Night-Time Train Travel

A Stated-Preference study into the Willingness to Use night trains for European long-distance travel

M.K. Heufke Kantelaar

TIL Thesis

Night-Time Train Travel

A Stated-Preference study into the Willingness to Use night trains for European long-distance travel

by

M.K. Heufke Kantelaar

to obtain the degree of Master of Science in

Transport, Infrastructure & Logistics

at Delft University of Technology,
to be defended publicly on Monday August 26, 2019 at 12:45 PM.

Student number:	4235193	
Project duration:	March 4, 2019 – August 26, 2019	
Thesis committee:	Prof. dr. G. P. van Wee,	TU Delft, committee chair
	Dr. E. J. E. Molin,	TU Delft, supervisor
	Dr. O. Cats,	TU Delft, supervisor
	Ir. B. J. H. F. Donners,	Royal HaskoningDHV

The cover photo is the courtesy of Austrian Federal Railways ÖBB.

TU Delft

Royal HaskoningDHV
Enhancing Society Together

Preface

Dear reader,

This document is the final report for the master thesis. The thesis is the final capstone project of the education program Transport, Infrastructure and Logistics, at Delft University of Technology, which is to be awarded with a Master of Science. The thesis is carried out as a graduation internship at Royal HaskoningDHV.

The subject of this thesis is night-time train travel and investigates the Willingness to Use this as a travel mode for long-distance European travel. This report is the complete work and includes an introduction to the topic, methodology, theoretic framework, survey design, results and application sections.

For now I would like to thank my supervisors from TU Delft for always being very approachable for feedback and critical, yet very helpful comments. Furthermore, I would like to thank (employees of) Royal HaskoningDHV for the opportunity, their enthusiasm and helpful talks we had about my research. Additionally, I would like to express my gratefulness to Dutch Railways, and especially to Mrs Van de Kar for helping in distributing the survey. A thank you to the people who helped with testing the draft survey. At last, a thank you to all friends and family who supported me through the process of writing the thesis and were open for discussion and provided distraction when needed.

I hope you enjoy reading the report,

M.K. Heufke Kantelaar
Delft, August 2019

Contents

List of Figures

List of Tables

1 Introduction

Flying has become one of our favourite travel modes and takes a central place in our society. The average Dutch person flies on average 1.3 times a year (Zijlstra & Huibregtse, 2018). Numbers from the Schiphol Group (2018) show that in 2018 71.1 million passengers used Schiphol Amsterdam Airport to reach their destination. The majority of these passengers (70.4%) travelled within the European continent. In addition, 63.4% of the total number of passengers consisted of direct, non-transfer traffic. Therefore, it can be concluded that the majority of the passengers using Schiphol Airport flies directly from city to city within Europe over a relatively short distance. This is a huge amount of passengers and has serious consequences for the environment.

Air travel is a threat to the environment as it is a major contributor to human-induced greenhouse gas emissions, which is the driving force for anthropogenic climate change. Climate change is one of the most challenging issues for mankind in this era. The Paris Agreement of 2015 showed that almost all countries around the world are committed to limiting climate change to two degrees Celsius (United Nations, 2018). However, both Boeing and Airbus forecast an annual average growth rate in air passengers of 3.3% and 4.0% for Europe until 2037 (Airbus, 2018; Boeing, 2018). Clearly, to reduce the impacts of climate change a switch to more sustainable mobility is needed.

Travelling by train is a much more environmentally friendly option and can, therefore, play a significant role in this switch to sustainable mobility. Conventional train journeys emit four to five times less CO_2 emissions than an aeroplane journey along that same corridor (Álvarez, 2010). This saving of emissions could increase even further when operating high-speed railway lines, which consume on average 29% less energy compared to conventional trains (Álvarez, 2010). Therefore, replacing flights within the European continent by train journeys could result in a significant step towards more sustainable mobility.

However, an important limitation for this substitution is the covered distance. Flying is perceived as a quick way to travel and during the flight, a large distance can be covered in a short amount of time. Trains are usually far slower. This distance problem can be solved in roughly two ways. One way implies investing in a major European high-speed railway network. Another option is to increase the usage of travelling by night train. Night trains can travel long distances while passengers are asleep. For journeys of between 12-14 hours (depart in the evening, arrive in the morning) this could pose an interesting alternative, which results in distances of roughly 1000 to 1200 kilometres (when considering conventional railway tracks). This distance range could connect city pairs such as Amsterdam - Vienna (+- 1175 km by rail) and Amsterdam - Milan (+-1090 km by rail). To place this group of travellers into perspective, both these destinations are in the 2018 top 20 destinations from Schiphol Airport and in 2017 the number of aeroplane passengers to both of these destinations was around 750.000 (Eurostat, 2017). Additionally, the total number of commercial flights between Schiphol and Milan is about 9.150 (combined for both airports serving Milan), which makes it the 2^{nd} busiest destination from Schiphol (Eurostat, 2017).

Another interesting aspect of travelling by night train is that the number of flights between these city pairs can be reduced. Multiple large European airports are capacity constrained. Last year, Schiphol had 499.000 movements, almost reaching its maximum capacity of 500.000 flight movements a year (Centraal Bureau voor de Statistiek, 2019). Switching from air travel to train travel can save a considerable amount of flights. A report from Royal HaskoningDHV points out that improving tickets & service, passenger rights and optimising train services with the different operators can save up to 89.000 flight movements from Schiphol, and investing in a major European high-speed rail network even an additional 43.000 movements (Donners, 2018). It can be argued that the night train might even act as a sort of 'multiplier' for these numbers, as during the night even more distant destinations can be reached. For Schiphol, this opens up the possibility to solve part of their congestion problem and allows the airport to focus on further developing its intercontinental network.

However, it is unclear if people are willing to use the night train as a travel alternative to get to their destination. At first sight, travelling by night train seems attractive. For the traveller, it has several advantages: it allows the traveller to sleep or do other pleasant activities during the trip, arrive rested in the morning at their destination and save on hotel costs. However, the modality choice for using night trains should be considered in a broader context.

