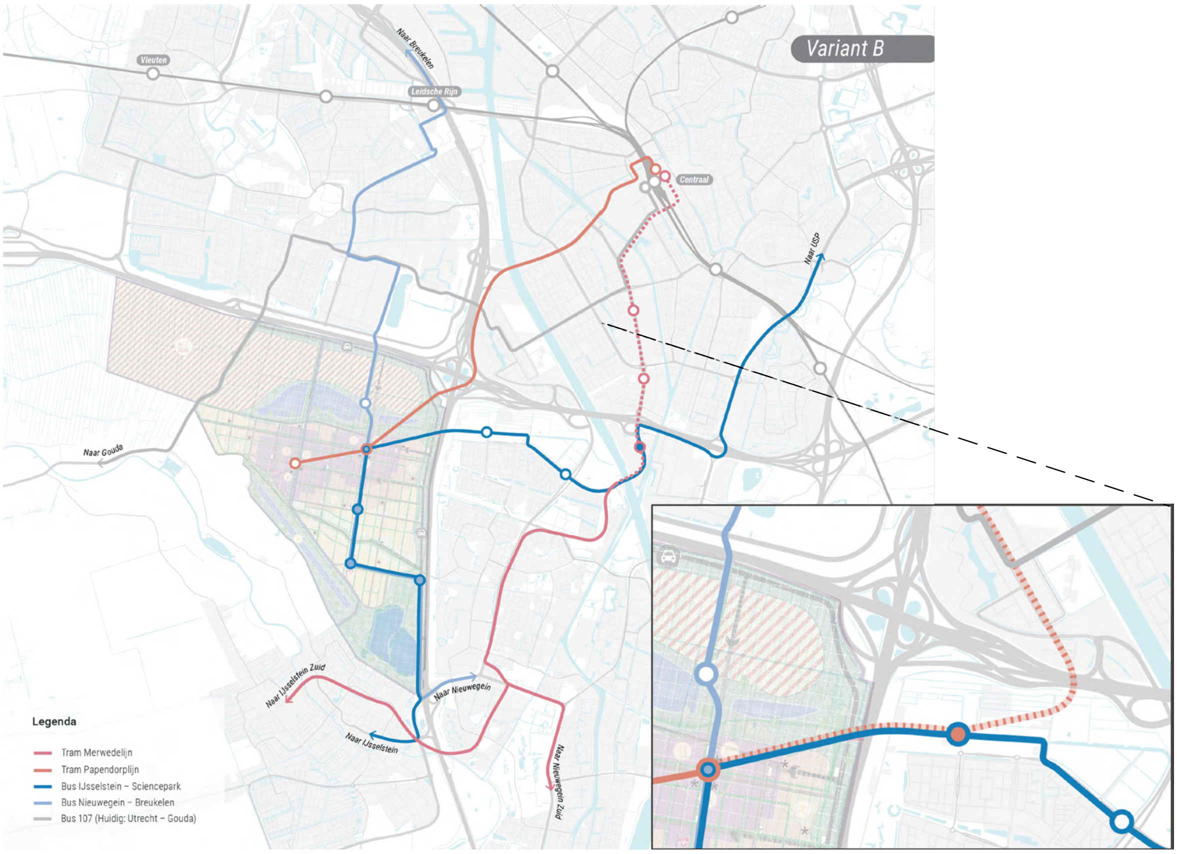


Figure A.2: Potential light rail connection for Rijnenburg (Wheel) (Studio Bereikbaar, 2023)

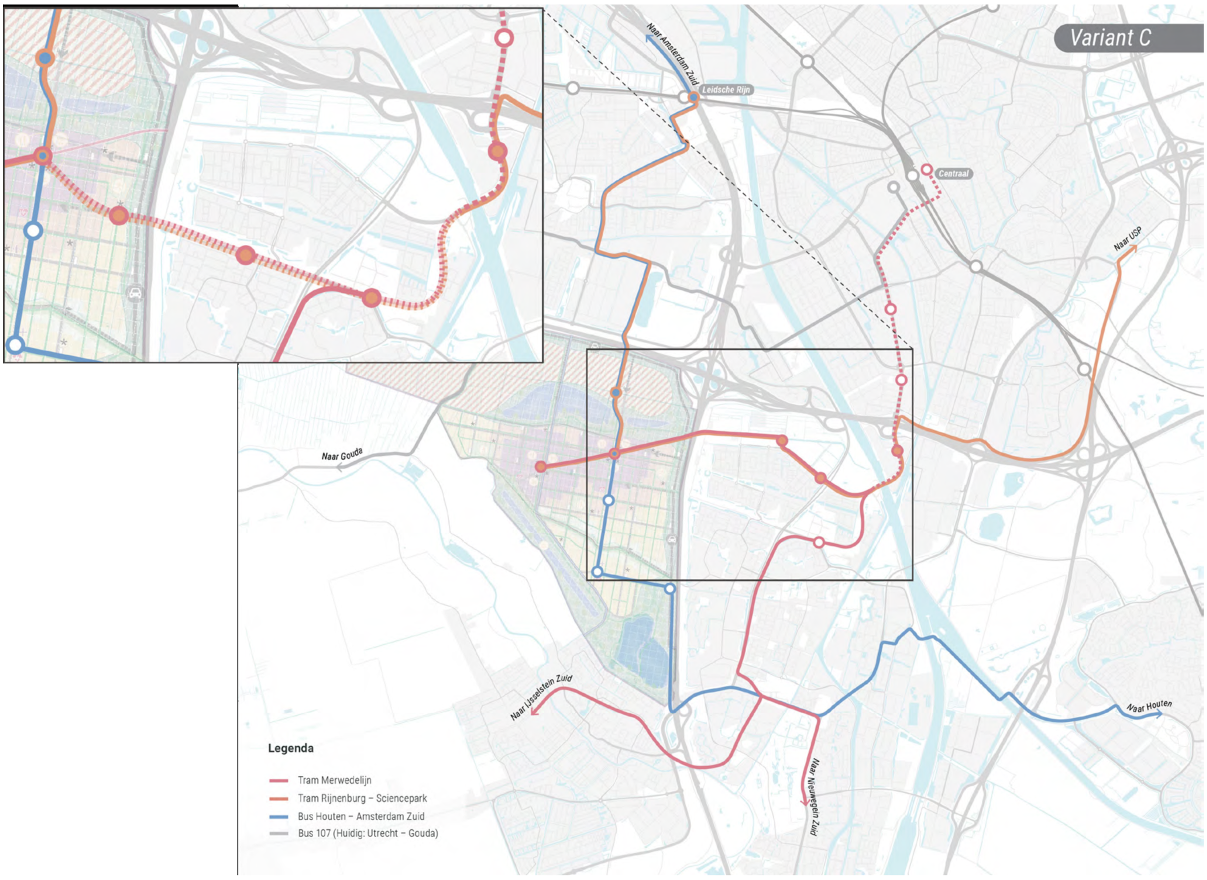


Figure A.3: Potential light rail connection for Rijnenburg (Branching) (Studio Bereikbaar, 2023)

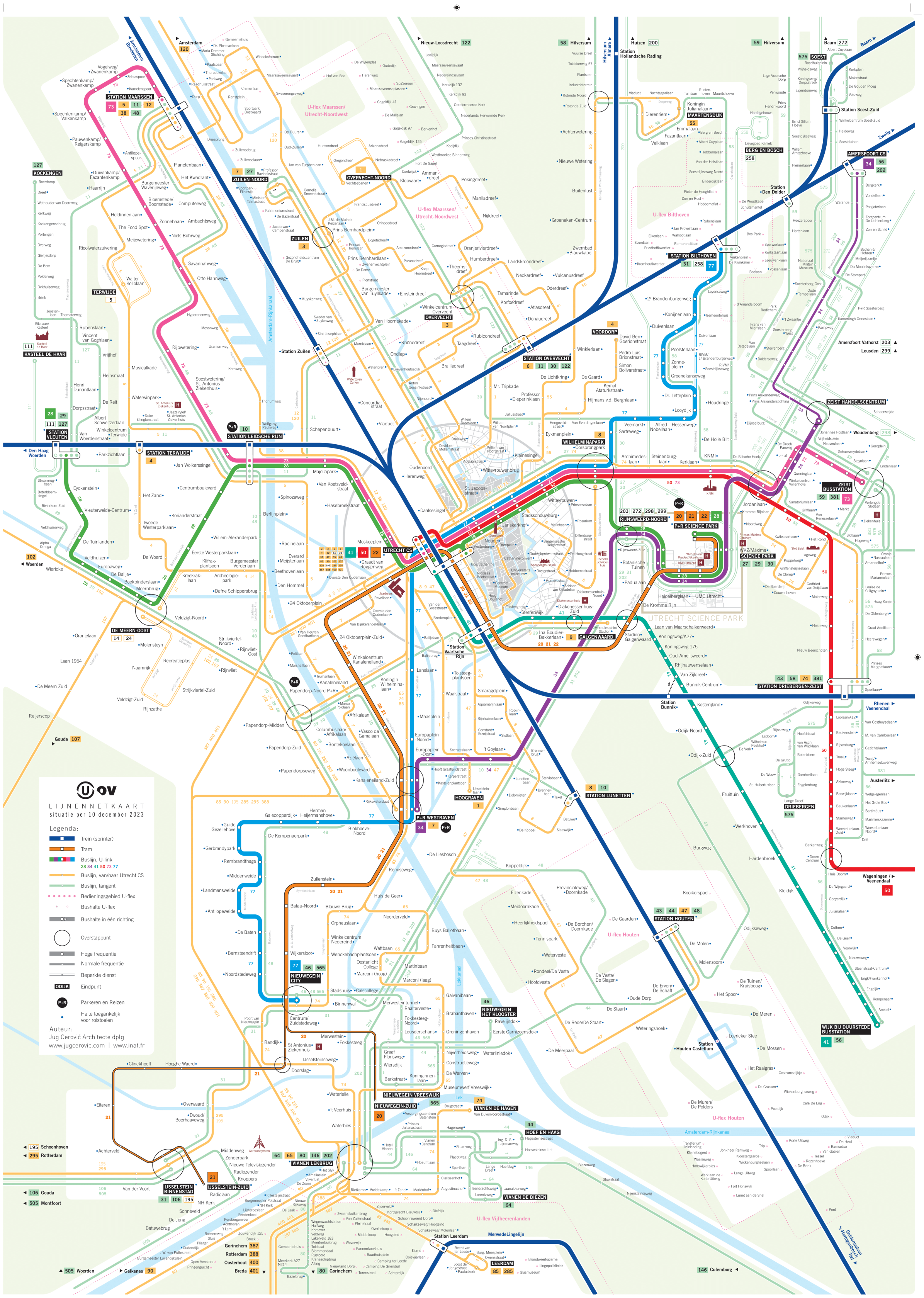


Figure A.4: Current public transportation network of Utrecht (U-OV, 2023)