Traveller behaviour of individuals can be described using various perspectives. One of these perspectives is the econometric one, which focuses on minimising disutility of travel (Arentze & Molin, 2013). This perspective takes attributes into account such as 'travel time' and 'cost'. Furthermore, the marketing perspective adds possibilities to identify user groups and adapt the policy to their needs (Anable, 2005).

With travelling, several choices are involved (Pel, 2017). The focus of this study is the mode choice decision. Often included attributes in mode choice models are trip time and trip costs. Additionally, travelling by night train has several advantages that could result in people preferring it over flying. The night train typically arrives in the morning at the destination, therefore the Willingness to Use the night train is likely influenced by the preferred arrival time at the destination. This thesis will look into this by using contexts with different arrival times in the morning. By doing so, the influence of the arrival time can be observed. The travel options will consist of taking the night train, flying in the early morning, or the day before and stay in a hotel (ignoring private travel modes).

Scientifically, there is little to no knowledge about the Willingness to Use night trains as a travel mode for long-distance trips. To the best of the authors' knowledge, no study into this subject has been done. However, some sub-parts (like mode choice, time valuation, the influence of socio-demographics and travelling at night) that contribute to the concept of travelling by night train have been addressed by literature. A concise overview: standard mode choice models consider, among others, travel time and costs. Román, Espino, and Martín (2010) found that these factors also apply for long-distance trips such as Madrid-Barcelona. However, it is not clear if there are additional attributes that need to taken into account when specifically considering night trains. Furthermore, there is some evidence that entertainment, comfort and conducting activities during travel contribute to a less negative utility of travelling, which could affect mode choice (Li, 2015; Mokhtarian & Salomon, 2001). In addition, the valuation of travel time differs with trip characteristics and socio-demographics (Börjesson & Eliasson, 2014; Ozbay & Yanmaz-Tuzel, 2008; Small, 2012). Limtanakool, Dijst, and Schwanen (2006a) investigated the effect of socio-demographics on mode choice and indeed found a strong relationship between personal and household attributes, such as gender, age and educational level, on mode choice.

Deriving insights into the choices and trade-offs individuals make and estimating resulting parameters for utility functions can be done by using Discrete Choice Modelling. Application of these models gives the researcher information about the willingness to use night trains and potential modality shares (McFadden, 1972). Data for discrete choice models can be divided into two main categories: Revealed Preference (RP) data or Stated Preference (SP) data. At this moment, there are no long-distance night trains in operation in the Netherlands, making the collection of RP data impossible. Using SP data, this issue is circumvented as it offers complete control of the choice set and information shown to decision-makers.

1.1. Research objective and questions

This thesis is written as a graduation project at Royal HaskoningDHV. This research aims to provide more knowledge about the Willingness to Use night trains, as a substitution for flying, for European long-distance travel. The company considers itself the market leader in the long-distance sustainable mobility field. It has identified several steps (the aforementioned improvements in tickets & service, optimisation of services and investment in a European high-speed network) that would contribute to the train being a serious competitor for travelling by aeroplane. However, it has little knowledge about the willingness of individuals to use night trains, and which factors they take into account when considering the night train travel option.

The company would like to gather this knowledge as it expects the night train could potentially enhance the effect of the previously mentioned steps. Furthermore, the company would like to use these insights to advise its customers on topics within the night train domain, such as the introduction of new night train services. To be able to fulfil this aim properly, a combination of a qualitative and quantitative study is chosen. The qualitative part to get preliminary insights into which factors and attributes people take into account when considering the night train travel option. The quantitative part to provide more insight into the actual choice behaviour of individuals, which allows for estimating choice models and predicting potential modal shares for varying scenarios.

Combining the identified problems, context and research aim, the main research question is formulated as:

To what extent are Dutch travellers willing to use night trains, as an alternative for flying, for European long-distance trips?

With the following sub-questions:

1. What are the comfort determinants of the night train that Dutch travellers consider for European trips with a duration of 12-14 hours?
2. What explanatory variables do travellers take into account when choosing between night train and aeroplane alternatives for European long-distance travel?
3. To what extent do socio-demographics and other background variables influence the night train comfort and mode choice?
4. Which segments for mode choice can be identified?
5. What is the sensitivity of the modality share for combinations of explanatory variables?
6. What are the predicted market shares of the different modes for certain scenarios?

The first question is about getting to know what the determinants of comfort are that individuals take into account when considering the night train travel option. It is expected that various comfort factors (such as private or shared accommodation, the possibility to shower, the number of stops etc.) affect the perceived comfort level of decision-makers. However, it is not yet known what factors contribute to this 'level of comfort'. This is answered by qualitative and quantitative research. The second sub-question aims to add knowledge about which explanatory variables are considered when people are faced with the night train travel option. Third, it will be explored to what extent socio-demographics and other background variables influence this comfort and mode choice. These same socio-demographics can be used to identify market segments who are most likely to use the night train. This provides insight from a marketing perspective. The last two questions are about the application of the estimated model. The sensitivity of the modality shares for changes in explanatory variables is explored as well as the potential market share for the different modes in several scenarios.

1.2. Relevance

This topic has relevance from society, science and the internship company perspective. As mentioned earlier on, for society the topic is relevant as it can help in assessing if individuals are willing to use the night train for European long-distance travel and switch to more sustainable mobility. Furthermore, this could potentially mitigate airport congestion by reducing the number of short-haul flights. Also, the results can help railway undertakings in their decision to reintroduce night trains. Scientifically it will be the first study into the Willingness to Use night trains for European long-distance travel, adding knowledge about this topic and its sub-parts such as determinants of comfort and explanatory variables that need to be taken into account for night-time train travel. Finally, the topic is relevant for the company as it increases its expertise in sustainable long-distance mobility and the ability to advise its clients about night train decisions and policies. Table 1.1 summarises the relevance from the different perspectives.