B

Ngene Syntax & Design

B.1. Choice set generation

```
Design
;alts = PT, car, bike
;rows = 9
;orth = sim
;block = 4
;model:

U(PT) = bPT_access * PT_access[5,7,10] + bPT_IVtime * PT_IVtime[31,41] + bPT_waiting * PT_waiting[3,5,7.5]
+ bTransfer * Transfer[0,1] + bPT_type * PT_type[0,1] /

U(car) = ASC_car + bCar_access * Car_access[0,5,15] + bCarIVtime * Car_IVtime[13,43] + bCar_parking
* Car_parking[10,60] /

U(bike) = ASC_bike + bBike_IVtime * Bike_IVtime[50,38]

    $
```

B.2. Ngene output

Table B.1: Ngene output - Final design

Choice situation	pt.pt_access	pt.pt_ivtime	pt.pt_waiting	pt.transfer	pt.pt_type	car.car_access	car.car_ivtime	car.car_parking	bike.bike_ivtime	Block
1	7	31	5	0	0	15	13	10	50	1
2	10	31	7.5	0	0	5	43	10	38	3
3	10	41	7.5	1	1	5	43	10	50	3
4	7	41	5	1	0	15	43	10	50	2
5	5	41	5	0	0	5	43	60	38	4
6	7	31	3	1	0	0	13	60	50	4
7	7	31	3	0	1	0	43	60	50	2
8	5	41	5	0	1	5	13	60	50	3
9	10	31	3	1	1	15	13	10	38	4
10	5	41	7.5	0	1	0	13	10	38	2
11	5	41	7.5	1	0	0	13	60	38	1
12	10	31	3	1	1	15	43	60	38	1
13	10	31	7.5	0	0	0	13	10	50	1
14	5	31	3	0	0	15	43	10	38	3
15	5	41	3	1	1	15	43	10	50	3
16	10	41	7.5	1	0	0	43	10	50	2
17	7	41	7.5	0	0	15	43	60	38	4
18	10	31	5	1	0	5	13	60	50	4
19	10	31	5	0	1	5	43	60	50	2
20	7	41	7.5	0	1	15	13	60	50	3
21	5	31	5	1	1	0	13	10	38	4
22	7	41	3	0	1	5	13	10	38	2
23	7	41	3	1	0	5	13	60	38	1
24	5	31	5	1	1	0	43	60	38	1
25	5	31	3	0	0	5	13	10	50	1
26	7	31	5	0	0	0	43	10	38	3
27	7	41	5	1	1	0	43	10	50	3
28	5	41	3	1	0	5	43	10	50	2
29	10	41	3	0	0	0	43	60	38	4
30	5	31	7.5	1	0	15	13	60	50	4
31	5	31	7.5	0	1	15	43	60	50	2
32	10	41	3	0	1	0	13	60	50	3
33	7	31	7.5	1	1	5	13	10	38	4
34	10	41	5	0	1	15	13	10	38	2
35	10	41	5	1	0	15	13	60	38	1
36	7	31	7.5	1	1	5	43	60	38	1

C Survey

English

Dear respondent,

You are kindly invited to participate in the graduation research on urban mobility. This research is conducted by Kim Pham, a Civil Engineering student at TU Delft.

The aim of this research is to identify the characteristics and preferences of commuting behavior. Based on this, design strategies can be derived for transportation infrastructure in new development areas. This survey will take approximately 8 minutes to complete, and the collected data will be used for analysis and a written master's thesis. You will be asked to provide some socio-demographic information and make various choices about how you want to travel between your home and work.

As with any online activity, there is a risk of data breaches. This risk is minimized by storing your data on a secure server in accordance with Dutch GDPR regulations, anonymizing the data before analyzing and reporting the results. Furthermore, your personal data will be deleted at the end of the study (August 2024). Your participation in this research is completely voluntary, and you may withdraw at any time without giving a reason. You are free to choose not to answer questions. If you have any questions or concerns about this research, you can contact me directly at D.K.A.Pham@student.tudelft.nl.

By proceeding to the next section, you agree to the above opening statement and consent to the collection and storage of your personal data for the duration of the study.

Best regards,
Kim Pham

English

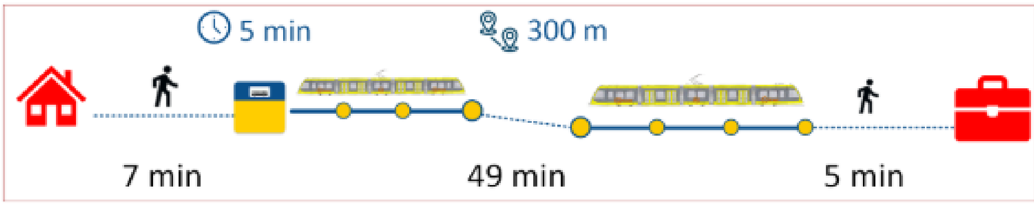
The first part will focus on how you would like to travel from your home to your workplace. The route between Rijnenburg and the Utrecht Science Park (USP) will be used as an example.

Imagine you live in Rijnenburg and travel to your work at the USP every morning. Rijnenburg is a busy residential area outside Utrecht, surrounded by two major freeways, the A2 and A12, making it very convenient for driving car to the USP via the Utrecht Ring Road. The distance for cycling between Rijnenburg and USP is approximately 15 km, which is quite a distance to cycle. There is no direct train connection, but there may be tram or bus lines available.

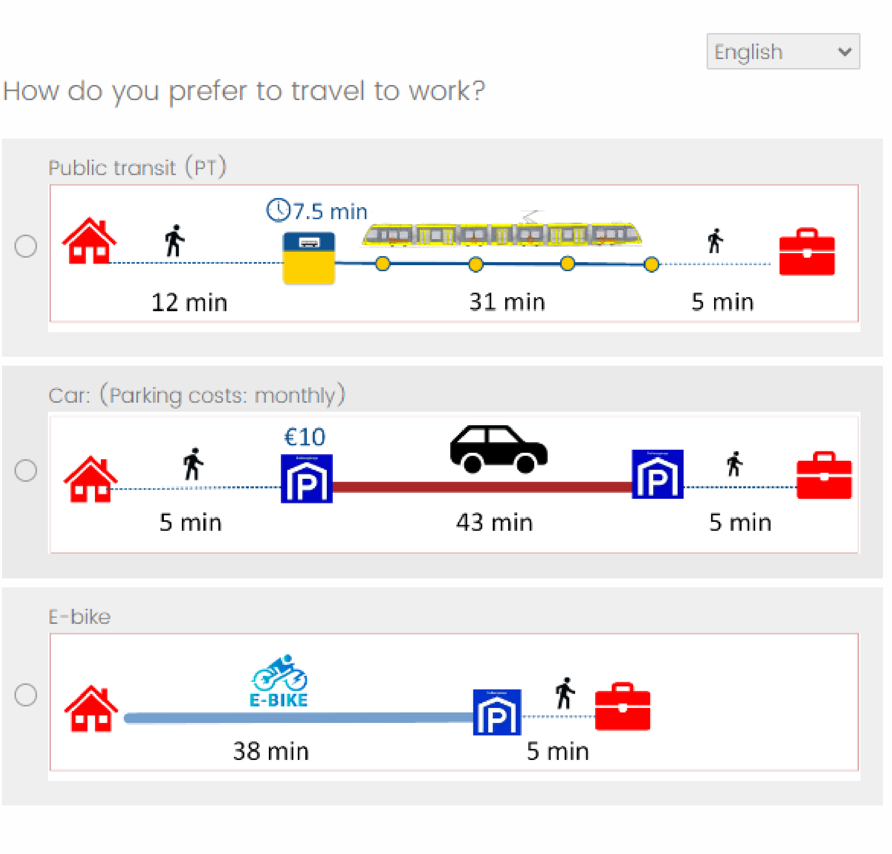
On the map below, you can see where your home, workplace, and the city center with the train station are located. You will also find the railway lines and highways marked on the map.