Table 1.1: Societal, scientific and company relevance of the topic

Societal Relevance
Results can be used to discover if people are willing to use the night train for sustainable long-distance European transport
Help in partially solving airport congestion by reducing short-haul flights
Results can be used by railway undertakings in their decision to reintroduce night trains
Scientific Relevance
First study into the willingness to use night trains for European long-distance travel
Add to the knowledge about what determinants of comfort and explanatory variables to include for night-time train travel
Company Relevance
Increase the company expertise in long-distance sustainable mobility
Increase the ability to advise clients about night train decisions and policies

1.3. Thesis outline

The first part of the thesis will introduce and define the night train concept in Chapter 2. Chapter 3 examines in more depth the methodology that will be applied. The theoretical framework in Chapter 4 discusses the elements of the choice model from a literature viewpoint, completed with input from focus groups and presents a framework. The second part of the study is about the data collection and modelling phase. The survey will be designed in Chapter 5. The collected data will be analysed for sample and travel characteristics in Chapter 6. The model estimation process and results are detailed in Chapter 7. Estimated models will be used for sensitivity analysis and exploring the effects of scenarios on modal share in Chapter 8. At last, key findings, implications, limitations and recommendations of the study for further research will be presented in Chapter 9.

2

Night train concept

Night trains, also sometimes called 'sleeper trains', are currently not in operation in the Netherlands. Therefore, for those unfamiliar, this chapter will introduce the concept of night trains. The definition is given in Section 2.1, followed by an introduction of current service levels and differences between them in Section 2.2. Additionally, in Section 2.3 a brief history about night train operations in Europe is presented. A concise summary can be found in Section 2.4 .

2.1. Definition of a night train

Defining the night train travel concept involves the train type, typical operating hours and service classes. Figure 2.1 shows several different concepts regarding the trip duration and train type that can be used (DB International & International Union of Railways (UIC), 2013b). Three different kinds of 'night train models' can be identified. The first one is the 'standard day train' model, which uses the regular day-time rolling stock for running overnight services. This type of service is most similar to the 'Nachtnet' service operated by the Nederlandse Spoorwegen (Dutch Railways, in the remainder abbreviated as NS). The second category of night train services is the 'conventional' night train, which is still operated on services within Central-Europe. Differences in this category are mainly due to the offered accommodation types. The last category is the so-called 'touristic trains' which are often luxury special trains or trains making multi-day trips. The remainder of this thesis will use the definition of the 'Traditional Night train', with the addition of seating accommodation beside the couchette and sleeper cars.

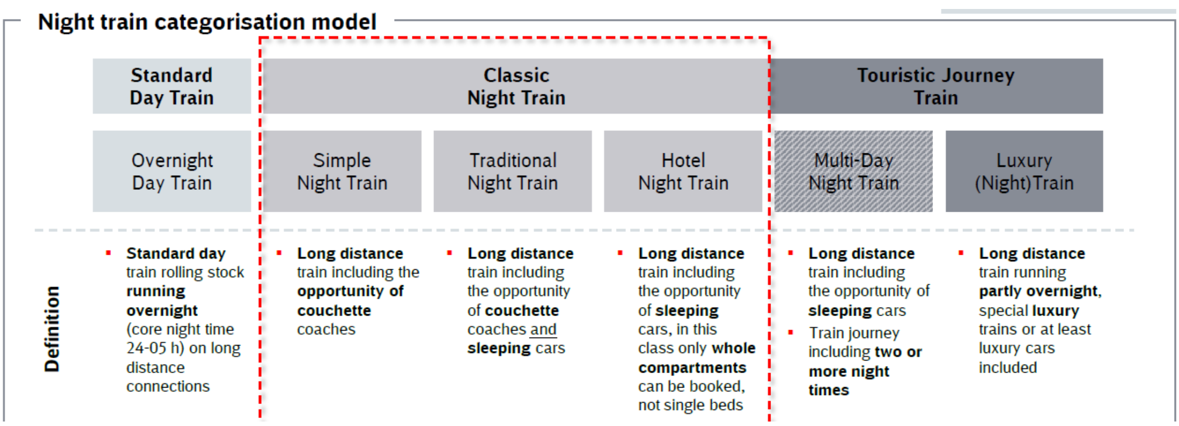


Figure 2.1: Different night train categories (adapted from DB International and International Union of Railways (UIC) (2013b))

Regarding the hours of operation: the ideal overnight train departs after 22:00 and arrives before 08:00, which means a trip time of fewer than 10 hours and almost all passengers travel long distances (Bird et al., 2017). However, for travelling to destinations such as Vienna and Milan, the trip time will be longer than 10 hours, which means that the train departs before 22:00 or arrives after 08:00. This could result in the night train carrying also late evening or early morning commuters, especially if the train stops frequently upon leaving the first station or nearing its final destination (Bird et al., 2017).

The trip of a night train can be divided into three phases, which are illustrated in Figure 2.2. The first phase is the so-called 'boarding area', during the phase the train stops several times along the line to offer several origins for joining the service to its final destination. For high-quality night train services this 'boarding area' phase ends around midnight. After midnight, the 'ideal' night train runs uninterrupted to cover long distances until about 06:00. At last, when the train reaches the 'de-boarding area' the train will stop more often when it gets closer to its final destination to drop off passengers. Inspection of several night train schedules from Austrian operator ÖBB indicates that most night trains do stop during the middle of the night, deviating from an ideal service. Timetables indicate that services stop generally every hour to every three hours between midnight and early morning (ÖBB, 2019). This is likely to result in a degraded sleep quality due to the noise of passengers boarding and deboarding. During this thesis, the effect of the number of stops on the travel comfort will be explored.