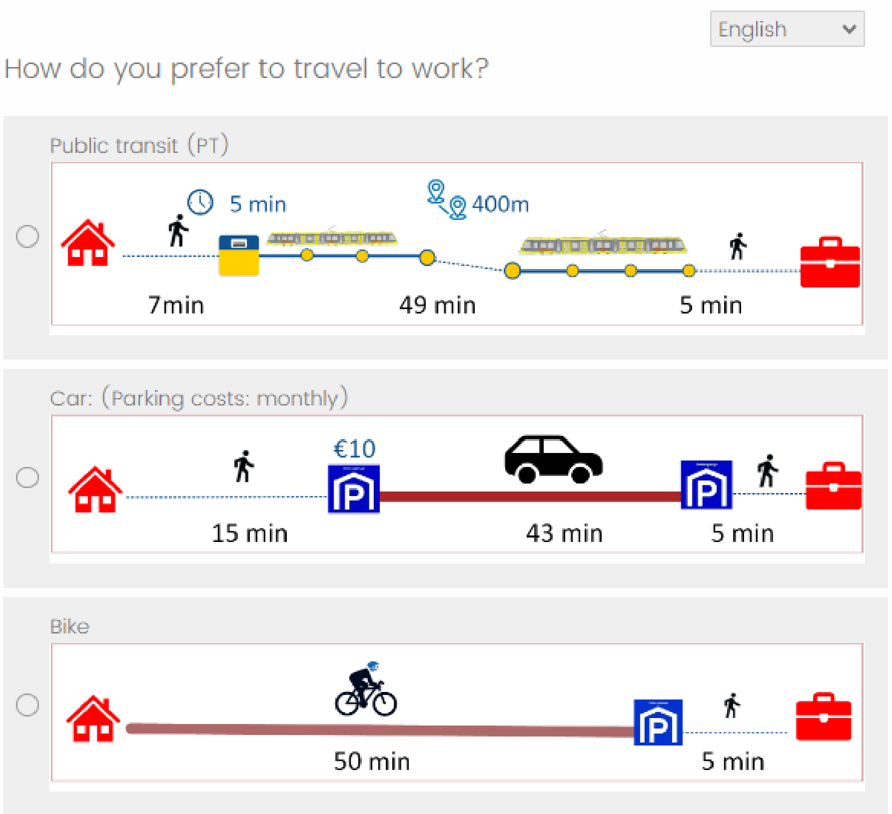
You will be presented with several scenarios, and for each scenario, you will be asked to choose one mode of transportation: tram, bus, car, bicycle, or electric bike (e-bike). We assume that the travel fare of the trip is covered by your employer, maintenance costs and motor vehicle tax for the car option will be €300 per month in total.

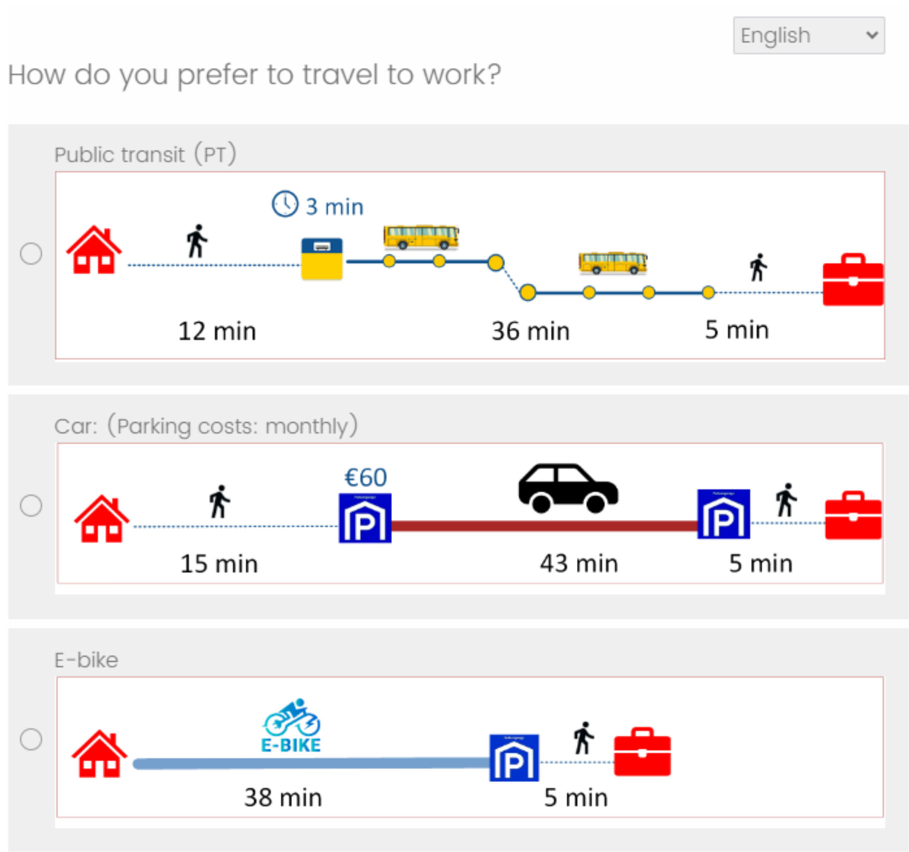
Each situation indicates the time it takes to walk, wait, and transfer. Additionally, parking costs and types of vehicles are also provided. Below is an example where you first walk 7 minutes from home to the tram stop, then wait for 5 minutes for your tram, followed by a 49-minute tram ride with transferdisatcne of 300 metre. Upon arrival, you walk 5 minutes from the tram stop to your workplace.

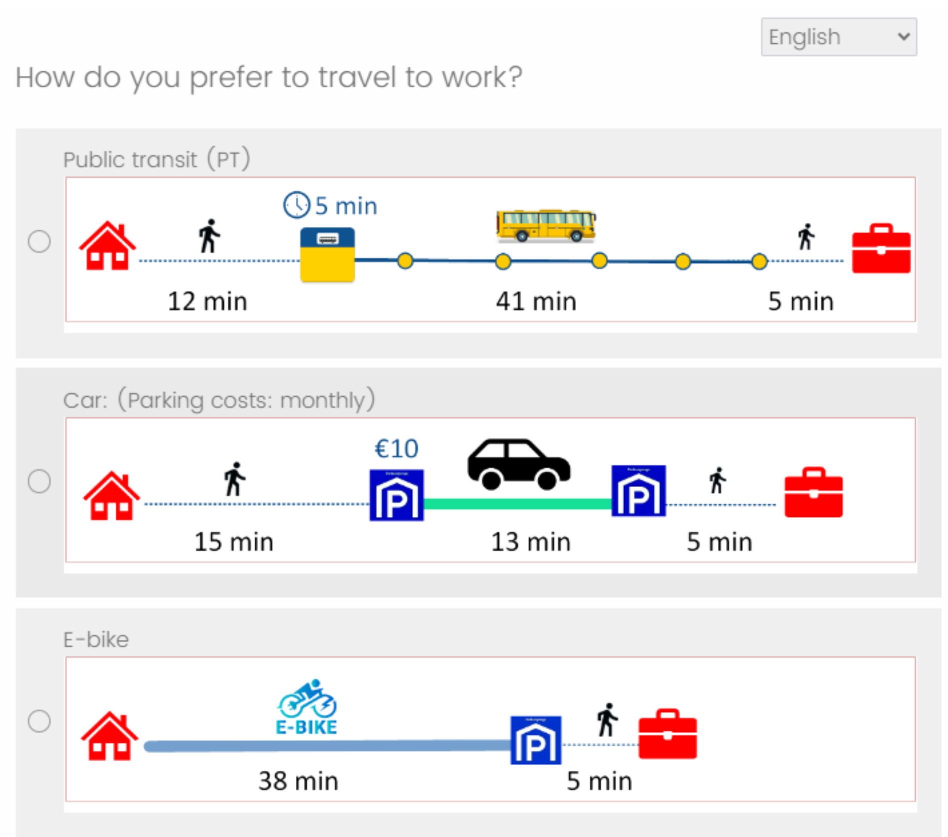


Carefully consider each situation and choose the mode of transportation you would select!


English
How do you prefer to travel to work?
Public transit (PT)
7.5 min
12 min
31 min
5 min
Car: (Parking costs: monthly)
€10
5 min
43 min
5 min
E-bike
E-BIKE
38 min
5 min


English
How do you prefer to travel to work?
Public transit (PT)
5 min
400m
7min
49 min
5 min
Car: (Parking costs: monthly)
€10
15 min
43 min
5 min
Bike
50 min
5 min


English
How do you prefer to travel to work?
Public transit (PT)
3 min
12 min
36 min
5 min
Car: (Parking costs: monthly)
€60
15 min
43 min
5 min
E-bike
E-BIKE
38 min
5 min


English
How do you prefer to travel to work?
Public transit (PT)
5 min
12 min
41 min
5 min
Car: (Parking costs: monthly)
€10
15 min
13 min
5 min
E-bike
E-BIKE
38 min
5 min

English

Given the location and the distance to your workplace, would you consider living in Rijnenburg?

- ◯ Yes
- ◯ No
- ◯ Other,

English

How old are you?

- ◯ Younger than 18
- ◯ 18 to 24
- ◯ 25 to 34
- ◯ 35 to 44
- ◯ 45 to 54
- ◯ 55 to 64
- ◯ Older than 64
- ◯ I prefer not to answer

What is the composition of your household?

- ◯ Single without children
- ◯ Single with children
- ◯ Living together/married without children
- ◯ Living together/married with children
- ◯ Other/ I prefer not to answer

What is your highest educational qualification?