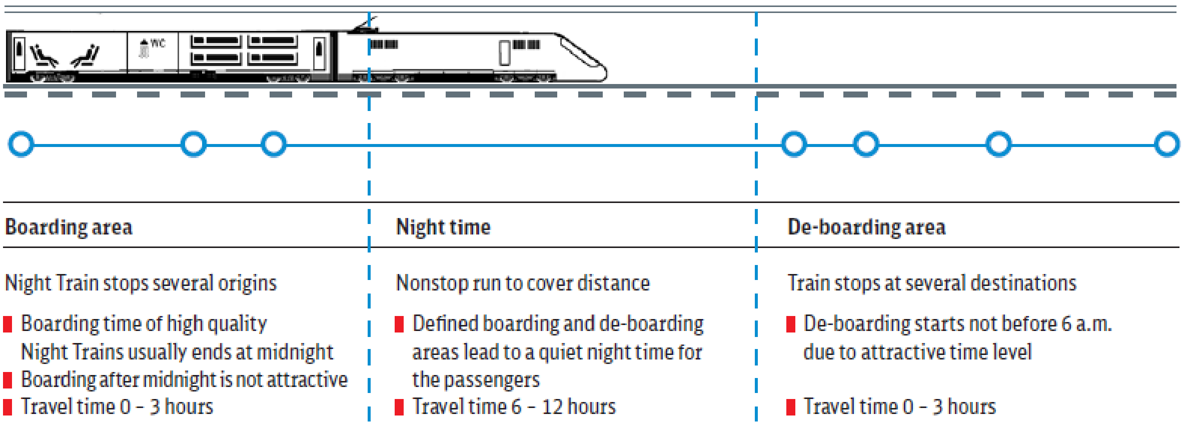


Figure 2.2: Phases of an ideal night train trip (adapted from DB International and International Union of Railways (UIC) (2013a))

2.2. Travel experience of the night train

To illustrate the travel experience on board the night train, information from ÖBB is used. Reconnecting the Netherlands to the European night train network would likely imply extending the ÖBB lines from Dusseldorf towards Arnhem, Utrecht and Amsterdam. This is the reason why information from ÖBB is used to illustrate the travel experience. As mentioned in Section 2.1, the night trains operated within the ÖBB network do not offer a non-stop service between midnight and early morning, which could lead to interrupted sleep. The night trains that are operated offer several classes of service. Usually, there are three kinds of services available: seats, couchettes and sleeper carriages. Couchettes and sleeper carriages both offer the possibility to lie-down for sleeping. Couchette carriages are more 'hostel' style, while sleeper cabins have nicer beds, are more private and can offer more facilities like a private bathroom. Furthermore, night trains do have a railway carriage that includes a catering section that sells food or drinks during the journey. An overview of typical services from ÖBB trains is provided in Table 2.1, the prices reflect starting one-way prices.

Table 2.1: Overview of properties of different service classes on typical night trains

Accomodation type / Attribute	Seat	Couchette	Sleeper
Capacity of compartment:	6 seats	4 or 6 beds	2 or 3 beds
Lockable compartment:	No	Yes	Yes
Bathroom:	Toilets and washbasin in carriage	Toilets and washbasin in carriage	Standard: washbasin in compartment, toilets and showers in carriage Deluxe: private washbasin, toilet and shower in compartment
Breakfast included:	No	Light breakfast	Extended breakfast
Wake-up service:	No	No	Yes
Price (starting, per person):	€29	€49 (6 beds) €59 (4 beds)	Standard: €69 (3 beds) / €89 (2 beds) Deluxe: €99 (3 beds), €119 (2 beds)

2.3. History and current operations of night trains in Europe

The usage of night trains in Europe goes back decades of years. Night train travel was very common in Europe during the nineteenth and beginning of the twentieth century. By the 1930s the Compagnie Internationale des Wagon-Lits operated an extensive network of sleeper trains within Europe, even all the way to Istanbul (Bird et al., 2017). However, after the Second-World War, technology developments made it possible to travel long distances by private car on motorways and commercial flying took off. Bird et al. (2017) identified that starting from the 1980s, there has been a downward trend in the number of passenger night-time services. This has even led to the withdrawal of the connection Amsterdam - Dusseldorf in 2016, citing ’low demand’ and ’investment in rolling stock’ reasons according to Deutsche Bahn (DB). As a consequence the Netherlands was disconnected from the European night train network (Nederlandse Spoorwegen, 2016; OVPro, 2016).

Bird et al. (2017) identified several causes for this decline in passenger night train services. First, it is stated that operating night trains involves higher costs per passenger space. This is because the rolling stock is more complex, built-in smaller volumes and carries fewer passengers per wagon. Additionally, staff costs are higher as the staff is required to work overnight and additional services such as bed-making are needed. Thirdly, changing social norms (i.e. not willing to share bunk beds, no direct access to a bathroom) are identified as causes. Furthermore, growing competition from other modes, such as low-cost plane tickets and a liberated international night bus market is mentioned. Another identified reason for the decrease in service coverage is the increase in daytime high-speed train services. At last, it is mentioned that there is limited market awareness of night train travel options. Only regular or local (in terms of departure or arrival station) passengers might be aware of those travel options, whereas the night train option is invisible for other travellers. (Bird et al., 2017).

However, the future of night-time train services may not be as negative as it may appear. Since the Austrian Federal Railways ÖBB took over night-time services from Deutsche Bahn, the tide seems to have turned. The company calls its night-train travel concept, ’Nightjet’, a success. In 2017 more than 1.4 million passengers travelled by ’Nightjet’ all over central Europe (Lechner, 2018). For ÖBB the ’successful night traffic concept’ clearly has a future as it is (re)introducing new destinations to its network. The night train between Berlin and Vienna was reintroduced in 2018 (Lechner, 2018). An overview of the ÖBB night-time network is provided in Figure 2.3. Note that there are some other night trains operating in Europe, which are not shown in Figure 2.3. However, the network mainly radiates out from Austria. Furthermore, ÖBB is investing in new carriages to cater to the needs of the modern traveller (note that that is one of the reasons mentioned for the decline). Starting from 2021, new Nightjet carriages will enter service, providing several service levels, including small private cabins for solo travellers and family compartments (OBB, 2018b).