- ◯ No (formal) education
- ◯ High school
- ◯ Vocational Education and Training (VET)
- ◯ HBO/WO Bachelor's degree
- ◯ HBO/WO Master/PhD
- ◯ Other/ I prefer not to answer

What is your employment status?

- ◯ Student
- ◯ Job seeker
- ◯ Full time employed
- ◯ Part time employed
- ◯ Self employed
- ◯ Retired
- ◯ Other

What is your monthly gross income?

- ◯ Lower than €833
- ◯ Between €833 and €2.310
- ◯ Between €2.310 and €3.975
- ◯ Between €3.975 and €4.798
- ◯ Higher than €4.798
- ◯ I prefer not to answer

English

Do you have any physical or health condition affecting your ability to travel?

- ◯ Yes
- ◯ No
- ◯ I prefer not to answer

Do you have car driving license?

- ◯ Yes
- ◯ No
- ◯ I prefer not to answer

What modes of transportation do you have? (Multiple choices possible)

- ☐ I own a car (shared with my partner)
- ☐ I have a leased car through my work
- ☐ I own a bike
- ☐ I own an electronic (e-bike)
- ☐ I have a scooter and/or a moped
- ☐ I have a public transportation subscription
- ☐ Other/ I prefer not to answer

Do you receive travel reimbursement from your work?

- [ ] No
- [ ] Yes, mileage reimbursement
- [ ] Yes, a leased car
- [ ] Yes, public transport subscription by employer
- [ ] Yes, other form of reimbursement
- [ ] Other/ I prefer not to answer

How do you usually travel to work?

- ○ By public transportation (train, tram, bus, metro)
- ○ By my own car
- ○ Cycling
- ○ Other

English

Thank you for taking the time to participate in my research! I appreciate your effort and would love to receive your feedback.

Where did you find this survey?

- ○ Information session U Ned
- ○ U Ned
- ○ Facebook / LinkedIn
- ○ Utrecht Science Park
- ○ Utrecht University
- ○ TU Delft
- ○ Other,

D

Descriptive Statistics

D.1. Characteristics differences

Table D.1: Chi-square statistical test for characteristics differences between collected sample and population of Province Utrecht.

Variables	Observed frequency O_i	Percentage population	Expected frequency E_i	χ^2 $(O_i - E_i)^2/E_i$
Age group				
15 - 24 yr	93	12.6	25.2	182.414
25 - 44 yr	61	27	54.0	0.907
45 - 64 yr	38	25.9	51.8	3.676
Older than 64 yr	5	17.8	35.6	26.302
Total	**200**			**213.300**
Household composition				
Single without children	74	39.2	78.4	0.247
Couple without children	54	27.5	55.0	0.018
Presence of children	33	33.3	66.6	16.951
Total	**200**			**17.216**
Educational levels				
High school	43	48	96.0	29.260
Bachelor's degree	91	52	104.0	20.346
Master/PhD	59			
Total	**200**			**49.607**
Employment status				
Student	95	20.9	41.8	67.709
Employment (parttime & fulltime)	83	53	106.0	4.991
Total	**200**			**72.700**

D.2. Difference in mode choice

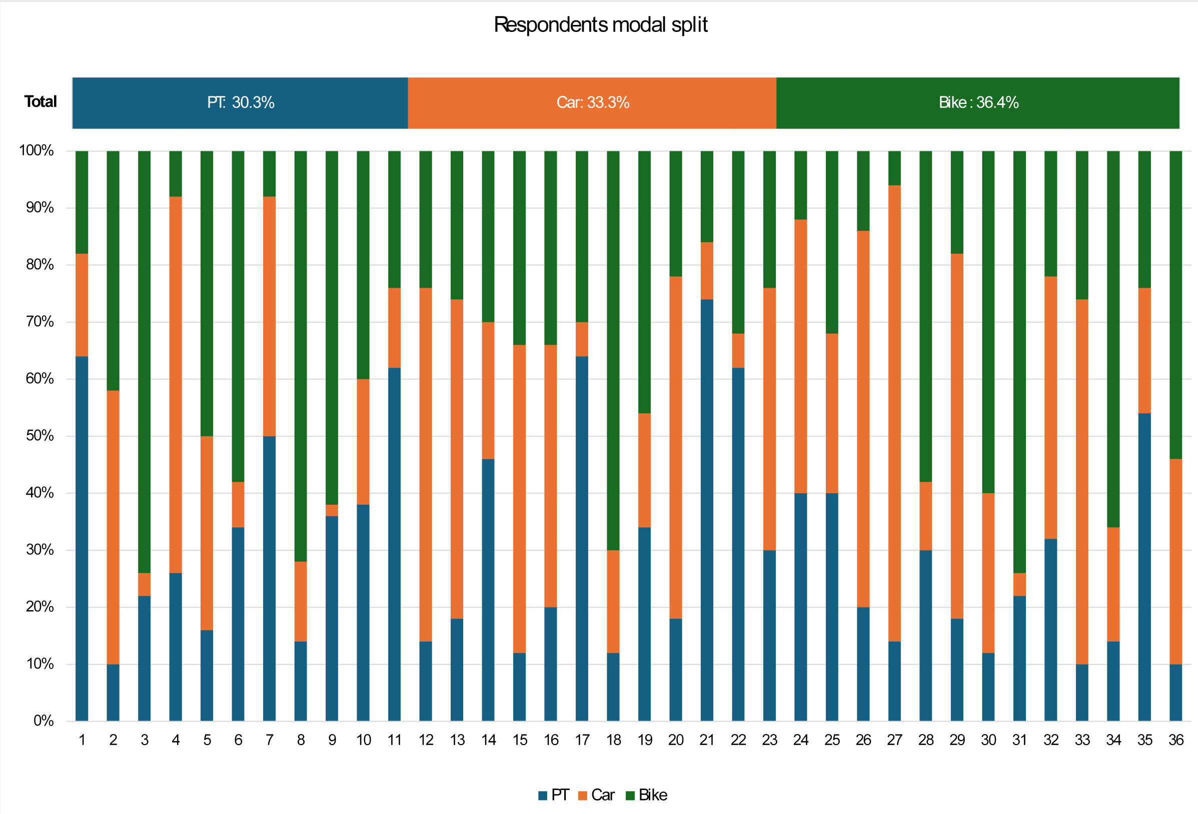


Figure D.1: Modal preference for total and for each choice situation

D.3. Chi-square test for differences in mode preferences

Table D.2: Chi-square test for mode choice differences in age groups

	Mode	15-24 yr	25-44 yr	45-64 yr	Total
Observed Frequencies	PT	246	171	92	509
	Car	334	163	93	590
	Bike	256	212	155	623
Expected Frequencies	PT	247.11	161.39	100.5	509
	Car	286.43	187.07	116.49	590
	Bike	302.46	197.54	123.01	623
Chi-square Statistic	PT	0	0.57	0.72	1.29
	Car	7.9	3.1	4.74	15.74
	Bike	7.14	1.06	8.32	16.52

Table D.3: Chi-square test for mode choice differences household composition

	Mode	Single	Couple	Presence of Children	Total
Observed Frequencies	PT	227	143	74	444
	Car	224	160	88	472
	Bike	213	181	134	528
Expected Frequencies	PT	204.17	148.82	91.01	444
	Car	217.04	158.2	96.75	472
	Bike	242.79	176.98	108.23	528
Chi-square Statistic	PT	2.55	0.23	3.18	5.96
	Car	0.22	0.02	0.79	1.03
	Bike	3.66	0.09	6.14	9.89

Table D.4: Chi-square test for mode choice differences in educational levels

	Mode	Low	Bachelor	Master	Total
Observed Frequencies	PT	106	259	161	526
	Car	139	268	169	576
	Bike	141	290	198	629
Expected Frequencies	PT	117.29	248.26	160.44	526
	Car	128.44	271.86	175.69	576
	Bike	140.26	296.88	191.86	629
Chi-square Statistic	PT	1.09	0.46	0	1.55
	Car	0.87	0.05	0.25	1.17
	Bike	0	0.16	0.2	0.36

Table D.5: Chi-square test for mode choice differences in age groups

	Mode	Lower than €2,310	€2,310 - €3,975	Higher than €3,975	Total
Observed Frequencies	PT	308	90	104	502
	Car	347	109	91	547
	Bike	963	267	412	1642
Expected Frequencies	PT	301.83	86.93	113.23	502
	Car	328.89	94.72	123.38	547
	Bike	987.27	284.34	370.38	1642
Chi-square Statistic	PT	0.13	0.11	0.75	0.99
	Car	1	2.15	8.5	11.65
	Bike	0.6	1.06	4.68	6.34

Table D.6: Chi-square test for mode choice differences in employment status

Employment stattus	Mode	Student	Employment	Total
Observed Frequencies	PT	262	212	474
	Car	312	213	525
	Bike	281	298	579
Expected Frequencies	PT	256.83	217.17	474
	Car	284.46	240.54	525
	Bike	313.72	265.28	579
Chi-square Statistic	PT	0.1	0.12	0.22
	Car	2.67	3.15	5.82
	Bike	3.41	4.04	7.45

Table D.7: Chi-square test for mode choice differences in mobility availability

Income	Mode	Disability	Driving License	PT Subscription
Observed Frequencies	PT	23	398	369
	Car	29	468	353
	Bike	11	538	421
Expected Frequencies	PT	19.18	427.37	347.92
	Car	20.5	456.79	371.87
	Bike	23.33	519.84	423.21
Chi-square Statistic	PT	0.76	2.02	1.28
	Car	3.52	0.28	0.96
	Bike	6.52	0.63	0.01

Income	Mode	Car Ownership	Bike Ownership	E-bike Ownership	Total
Observed Frequencies	PT	210	461	369	1830
	Car	277	476	353	1956
	Bike	260	575	421	2226
Expected Frequencies	PT	227.38	460.24	347.92	1830.01
	Car	243.04	491.93	371.87	1956
	Bike	276.58	559.83	423.21	2226
Chi-square Statistic	PT	1.33	0	1.28	6.67
	Car	4.75	0.52	0.96	10.99
	Bike	0.99	0.41	0.01	8.57

E

Modelling

E.1. Model estimation

MNL model estimation

```
# %% Import packages
import pandas as pd
%matplotlib inline
import numpy as np
import matplotlib.pyplot as plt
import biogeme.database as db
import biogeme.biogeme as bio
from biogeme import models
from biogeme.expressions import Beta, Variable
from sklearn.model_selection import train_test_split
from sklearn.model_selection import StratifiedShuffleSplit

# Read the data
data = pd.read_csv('train_data.csv')
df = pd.DataFrame(data)
train_data = db.Database("train_data", df)
globals().update(train_data.variables)


# Define variables

#Socio characteristics
Age_25_44 = Variable('Age(25-44)')
Age_45_64 = Variable('Age(45-64)')
Age_65 = Variable('Age(65)')
Single = Variable('Single')
Couple = Variable('Couple')
Bachelor = Variable('Bachelor')
High_Edu = Variable('High_Edu')
Employment = Variable('Employment')
Mid_income = Variable('Mid-income')
High_income = Variable('High-income')
Disability = Variable('Disability')
Driving_license = Variable('Driving_license')
Availability_PT = Variable('Availability_PT')
Availability_car = Variable('Availability_car')
```

```
Availability_bike = Variable('Availability_bike')
Frequent_car = Variable('Frequent_car')
Frequent_bike = Variable('Frequent_bike')
Reimbursement = Variable('Reimbursement')
Reimbursement_PT = Variable('Reimbursement_PT')
Reimbursement_mileage = Variable('Reimbursement_mileage')
Reimbursement_leased = Variable('Reimbursement_leased')


#Trip characteristics
PT_access = Variable('PT_access')
PT_waiting = Variable('PT_waiting')
PT_IVtime = Variable('PT_IVtime')
PT_TTtime = Variable('PT_TTtime')
Transfer = Variable('Transfer')
Transfer_time = Variable('Transfer_time')
Transfer_distance = Variable('Transfer_distance')
PT_type = Variable('PT_type')


Car_access = Variable('Car_access')
Car_IVtime = Variable('Car_IVtime')
Car_parking = Variable('Car_parking')

Bike_IVtime = Variable('Bike_IVtime')

av = Variable('Av')
choice = Variable('Choice')

# Model Parameters
# 1 Name for report. Typically, same as variable
# 2 Starting value
# 3 Lower bound
# 4 Upper bound
# 5 0: estimate the parameter, 1: keep it fixed
ASC_PT = Beta('ASC_PT',0,None,None,1) # fixed to 0
ASC_Car = Beta('ASC_Car',0,None,None,0)
ASC_Bike = Beta('ASC_Bike',0,None,None,0)

#Generic attributes
B_Age_25_44 = Beta('B_Age_25_44', 0, None, None, 0)
B_Age_45_64 = Beta('B_Age_45_64', 0, None, None, 1)
B_Age_65 = Beta('B_Age_65', 0, None, None, 1)
B_Single = Beta('B_Single', 0, None, None, 1)
B_Couple = Beta('B_Couple', 0, None, None, 1)
B_Bachelor = Beta('B_Bachelor', 0, None, None, 1)
B_High_Edu = Beta('B_High_Edu', 0, None, None, 1)
B_Employment = Beta('B_Employment', 0, None, None, 1)
B_Mid_income = Beta('B_Mid_income', 0, None, None, 1)
B_High_income = Beta('B_High_income', 0, None, None, 1)
B_Disability = Beta('B_Disability', 0, None, None, 0)
B_Driving_license = Beta('B_Driving_license', 0, None, None, 0)
B_Availability_PT = Beta('B_Availability_PT', 0, None, None, 1)
B_Availability_car = Beta('B_Availability_car', 0, None, None, 1)
B_Availability_bike = Beta('B_Availability_bike', 0, None, None, 1)
B_Frequent_car = Beta('B_Frequent_car', 0, None, None, 1)
```

```
B_Frequent_bike = Beta('B_Frequent_bike', 0, None, None, 1)
B_Reimbursement = Beta('B_Reimbursement', 0, None, None, 1)
B_Reimbursement_PT = Beta('B_Reimbursement_PT', 0, None, None, 1)
B_Reimbursement_mileage = Beta('B_Reimbursement_mileage', 0, None, None, 1)
B_Reimbursement_leased = Beta('B_Reimbursement_leased', 0, None, None, 1)

#Specific attributes
B_PT_Age_25_44 = Beta('B_PT_Age_25_44  ', 0, None, None, 1)
B_Car_Age_25_44 = Beta('B_Car_Age_25_44', 0, None, None, 0)
B_Bike_Age_25_44 = Beta('B_Bike_Age_25_44', 0, None, None,1)

B_PT_Age_45_64 = Beta('B_PT_Age_45_64', 0, None, None, 1)
B_Car_Age_45_64 = Beta('B_Car_Age_45_64', 0, None, None, 1)
B_Bike_Age_45_64 = Beta('B_Bike_Age_45_64', 0, None, None, 1)

B_PT_Age_65 = Beta('B_PT_Age_65', 0, None, None, 1)
B_Car_Age_65 = Beta('B_Car_Age_65', 0, None, None, 1)
B_Bike_Age_65 = Beta('B_Bike_Age_65', 0, None, None, 1)

B_PT_Single = Beta('B_PT_Single', 0, None, None, 1)
B_Car_Single = Beta('B_Car_Single', 0, None, None, 1)
B_Bike_Single = Beta('B_Bike_Single', 0, None, None, 1)

B_PT_Couple = Beta('B_PT_Couple', 0, None, None, 1)
B_Car_Couple = Beta('B_Car_Couple', 0, None, None, 1)
B_Bike_Couple = Beta('B_Bike_Couple', 0, None, None, 1)

B_PT_Bachelor = Beta('B_PT_Bachelor', 0, None, None, 1)
B_Car_Bachelor = Beta('B_Car_Bachelor', 0, None, None, 1)
B_Bike_Bachelor = Beta('B_Bike_Bachelor', 0, None, None, 1)

B_PT_High_Edu = Beta('B_PT_High_Edu', 0, None, None, 1)
B_Car_High_Edu = Beta('B_Car_High_Edu', 0, None, None, 1)
B_Bike_High_Edu = Beta('B_Bike_High_Edu', 0, None, None, 1)

B_PT_Employment = Beta('B_PT_Employment', 0, None, None, 1)
B_Car_Employment = Beta('B_Car_Employment', 0, None, None, 1)
B_Bike_Employment = Beta('B_Bike_Employment', 0, None, None, 1)

B_PT_Mid_income = Beta('B_PT_Mid_income', 0, None, None, 0)
B_Car_Mid_income = Beta('B_Car_Mid_income', 0, None, None, 1)
B_Bike_Mid_income = Beta('B_Bike_Mid_income', 0, None, None, 1)

B_PT_High_income = Beta('B_PT_High_income', 0, None, None, 0)
B_Car_High_income = Beta('B_Car_High_income', 0, None, None, 1)
B_Bike_High_income = Beta('B_Bike_High_income', 0, None, None, 1)

B_PT_Disability = Beta('B_PT_Disability', 0, None, None, 1)
B_Car_Disability = Beta('B_Car_Disability', 0, None, None, 0)
B_Bike_Disability = Beta('B_Bike_Disability', 0, None, None, 1)

B_PT_Driving_license = Beta('B_PT_Driving_license', 0, None, None, 1)
B_Car_Driving_license = Beta('B_Car_Driving_license', 0, None, None, 0)
B_Bike_Driving_license = Beta('B_Bike_Driving_license', 0, None, None, 1)

B_PT_Availability_PT = Beta('B_PT_Availability_PT', 0, None, None, 1)
```

```
B_Car_Availability_PT = Beta('B_Car_Availability_PT', 0, None, None, 1)
B_Bike_Availability_PT = Beta('B_Bike_Availability_PT', 0, None, None, 0)

B_PT_Availability_car = Beta('B_PT_Availability_car', 0, None, None, 1)
B_Car_Availability_car = Beta('B_Car_Availability_car', 0, None, None, 1)
B_Bike_Availability_car = Beta('B_Bike_Availability_car', 0, None, None, 0)

B_PT_Availability_bike = Beta('B_PT_Availability_bike', 0, None, None, 0)
B_Car_Availability_bike = Beta('B_Car_Availability_bike', 0, None, None, 0)
B_Bike_Availability_bike = Beta('B_Bike_Availability_bike', 0, None, None, 1)