Figure 2.3: Overview of the ÖBB night-time network (OBB, 2018a)

At last, also in government, the perception of night trains is changing. A report of the European Environment Agency, part of the European Commission, mentions night-time train travel as an alternative for flying (European Environment Agency, 2018). Also the Dutch government last year overwhelmingly agreed to a resolution stressing the importance of reconnecting the Netherlands to the European night train network (Kröger, 2018).

2.4. Summary

Main takeaways

- The used night train definition is that of the 'traditional night train', offering three types of accommodation: seats, couchettes and sleeper.
- The ideal night train journey takes about 10 hours, in reality, this is often longer.
- Since the 1980s the number of night train services has steadily declined. The Netherlands is no longer serviced since 2016.
- Identified causes for the demise are higher operating costs, changing social norms, competition from other modes and limited awareness.
- During recent years, night trains operated by ÖBB are making a come-back. Politicians are once again in favour of reintroducing services.

The night train definition that will be used in this thesis is that of a 'traditional night train', which means a night train that operates using specialised equipment offering several service classes. These service classes are seat, couchette and sleeper. The ideal night train operation takes about 10 hours, resulting in mostly long-distance passengers, and runs non-stop during the middle of the night. Nevertheless, currently, most night train services operate for a longer duration and make several stops during the night. This will be taken into account in this study. Service classes differ in price, offered accommodation, the number of people in the compartment, the possibility to lock your compartment, bathroom facilities, if breakfast is included and other services. At last, an overview of the history of night train operations within Europe and possible causes for its demise were identified. These are higher operating costs, changing social norms (not willing to share bunk beds), competition from low-cost flights and busses, the increase in high-speed train services, and limited awareness of travel options. However, during the last years, the night train operated by ÖBB is making a come-back. Furthermore, politicians in the Netherlands, as well as Europe, are in favour of reintroducing night trains. It is considered to be a 'green alternative' for flying.

3
Methodology

This chapter introduces the methodology that will be applied for answering the research questions. First, Section 3.1 highlights the steps taken for literature research. Section 3.2 introduces the Stated-Preference survey for collecting data, and includes information about the rating experiment in subsection 3.2.1. Section 3.3 will introduce the Discrete Choice Modelling that will be used to analyse the collected data. A summary is given in section 3.4.

3.1. Literature research

Researching literature was done by using a combination of different search engines. Scopus, WorldCat Discovery (via TU Delft Library) and Google Scholar were all used to find relevant scientific articles about the topic. Examples of keywords that were used are: train, night, night time train travel, mode choice, multimodal, long-distance, aeroplane, high speed, departure time choice, travel comfort, Value of Time, heterogeneity, Hierarchical Information Integration, Stated-Preference, hypothetical bias and Discrete Choice Modelling. Additionally, search operators such as AND were used. Furthermore, when reviewing search results, attention was paid to the source, journal it was published, year of publication and number of citations. If after reading the abstract and conclusion the paper was considered to be useful a more in-depth investigation was done. In that case, 'snowballing' was also applied to discover other relevant papers about the topic and papers in which the work was cited. In some cases the usage of 'grey literature' could not be avoided, this was for example used in the night train definition for which a study by Deutsche Bahn was used.

3.2. Stated-Preference survey

Data has to be collected before a choice model can be estimated and the Willingness to Use night trains be assessed. Data for estimating choice models can be collected in two ways, one can collect Revealed-Preference (RP) data or Stated-Preference (SP) data. Whitehead, Pattanayak, Van Houtven, and Gelso (2008) summarise the advantages and disadvantages of both methods in their paper. A major limitation of collecting RP data is that only existing alternatives can be evaluated, and therefore cannot be used to predict market shares of new modes. As no night trains are operating in the Netherlands at this moment, it is therefore impossible to collect RP data. For this reason, the usage of SP data is chosen. This allows for complete control over the experiment, the included alternatives, attributes and range of attribute values. Furthermore, the survey can be designed in such a way that there are low correlations between (main) attributes. The survey design process will be elaborated on in Chapter 5.

One drawback of SP data is that it is prone to hypothetical bias. This phenomenon is described as the difference between observed behaviour and stated behaviour. Studies, such as the often-cited paper by Brownstone and Small (2005), have shown that there can be significant differences in the Willingness to Pay values derived from RP and SP studies. Hensher (2010) gives several suggestions, such as using a reference alternative and clearly explaining study objectives, for survey design to limit hypothetical bias.

Additionally, M. Ben-Akiva, McFadden, and Train (2019) present a list of items that require attention when designing SP experiment. This to ensure that from an economist viewpoint the outcomes are useful and can be used to assess the Willingness to Use night trains for European long-distance travel. The following list details how in this study these points are addressed:

- **Familiarity**, respondents need to be familiar with the night train concept. Because this is currently a non-existing alternative in the Netherlands, respondents will be introduced to the concept of travelling by night train. Care is taken to minimise manipulation.
- **Sampling, Recruitment, Background** addresses concerns regarding the sample. Information about socio-demographics and other background variables will be collected to check the composition of the sample.
- **Menu design** highlights points regarding the design of the survey, such as the number of alternatives, attributes per alternative and realism of attribute levels. To make sure the survey is well designed, choice sets are generated using Ngene and dedicated survey software is used from SurveyGizmo.
- **Attribute formatting**, the choice set will be limited to 3 alternatives and does not include more than 6 attributes for both the perceived comfort rating experiment (see Section 3.2.1) and the main mode choice experiment. The attribute values mimic market conditions that decision-makers could encounter in real life.
- **Calibration and testing** recommend testing and calibrating the study results to consumer behaviour in real markets, so using RP data to compare the SP study results. However, due to the lack of RP data, this will be impossible. Therefore, one should aim to use a model that fits the data best.