B_PT_Frequent_car = Beta('B_PT_Frequent_car', 0, None, None, 1)
B_Car_Frequent_car = Beta('B_Car_Frequent_car', 0, None, None, 1)
B_Bike_Frequent_car = Beta('B_Bike_Frequent_car', 0, None, None, 1)

B_PT_Frequent_bike = Beta('B_PT_Frequent_bike', 0, None, None, 1)
B_Car_Frequent_bike = Beta('B_Car_Frequent_bike', 0, None, None, 1)
B_Bike_Frequent_bike = Beta('B_Bike_Frequent_bike', 0, None, None, 1)

B_PT_Reimbursement = Beta('B_PT_Reimbursement', 0, None, None, 1)
B_Car_Reimbursement = Beta('B_Car_Reimbursement', 0, None, None, 1)
B_Bike_Reimbursement = Beta('B_Bike_Reimbursement', 0, None, None, 1)

B_PT_Reimbursement_PT = Beta('B_PT_Reimbursement_PT', 0, None, None, 0)
B_Car_Reimbursement_PT = Beta('B_Car_Reimbursement_PT', 0, None, None, 1)
B_Bike_Reimbursement_PT = Beta('B_Bike_Reimbursement_PT', 0, None, None, 1)

B_PT_Reimbursement_mileage = Beta('B_PT_Reimbursement_mileage', 0, None, None, 1)
B_Car_Reimbursement_mileage = Beta('B_Car_Reimbursement_mileage', 0, None, None, 1)
B_Bike_Reimbursement_mileage = Beta('B_Bike_Reimbursement_mileage', 0, None, None, 1)

B_PT_Reimbursement_leased = Beta('B_PT_Reimbursement_leased', 0, None, None, 0)
B_Car_Reimbursement_leased = Beta('B_Car_Reimbursement_leased', 0, None, None, 1)
B_Bike_Reimbursement_leased = Beta('B_Bike_Reimbursement_leased', 0, None, None, 1)

# Trip characteristics
B_PT_access = Beta('B_PT_access', 0, None, None, 0)
B_waiting = Beta('B_waiting', 0, None, None, 1)
B_PT_IVtime = Beta('B_PT_IVtime', 0, None, None, 0)
B_transfer = Beta('B_transfer', 0, None, None, 0)
B_transfer_time = Beta('B_transfer_time', 0, None, None, 1)
B_transfer_distance = Beta('B_transfer_distance', 0, None, None, 1)
B_PT_TTtime = Beta('B_PT_TTtime', 0, None, None, 1)
B_PT_type = Beta('B_PT_type', 0, None, None, 1)

B_car_access = Beta('B_car_access', 0, None, None, 0)
B_car_IVtime = Beta('B_car_IVtime', 0, None, None, 0)
B_car_parking = Beta('B_car_parking', 0, None, None, 0)

B_bike_IVtime = Beta('B_bike_IVtime', 0, None, None, 0)

#Interaction

# Utility functions
V_socio = ( B_Age_25_44 * Age_25_44
        + B_Age_45_64 * Age_45_64
```

```
          + B_Age_65 * Age_65
          + B_Single * Single
          + B_Couple * Couple
          + B_Bachelor * Bachelor
          + B_High_Edu * High_Edu
          + B_Employment * Employment
          + B_Mid_income * Mid_income
          + B_High_income * High_income
          + B_Disability * Disability
          + B_Driving_license * Driving_license
          + B_Availability_PT * Availability_PT
          + B_Availability_car * Availability_car
          + B_Availability_bike * Availability_bike
          + B_Frequent_car * Frequent_car
          + B_Frequent_bike * Frequent_bike
          + B_Reimbursement * Reimbursement
          + B_Reimbursement_PT * Reimbursement_PT
          + B_Reimbursement_mileage * Reimbursement_mileage
          + B_Reimbursement_leased * Reimbursement_leased)

V_PT_socio = (B_PT_Age_25_44 * Age_25_44
         + B_PT_Age_45_64 * Age_45_64
         + B_PT_Age_65 * Age_65
         + B_PT_Single * Single
         + B_PT_Couple * Couple
         + B_PT_Bachelor * Bachelor
         + B_PT_High_Edu * High_Edu
         + B_PT_Employment * Employment
         + B_PT_Mid_income * Mid_income
         + B_PT_High_income * High_income
         + B_PT_Disability * Disability
         + B_PT_Driving_license * Driving_license
         + B_PT_Availability_PT * Availability_PT
         + B_PT_Availability_car * Availability_car
         + B_PT_Availability_bike * Availability_bike
         + B_PT_Frequent_car * Frequent_car
         + B_PT_Frequent_bike * Frequent_bike
         + B_PT_Reimbursement * Reimbursement
         + B_PT_Reimbursement_PT * Reimbursement_PT
         + B_PT_Reimbursement_mileage * Reimbursement_mileage
         + B_PT_Reimbursement_leased * Reimbursement_leased)

V_Car_socio = (B_Car_Age_25_44 * Age_25_44
          + B_Car_Age_45_64 * Age_45_64
          + B_Car_Age_65 * Age_65
          + B_Car_Single * Single
          + B_Car_Couple * Couple
          + B_Car_Bachelor * Bachelor
          + B_Car_High_Edu * High_Edu
          + B_Car_Employment * Employment
          + B_Car_Mid_income * Mid_income
          + B_Car_High_income * High_income
          + B_Car_Disability * Disability
          + B_Car_Driving_license * Driving_license
          + B_Car_Availability_PT * Availability_PT
          + B_Car_Availability_car * Availability_car
```

```
          + B_Car_Availability_bike * Availability_bike
          + B_Car_Frequent_car * Frequent_car
          + B_Car_Frequent_bike * Frequent_bike
          + B_Car_Reimbursement * Reimbursement
          + B_Car_Reimbursement_PT * Reimbursement_PT
          + B_Car_Reimbursement_mileage * Reimbursement_mileage
          + B_Car_Reimbursement_leased * Reimbursement_leased)


V_Bike_socio = (B_Bike_Age_25_44 * Age_25_44
           + B_Bike_Age_45_64 * Age_45_64
           + B_Bike_Age_65 * Age_65
           + B_Bike_Single * Single
           + B_Bike_Couple * Couple
           + B_Bike_Bachelor * Bachelor
           + B_Bike_High_Edu * High_Edu
           + B_Bike_Employment * Employment
           + B_Bike_Mid_income * Mid_income
           + B_Bike_High_income * High_income
           + B_Bike_Disability * Disability
           + B_Bike_Driving_license * Driving_license
           + B_Bike_Availability_PT * Availability_PT
           + B_Bike_Availability_car * Availability_car
           + B_Bike_Availability_bike * Availability_bike
           + B_Bike_Frequent_car * Frequent_car
           + B_Bike_Frequent_bike * Frequent_bike
           + B_Bike_Reimbursement * Reimbursement
           + B_Bike_Reimbursement_PT * Reimbursement_PT
           + B_Bike_Reimbursement_mileage * Reimbursement_mileage
           + B_Bike_Reimbursement_leased * Reimbursement_leased)

V_PT = (ASC_PT + B_PT_access * PT_access
        + B_waiting * PT_waiting
        + B_PT_IVtime * PT_IVtime
        + B_transfer * Transfer
        + B_transfer_time * Transfer_time
        + B_transfer_distance * Transfer_distance
        + B_PT_TTtime * PT_TTtime
        + B_PT_type * PT_type
        + V_PT_socio)

V_Car =  (ASC_Car + B_car_access * Car_access
          + B_car_IVtime * Car_IVtime
          + B_car_parking * Car_parking
          + V_Car_socio)

V_Bike = (ASC_Bike + B_bike_IVtime * Bike_IVtime
           + V_Bike_socio)


# MNL with availability conditions
V = {0: V_PT, 1: V_Car, 2: V_Bike}
av = {0: av, 1: av, 2: av}
logprob = models.loglogit(V, av, choice)

biogeme = bio.BIOGEME(train_data, logprob)
```

```
biogeme.modelName = 'Final MNL'

# Get results
results = biogeme.estimate()

# Get the results in a pandas table
pandasResults = results.getEstimatedParameters()
pandasResults
```

Panel model estimation

```
# %% import the packages
import biogeme.biogeme as bio
import biogeme.database as db
import biogeme.models as models
import pandas as pd
from sklearn.model_selection import train_test_split
from biogeme.expressions import Beta, RandomVariable, Integrate, bioNormalCdf,
log, bioDraws
from biogeme.expressions import Beta, Variable, bioDraws,
log, MonteCarlo, exp, bioMultSum, PanelLikelihoodTrajectory

%matplotlib inline
import numpy as np
import matplotlib.pyplot as plt

from biogeme.expressions import Beta, Variable
from sklearn.model_selection import train_test_split
from sklearn.model_selection import StratifiedShuffleSplit

# %% load the data# %% Import packages


data = pd.read_csv('ML_Data.csv', delimiter=';')
# data = pd.read_csv('ML_train_data.csv')
# data = pd.read_csv('train_data.csv')
# data = pd.read_csv('Data.csv', delimiter = ';')
df = pd.DataFrame(data)
train_data = db.Database("train_data", df)
globals().update(train_data.variables)

# print(df.describe())

# print(df.isnull().sum())

# print(df[df.isin([np.nan, np.inf, -np.inf]).any(1)])

# %%Sort the data by individual ID to ensure panel data is correctly formatted
train_data.data = train_data.data.sort_values(by=['ID'])

# Define panel data structure
train_data.panel("ID")

globals().update(train_data.variables)
```

```
# %%Define variables

#Socio characteristics
Age_25_44 = Variable('Age(25-44)')
Age_45_64 = Variable('Age(45-64)')
Age_65 = Variable('Age(65)')
Single = Variable('Single')
Couple = Variable('Couple')
Bachelor = Variable('Bachelor')
High_Edu = Variable('High_Edu')
Employment = Variable('Employment')
Mid_income = Variable('Mid-income')
High_income = Variable('High-income')
Disability = Variable('Disability')
Driving_license = Variable('Driving_license')
Availability_PT = Variable('Availability_PT')
Availability_car = Variable('Availability_car')
Availability_bike = Variable('Availability_bike')
Frequent_car = Variable('Frequent_car')
Frequent_bike = Variable('Frequent_bike')
Reimbursement = Variable('Reimbursement')
Reimbursement_PT = Variable('Reimbursement_PT')
Reimbursement_mileage = Variable('Reimbursement_mileage')
Reimbursement_leased = Variable('Reimbursement_leased')


#Trip characteristics
PT_access = Variable('PT_access')
PT_waiting = Variable('PT_waiting')
PT_IVtime = Variable('PT_IVtime')
PT_TTtime = Variable('PT_TTtime')
Transfer = Variable('Transfer')
Transfer_time = Variable('Transfer_time')
Transfer_distance = Variable('Transfer_distance')
PT_type = Variable('PT_type')


Car_access = Variable('Car_access')
Car_IVtime = Variable('Car_IVtime')
Car_parking = Variable('Car_parking')

Bike_IVtime = Variable('Bike_IVtime')

av = Variable('Av')
choice = Variable('Choice')

# Model Parameters Specification
# 1 Name for report. Typically, same as variable
# 2 Starting value
# 3 Lower bound
# 4 Upper bound
# 5 0: estimate the parameter, 1: keep it fixed
ASC_PT = Beta('ASC_PT',0,None,None,1) # fixed to 0
ASC_Car = Beta('ASC_Car',0,None,None,0)
ASC_Bike = Beta('ASC_Bike',0,None,None,0)
```

```
#Generic attributes
B_Age_25_44 = Beta('B_Age_25_44', 0, None, None, 0)
B_Age_45_64 = Beta('B_Age_45_64', 0, None, None, 1)
B_Age_65 = Beta('B_Age_65', 0, None, None, 1)
B_Single = Beta('B_Single', 0, None, None, 1)
B_Couple = Beta('B_Couple', 0, None, None, 1)
B_Bachelor = Beta('B_Bachelor', 0, None, None, 1)
B_High_Edu = Beta('B_High_Edu', 0, None, None, 1)
B_Employment = Beta('B_Employment', 0, None, None, 1)
B_Mid_income = Beta('B_Mid_income', 0, None, None, 1)
B_High_income = Beta('B_High_income', 0, None, None, 1)
B_Disability = Beta('B_Disability', 0, None, None, 0)
B_Driving_license = Beta('B_Driving_license', 0, None, None, 0)
B_Availability_PT = Beta('B_Availability_PT', 0, None, None, 1)
B_Availability_car = Beta('B_Availability_car', 0, None, None, 1)
B_Availability_bike = Beta('B_Availability_bike', 0, None, None, 1)
B_Frequent_car = Beta('B_Frequent_car', 0, None, None, 1)
B_Frequent_bike = Beta('B_Frequent_bike', 0, None, None, 1)
B_Reimbursement = Beta('B_Reimbursement', 0, None, None, 1)
B_Reimbursement_PT = Beta('B_Reimbursement_PT', 0, None, None, 1)
B_Reimbursement_mileage = Beta('B_Reimbursement_mileage', 0, None, None, 1)
B_Reimbursement_leased = Beta('B_Reimbursement_leased', 0, None, None, 1)

#Specific attributes
B_PT_Age_25_44 = Beta('B_PT_Age_25_44 ', 0, None, None, 1)
B_Car_Age_25_44 = Beta('B_Car_Age_25_44', 0, None, None, 0)
B_Bike_Age_25_44 = Beta('B_Bike_Age_25_44', 0, None, None,1)

B_PT_Age_45_64 = Beta('B_PT_Age_45_64', 0, None, None, 1)
B_Car_Age_45_64 = Beta('B_Car_Age_45_64', 0, None, None, 1)
B_Bike_Age_45_64 = Beta('B_Bike_Age_45_64', 0, None, None, 1)

B_PT_Age_65 = Beta('B_PT_Age_65', 0, None, None, 1)
B_Car_Age_65 = Beta('B_Car_Age_65', 0, None, None, 1)
B_Bike_Age_65 = Beta('B_Bike_Age_65', 0, None, None, 1)

B_PT_Single = Beta('B_PT_Single', 0, None, None, 1)
B_Car_Single = Beta('B_Car_Single', 0, None, None, 1)
B_Bike_Single = Beta('B_Bike_Single', 0, None, None, 1)

B_PT_Couple = Beta('B_PT_Couple', 0, None, None, 1)
B_Car_Couple = Beta('B_Car_Couple', 0, None, None, 1)
B_Bike_Couple = Beta('B_Bike_Couple', 0, None, None, 1)

B_PT_Bachelor = Beta('B_PT_Bachelor', 0, None, None, 1)
B_Car_Bachelor = Beta('B_Car_Bachelor', 0, None, None, 1)
B_Bike_Bachelor = Beta('B_Bike_Bachelor', 0, None, None, 1)

B_PT_High_Edu = Beta('B_PT_High_Edu', 0, None, None, 1)
B_Car_High_Edu = Beta('B_Car_High_Edu', 0, None, None, 1)
B_Bike_High_Edu = Beta('B_Bike_High_Edu', 0, None, None, 1)

B_PT_Employment = Beta('B_PT_Employment', 0, None, None, 1)
B_Car_Employment = Beta('B_Car_Employment', 0, None, None, 1)
B_Bike_Employment = Beta('B_Bike_Employment', 0, None, None, 1)
```

```
B_PT_Mid_income = Beta('B_PT_Mid_income', 0, None, None, 0)
B_Car_Mid_income = Beta('B_Car_Mid_income', 0, None, None, 1)
B_Bike_Mid_income = Beta('B_Bike_Mid_income', 0, None, None, 1)

B_PT_High_income = Beta('B_PT_High_income', 0, None, None, 0)
B_Car_High_income = Beta('B_Car_High_income', 0, None, None, 1)
B_Bike_High_income = Beta('B_Bike_High_income', 0, None, None, 1)

B_PT_Disability = Beta('B_PT_Disability', 0, None, None, 1)
B_Car_Disability = Beta('B_Car_Disability', 0, None, None, 0)
B_Bike_Disability = Beta('B_Bike_Disability', 0, None, None, 1)

B_PT_Driving_license = Beta('B_PT_Driving_license', 0, None, None, 1)
B_Car_Driving_license = Beta('B_Car_Driving_license', 0, None, None, 0)
B_Bike_Driving_license = Beta('B_Bike_Driving_license', 0, None, None, 1)

B_PT_Availability_PT = Beta('B_PT_Availability_PT', 0, None, None, 1)
B_Car_Availability_PT = Beta('B_Car_Availability_PT', 0, None, None, 1)
B_Bike_Availability_PT = Beta('B_Bike_Availability_PT', 0, None, None, 0)

B_PT_Availability_car = Beta('B_PT_Availability_car', 0, None, None, 1)
B_Car_Availability_car = Beta('B_Car_Availability_car', 0, None, None, 1)
B_Bike_Availability_car = Beta('B_Bike_Availability_car', 0, None, None, 0)

B_PT_Availability_bike = Beta('B_PT_Availability_bike', 0, None, None, 0)
B_Car_Availability_bike = Beta('B_Car_Availability_bike', 0, None, None, 0)
B_Bike_Availability_bike = Beta('B_Bike_Availability_bike', 0, None, None, 1)

B_PT_Frequent_car = Beta('B_PT_Frequent_car', 0, None, None, 1)
B_Car_Frequent_car = Beta('B_Car_Frequent_car', 0, None, None, 1)
B_Bike_Frequent_car = Beta('B_Bike_Frequent_car', 0, None, None, 1)

B_PT_Frequent_bike = Beta('B_PT_Frequent_bike', 0, None, None, 1)
B_Car_Frequent_bike = Beta('B_Car_Frequent_bike', 0, None, None, 1)
B_Bike_Frequent_bike = Beta('B_Bike_Frequent_bike', 0, None, None, 1)

B_PT_Reimbursement = Beta('B_PT_Reimbursement', 0, None, None, 1)
B_Car_Reimbursement = Beta('B_Car_Reimbursement', 0, None, None, 1)
B_Bike_Reimbursement = Beta('B_Bike_Reimbursement', 0, None, None, 1)

B_PT_Reimbursement_PT = Beta('B_PT_Reimbursement_PT', 0, None, None, 0)
B_Car_Reimbursement_PT = Beta('B_Car_Reimbursement_PT', 0, None, None, 1)
B_Bike_Reimbursement_PT = Beta('B_Bike_Reimbursement_PT', 0, None, None, 1)

B_PT_Reimbursement_mileage = Beta('B_PT_Reimbursement_mileage', 0, None, None, 1)
B_Car_Reimbursement_mileage = Beta('B_Car_Reimbursement_mileage', 0, None, None, 1)
B_Bike_Reimbursement_mileage = Beta('B_Bike_Reimbursement_mileage', 0, None, None, 1)

B_PT_Reimbursement_leased = Beta('B_PT_Reimbursement_leased', 0, None, None, 0)
B_Car_Reimbursement_leased = Beta('B_Car_Reimbursement_leased', 0, None, None, 1)
B_Bike_Reimbursement_leased = Beta('B_Bike_Reimbursement_leased', 0, None, None, 1)

# Trip characteristics
B_PT_access = Beta('B_PT_access', 0, None, None, 0)
B_waiting = Beta('B_waiting', 0, None, None, 1)
B_PT_IVtime = Beta('B_PT_IVtime', 0, None, None, 0)
```

```
B_transfer = Beta('B_transfer', 0, None, None, 0)
B_transfer_time = Beta('B_transfer_time', 0, None, None, 1)
B_transfer_distance = Beta('B_transfer_distance', 0, None, None, 1)
B_PT_TTtime = Beta('B_PT_TTtime', 0, None, None, 1)
B_PT_type = Beta('B_PT_type', 0, None, None, 1)

B_car_access = Beta('B_car_access', 0, None, None, 0)
B_car_IVtime = Beta('B_car_IVtime', 0, None, None, 0)
B_car_parking = Beta('B_car_parking', 0, None, None, 0)

B_bike_IVtime = Beta('B_bike_IVtime', 0, None, None, 0)

#Panel effect
# Sigma_panel = Beta('Sigma_panel',0,-100,100,0)
# Zero = Beta('Zero',0,-100,100,1)
# Zero_sigma_panel = Zero + Sigma_panel * bioDraws('Zero_sigma_panel','NORMAL')

Sigma_panel = Beta('Sigma_panel', 0, -100, 100, 0)
Zero = Beta('Zero', 0, -100, 100, 1)
Zero_sigma_panel = Zero + Sigma_panel * bioDraws('Zero_sigma_panel', 'NORMAL')



V_PT = (B_PT_access * PT_access
        + B_PT_IVtime * PT_IVtime
        + B_transfer * Transfer
        + B_PT_Availability_bike * Availability_bike
        + B_PT_Reimbursement_leased * Reimbursement_leased
        + Zero_sigma_panel)

V_Car =  (ASC_Car + B_car_access * Car_access
         + B_car_IVtime * Car_IVtime
         + B_car_parking * Car_parking
         + B_Car_Availability_bike * Availability_bike
         + Zero_sigma_panel)

V_Bike = (ASC_Bike + B_bike_IVtime * Bike_IVtime
          )

# %% MNL with availability conditions
V = {0: V_PT, 1: V_Car, 2: V_Bike}
av = {0: av, 1: av, 2: av}

obsprob = models.logit(V,av,choice)
condprobIndiv = PanelLikelihoodTrajectory(obsprob)
logprob = log(MonteCarlo(condprobIndiv))

biogeme = bio.BIOGEME(train_data, logprob, numberOfDraws=5000)
biogeme.modelName = 'Final Panel ML'

# Get results
results = biogeme.estimate()

# Get the results in a pandas table
pandasResults = results.getEstimatedParameters()
pandasResults
```