3.2.1. Perceived comfort rating experiment

Comfort of travelling by night train is complicated to determine, several different attributes are expected to contribute to night train comfort and the weights of these attributes are different for each individual. One can say that each individual has their own 'perceived' comfort level of the night train travel alternative. If one would present all the attributes contributing to comfort at once in the main choice experiment, respondents could potentially only focus on a limited set of attributes or place too much weight onto the comfort attributes, as they are over-represented. The general notion is that decision-makers can only cope with a maximum number of seven attributes in choice experiments (Johnson & Orme, 1996).

As a solution for choice experiments where many more attributes play a role, such as in our case, the Hierarchical Information Integration (HII) theory was introduced by Louviere (1984). The theory is based on the assumption that decision-makers when faced with many attributes in a choice task, categorise them in so-called 'decision- constructs'. Individuals first trade-off the attributes that belong to each specific decision-construct in a sub experiment, after which a rating is awarded to that decision-construct. In the overall choice task, individuals are then presented with the different scores on the decision-constructs in a 'bridging experiment'.

This study uses an adapted version of the classical HII concept. Only one sub experiment (decision-construct) will be used, which determines the 'perceived comfort' of the night train alternative. Consequentially, because of the usage of only one sub experiment, the main mode choice experiment is also not a true 'bridging experiment'. Estimating a rating model for the sub experiment allows for predicting a 'perceived comfort' score, which can be used in the main choice experiment. The main mode choice experiment will show the 'perceived comfort' attribute among other attributes that were defined. This approach has been applied before, mainly in the transportation-related choice experiments. Bos, Van der Heijden, Molin, and Timmermans (2004); Molin and van Gelder (2008); Norojono and Young (2003) are all examples of studies that included decision-construct attributes, next to conventional attributes in the bridging experiment. Earlier research suggested that no decision-construct is necessary for the total trip time and trip cost (Bos et al., 2004). This same approach will be taken in this study. A conceptual overview of the applied approach is shown in Figure 3.1. As can be seen in Figure 3.1, in the 'perceived comfort' rating experiment, perceived comfort is the dependent variable, whereas in the main mode choice experiment 'perceived comfort' is an independent variable.

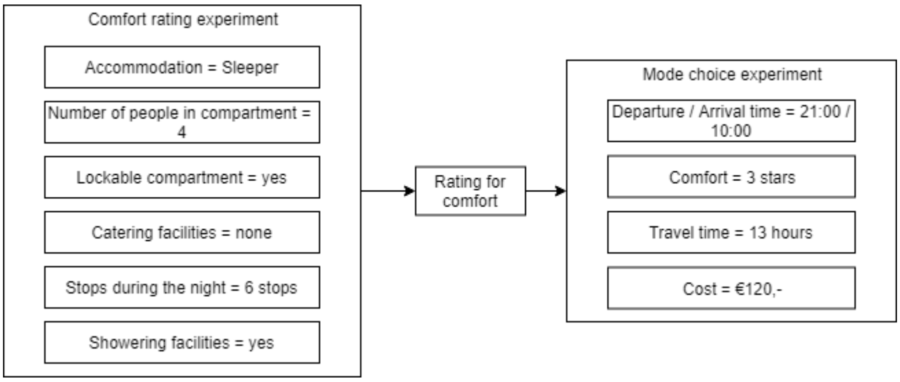


Figure 3.1: Conceptual overview of applied approach, showing both the 'perceived comfort' rating experiment and the main mode choice experiment. Note that in the figure only one alternative is shown

The remainder of this study will refer to the perceived comfort rating experiment as 'comfort rating experiment' and the choice experiment including, among others, the 'perceived comfort' attribute as the 'mode choice experiment'.

3.3. Discrete choice modelling

This section discusses the general methodology of Discrete Choice Modelling in subsection 3.3.1. This is followed by a discussion on various model types that will be used for this research in subsection 3.3.2.

3.3.1. Introduction to Discrete Choice Modelling

Discrete Choice Modelling (DCM) can be used to derive insights into the trade-offs individuals make and to estimate choice models for travelling (McFadden, 1972). Most choice models use the 'Random Utility Model' decision-rule to model decisions (M. Ben-Akiva & Bierlaire, 1999). Part of the econometric paradigm, see Section 4.2. The theory assumes that the preference of an individual decision-maker is captured by a certain value (called utility) and the decision-maker chooses the alternative that has his/her highest utility (McFadden, 1972). Most transport attributes have negative parameters (such as for travel time and travel cost). Therefore these models are also known as disutility minimisation models. In mathematical terms, this principle can be written as (note that the following formulas can also be written with a specific subscript for each individual n):

$$U_i > U_j, \quad i \neq j \quad \forall i, j \in Alternatives \tag{3.1}$$

Equation 3.1 describes that the decision-maker prefers alternative i if its utility (U_i) is larger than the utility of alternative j (U_j). This equation holds for all i and j in the choice set of Alternatives, while i and j cannot be equal. This formula assumes that there is perfect information and decision-making is completely rational. However, when analysing the data, only observed choices of the individuals are available, there simply is no perfect information at hand. Therefore, uncertainty has to be taken into account. To model this uncertainty, the utility is mathematically formulated as a deterministic part V_i and a random component ϵ_i which captures the uncertainty, see Equation 3.2.

$$U_i = V_i + \epsilon_i \quad \forall i \in Alternatives \tag{3.2}$$

The decision-maker evaluates all the attributes of the alternatives in the choice set. Each attribute is awarded a specific weight factor for that individual. When all parameters are considered to be linear, the deterministic part of the utility function can be formulated as shown in Equation 3.3.

$$V_i = \sum_{k=1}^{k} \beta_k \cdot X_{ik} \tag{3.3}$$

The equation shows that the deterministic part of the utility is the summation over all attributes k. Every alternative i has attributes of level X_{ik}, which are multiplied with the attribute specific weight factor β_k.