E.2. Model application

```
# %%Import packages
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import biogeme.database as db
import biogeme.biogeme as bio
import random

#%% Generate population
class Individual:
    def __init__(self, age_25_44, high_income, mid_income, disability, driving_
    license, availability_bike, availability_pt, availability_car, reimbursement
    _leased, reimbursement_pt):
        self.age_25_44 = age_25_44
        self.high_income = high_income
        self.mid_income = mid_income
        self.disability = disability
        self.driving_license = driving_license
        self.availability_bike = availability_bike
        self.availability_pt = availability_pt
        self.availability_car = availability_car
        self.reimbursement_leased = reimbursement_leased
        self.reimbursement_pt = reimbursement_pt

def assign_characteristic(probability):
    return 1 if random.randint(1, 1000) <= probability * 1000 else 0

def generate_population(size=52500):
    population = []
    for _ in range(size):
        age_25_44 = assign_characteristic(0.27)
        high_income = assign_characteristic(0.35)
        mid_income = assign_characteristic(0.25)
        disability = assign_characteristic(0.12)
        driving_license = assign_characteristic(0.61)
        availability_bike = assign_characteristic(1)
        availability_pt = assign_characteristic(1)
        availability_car = assign_characteristic(0.54)
        reimbursement_leased = assign_characteristic(1)
        reimbursement_pt = assign_characteristic(1)
        individual = Individual(age_25_44, high_income, mid_income, disability,
        driving_license, availability_bike, availability_pt, availability_car,
        reimbursement_leased, reimbursement_pt)
        population.append(individual)
    return population

# Generate the population
population = generate_population()

#%% Define MNL Model and attributes

ASC_Bike = 2.193053
ASC_Car = −1.349201
```

```
B_Bike_Availability_PT = −0.418877
B_Bike_Availability_car = −0.206820
B_Car_Age_25_44 = −0.000294
B_Car_Availability_bike = −0.625507
B_Car_Disability = −0.000174
B_Car_Driving_license = 0.000195
B_PT_Availability_bike = −0.625710
B_PT_High_income = −0.298933
B_PT_IVtime = −0.067529
B_PT_Mid_income = 0.298869
B_PT_Reimbursement_PT = 0.363886
B_PT_Reimbursement_leased = −0.363823
B_PT_access = −0.139800
B_bike_IVtime = −0.133990
B_car_IVtime = −0.041241
B_car_access = −0.109964
B_car_parking = −0.017368
B_transfer = −0.802567

# Input values for each scenario
# scenario = 'Conventional'
# PT_access = 10
# PT_IVtime = 41
# Transfer = 1
# Car_access = 0
# Car_IVtime = 43
# Car_parking = 10
# Bike_IVtime = 38

scenario = 'Sustainable'
PT_access = 7
PT_IVtime = 31
Transfer = 1
Car_access = 5
Car_IVtime = 13        #13/43
Car_parking = 10
Bike_IVtime = 50         #38/50

# scenario = 'Ambitious'
# PT_access = 5
# PT_IVtime = 31
# Transfer = 0
# Car_access = 15
# Car_IVtime = 13           #13/43
# Car_parking = 60
# Bike_IVtime = 38            #38/50

#%% Calculate probability

def calculate_probabilities(individual):
    V_PT = (B_PT_access * PT_access +
            B_PT_IVtime * PT_IVtime +
            B_transfer * Transfer +
            B_PT_High_income * individual.high_income +
            B_PT_Mid_income * individual.mid_income +
            B_PT_Reimbursement_PT * individual.reimbursement_pt +
```

```
            B_PT_Reimbursement_leased * individual.reimbursement_leased +
            B_PT_Availability_bike * individual.availability_bike)

    V_Car = (ASC_Car +
             B_car_access * Car_access +
             B_car_IVtime * Car_IVtime +
             B_car_parking * Car_parking +
             B_Car_Age_25_44 * individual.age_25_44 +
             B_Car_Availability_bike * individual.availability_bike +
             B_Car_Disability * individual.disability +
             B_Car_Driving_license * individual.driving_license)

    V_Bike = (ASC_Bike +
              B_bike_IVtime * Bike_IVtime +
              B_Bike_Availability_PT * individual.availability_pt +
              B_Bike_Availability_car * individual.availability_car)

    exp_V_PT = np.exp(V_PT)
    exp_V_Car = np.exp(V_Car)
    exp_V_Bike = np.exp(V_Bike)

    denominator = exp_V_PT + exp_V_Car + exp_V_Bike

    P_PT = exp_V_PT / denominator
    P_Car = exp_V_Car / denominator
    P_Bike = exp_V_Bike / denominator

    return P_PT, P_Car, P_Bike
# %% Choice  simulation
pop_size = 39585   #39585
population = generate_population(pop_size)  # Simulate a population of 100 individuals

# Function to simulate one run of mode choice
def simulate_mode_choice(population):
    count_PT = 0
    count_Car = 0
    count_Bike = 0

    for individual in population:
        P_PT, P_Car, P_Bike = calculate_probabilities(individual)
        modes = ['PT', 'Car', 'Bike']
        probabilities = [P_PT, P_Car, P_Bike]
        chosen_mode = random.choices(modes, probabilities)[0]

        if chosen_mode == 'PT':
            count_PT += 1
        elif chosen_mode == 'Car':
            count_Car += 1
        elif chosen_mode == 'Bike':
            count_Bike += 1

    return count_PT, count_Car, count_Bike

# Create an empty list to store the results
results_list = []
```

```
# Perform runs on the same population
num_runs = 10

for run in range(num_runs):
    count_PT, count_Car, count_Bike = simulate_mode_choice(population)

    # Store the results in a dictionary and add to the list
    results_list.append({
        'Run': run + 1,
        'PT_Count': count_PT,
        'Car_Count': count_Car,
        'Bike_Count': count_Bike
    })

# Convert the list of dictionaries to a DataFrame
results_df = pd.DataFrame(results_list)
# print(results_df)

# %% Calculate the required min sample size
average_counts = results_df.mean()
variance_counts = results_df.var()

def required_runs(mu, variance):
    n =( 1.96**2 / ((0.05*mu)**2) )* variance
    return n

print (f"Required minimum runs PT is {required_runs(average_counts['PT_Count'],
       variance_counts['PT_Count'])}")
print (f"Required minimum runs Car is {required_runs(average_counts['Car_Count'],
       variance_counts['Car_Count'])}")
print (f"Required minimum runs PT is {required_runs(average_counts['Bike_Count'],
       variance_counts['Bike_Count'])}")


# %% Results of application
print(f"For scenario {scenario}, Average number of persons choosing each alternative
      over {num_runs} runs:")
print(f"Average PT_Count: {average_counts['PT_Count']}, PT share:
      {average_counts['PT_Count']/pop_size:.4f}")
print(f"Average Car_Count: {average_counts['Car_Count']}, Car share:
      {average_counts['Car_Count']/pop_size:.4f}")
print(f"Average Bike_Count: {average_counts['Bike_Count']}, Bike share:
      {average_counts['Bike_Count']/pop_size:.4f}")
```