3.3.2. Discrete Choice Models

This subsection introduces several types of Discrete Choice Models that can be estimated. A basic model is the Multinomial Logit model (in short: MNL model), this is also the most used model. The mathematical formulation of the MNL model is shown in Equation 3.4. The model parameters are estimated using Maximum Likelihood estimation.

$$P(i|C) = \frac{e^{V_i}}{\sum_{j=1}^{j \in C} e^{V_j}} \tag{3.4}$$

Equation 3.4 shows the probability that a decision-maker picks alternative i from the choice set C. This is formulated as a function of the deterministic utility V and e is the base number of the natural logarithm. This deterministic utility of alternative i, V_i is divided by the sum of the deterministic utilities of all alternatives j in choice set C.

However, to arrive at this formulation some assumptions and simplifications are made. It is assumed that the error term is *independent and identical Gumbel distributed (i.i.d.)*, sometimes also referred to as Type 1 extreme value distributed (M. Ben-Akiva & Bierlaire, 1999). In other words, this means that the error term consists of unobserved factors that are uncorrelated over all alternatives and observations, and they all share the same variance. (C. Chorus, 2017).

Furthermore, the MNL model holds the Independent from Irrelevant Alternative (IIA) property (C. Chorus, 2017). These two properties are an issue if alternatives share unobserved factors, for example in an experiment that includes multiple public transport options versus a car option. In that case, the public transport alternatives share unobserved factors such as an individual dislike to public transport. This can result in the random errors of these alternatives to be correlated, leading to bias. By introducing a shared term for the alternatives that are correlated, the so-called Nested Logit model, the issue can be dealt with. In the case of this thesis, examples of possible nests are the two aeroplane alternatives or the night train and 'morning plane' both arriving in the early morning.

Another issue in traditional MNL is the ignoring of heterogeneity of the attribute weights (β_k) among individuals. This means everyone is assumed to have the same tastes, while in reality, this is likely to vary. By including an additional parameter reflecting the degree of unobserved taste variation of the attribute (σ_k) , one can account for this effect.

A fourth issue with the conventional MNL model is regarding observing multiple decisions from an individual. Someone who has a dislike for trains will express this mindset in every decision taken in the choice experiment. The conventional MNL model ignores the correlation between choices made by the same individual, ignoring the fact that someones' taste and preferences are more or less stable over time, this is called the 'panel' effect. As a result, the standard model overestimates the available information in the dataset, leading to biased results and underestimated standard errors (C. Chorus, 2017).

The Mixed Logit (ML) is a highly flexible model that can overcome these issues. The model can account for correlations between error components (the nesting) by including an additional error term (C. Chorus, 2017). By including additional σ estimates for both preferences and tastes one can include the unobserved heterogeneity in respondents. Furthermore, by considering the sequence of choices of an individual when estimating the model one can deal with the issue of ignoring the correlation between observed choices (C. Chorus, 2017). Results show that the overall model fit is better for models that include the panel effect (C. Chorus, 2017).

Mathematically, the panel-ML model is shown in Equation 3.5. Notice that the integral of the formula is outside the probability function. This reflects that the Panel-ML model searches for model parameters v_n, β_n, that make observing the combination of choices (of an individual) most likely.

$$\int_{v_n,\beta_n} \left(\prod_{t=1}^{T} ((P_{ni}^{t} | v_n, \beta_n) \cdot f(v_n, \beta_n)) dv_n d\beta_n \right) \tag{3.5}$$

One drawback of the ML method is that the integrals are open form. Therefore the model has to be estimated using simulation and for each simulation, a set of σ's has to be drawn from a distribution, the choice probabilities have to be evaluated and this needs to be repeated many times to arrive at the average, unconditional choice probabilities (C. Chorus, 2017). Therefore the computation time will increase considerably.

This study will take the MNL model as a reference model. It will be used to confirm if a more extended model does indeed fit the data better. Nested Logit will be used to see if the alternatives share unobserved factors. Furthermore, ML models will be estimated because they are more flexible and allow for accounting for the panel effect that is expected to be present due to the usage of SP data.

Additionally, one of the research questions is if there are several groups or segments present in the mode choice data. A latent class choice model is suitable for answering this question. The resulting choice model has several classes. Each of these classes could have a different set of parameters, resulting in different choice models. In the end, this means the overall model could explain more choices and fit the data better. An advantage compared to the also estimated Mixed Logit models, is that it does not need simulation to estimate the choice model. This means computation time is significantly reduced. Furthermore, the class membership function can incorporate various socio-demographic and other background variables. This provides the sought useful information as to which groups are most likely to use the night train and, in that way, investigates the choice behaviour from a marketing perspective. For this reason, a latent class choice model is also estimated.

3.4. Summary

Main takeaways

- Stated-Preference data will be collected using a survey.
- The survey uses an adapted HII set-up to combine the comfort rating and mode choice experiments.
- Discrete Choice Modelling will be used to analyse the survey data. Several models will be estimated, including a panel-ML model. A latent class choice model will be used to identify market segments in the dataset.

This chapter provided an elaboration on the methodology that will be applied. For collecting data, Stated-Preference data will be collected using a survey. Some items that require attention were identified to ensure the data is useful from an economist viewpoint and to limit the hypothetical bias. The study will use an adapted form of the Hierarchical Information Integration theory to combine both the comfort rating and mode choice experiment. In the comfort rating experiment 'perceived comfort' will be the dependent variable. In the mode choice experiment, this is an independent variable. For analysing the data, Discrete Choice Modelling will be used. A basic MNL-model will be applied as a reference model. The more advanced panel-ML model will be applied to deal with the limitations that are related to the MNL model. A latent class choice model will be estimated to get insights into the different market segments that are most willing to use the night train.

4

Theoretical framework

This chapter introduces the theoretical framework for the choice model. Section 4.1 argues why the mode choice step in the larger framework of travel behaviour is chosen. Similarly, Section 4.2 introduces from which perspective the behaviour will be researched. This is followed by the introduction of the variables that will be included in the model in Sections 4.3 - 4.6. Section 4.4 presents the results of the focus group meetings. Finally, the framework is presented in Section 4.7 and a summary in Section 4.8.

4.1. Modelling travel behaviour

The Willingness to Use night trains should be considered in a broader context. Modelling travel behaviour generally consists of four steps, which are shown in Figure 4.1 (Ortúzar S. & Willumsen, 2011). This so-called ’4 step model’ provides the basis. However, it is recognised that individuals might take decisions in a different order, combine steps or add additional steps, such as a time of day choice. Noland and Small (1995) presents a travel choice model that includes departure time choice. Including this step is proven to perform better for both work and leisure trips (Bhat, 1998).

The interest of this thesis is to find out if people are willing to use night trains, instead of taking the aeroplane, for European long-distance travel. This will be modelled by considering only the ’mode choice’ step. This is chosen as it is expected that the introduction of the night train travel option influences this step from the ’4-step model’ to the largest extent. Furthermore, the time component will be taken into account. The night train offers the possibility to leave your origin in the evening, travel at night and arrive in the morning at your destination and save on hotel costs. This will be compared to travelling by aeroplane which would mean travelling in the early morning or travelling the day before and spending the night in a hotel. Travellers are likely to make a trade-off between this departure/arrival time, mode and costs. This thesis will take the time aspect into account by using contexts with different arrival times.

4.2. Different perspectives of travel behaviour

Several perspectives can be used for looking at travel behaviour. Two examples are the ’econometric perspective’ and the ’marketing perspective’ (Anable, 2005; Arentze & Molin, 2013). The econometric perspective assumes travel behaviour is the result of a choice process. It uses Discrete Choice Models, which are based on mathematical equations to describe this choice behaviour. This makes the econometric perspective a very intuitive and powerful tool for predicting demand for new travel options. Additionally, because the model uses a mathematical way of describing choice behaviour, it is clear that varying the attribute levels influences the observed choices made, this means the causality direction is clear. Furthermore, Kroesen (2018) points out that the theoretic foundation is reasonable and consistent with welfare theory. The theory behind this paradigm is utility maximisation, also known as disutility minimisation. It assumes decision-makers use a decision rule that implies choosing the alternative that minimises their disutility, which in travel context is, for example, a combination of travel cost and time. This is called a Random Utility Model (RUM). While there are more decision-rules available,

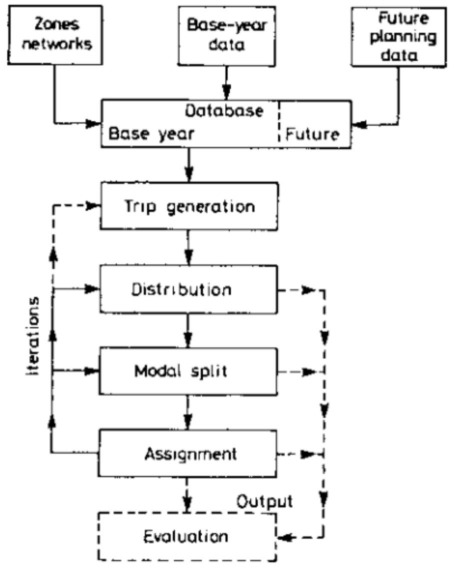


Figure 4.1: 4-step model, from Ortúzar S. and Willumsen (2011)

such as the Random Regret Minimisation (RRM), work from C. G. Chorus, Rose, and Hensher (2013) showed that when comparing both models, the differences in the performance can be very small. Additionally, limitations in the availability of software mean it is harder to design a RRM choice experiment and requires a completely different set-up. Combining this, it is chosen to make use of the more traditional RUM based econometric perspective (which has been used in most studies) for answering the research questions of this thesis.

Secondly, it is interesting to see if user groups can be identified. This allows the company (and clients thereof) to get to know more about the potential travellers and adapt marketing to these specific user groups, essentially this is an application of the 'marketing perspective' described by Anable (2005). An added benefit of including socio-demographics in the model is that it explains part of the unobserved heterogeneity and gives a better model fit (M. Ben-Akiva & Bierlaire, 1999).

4.3. Mode choice attributes included in the model

To assess the Willingness to Use night trains for European long-distance travel, this study uses a choice model which needs to be set up. Literature research was done to find out which variables are often included in mode choice experiments. The following attributes were found in completed studies and elaborated on if they are included in this study. If included, an indication is presented for the expected parameter sign.

- **Trip time** is almost always included in studies into the mode choice (see for example Hensher and Rose (2007); Morikawa, Ben-Akiva, and Mcfadden (2002); Paulssen, Temme, Vij, and Walker (2014); Román et al. (2010)). It differs for each study if this attribute refers to the complete trip time (from origin to destination) or, in the case of shared transport, if this is split between the in-vehicle time of the main mode and access and egress time. This study will include this attribute, considering the complete trip time. A division between main travel time, waiting time and egress time will be shown in the choice set. Following previous studies, the effect is expected to be negative, as more travel time, in general, is unfavourable.

- **Access / Egress time** is often included in studies about public transportation (see Hensher and Rose (2007); Morikawa et al. (2002); Román et al. (2010)). Sometimes a variant is used called access distance, such as in Paulssen et al. (2014). For assessing the Willingness to Use the night train it is assumed all modes depart from the same location. This means access times are assumed equal and therefore not taken into account. Egress time will be taken into account to make sure a fair comparison between modes is made, for example including the trip from the airport to the city centre, whereas the night train is assumed to arrive in the city centre.
- **Waiting time**, Hensher and Rose (2007) include the waiting time at the public transport location. This attribute will be incorporated in the division shown under 'trip time'.
- **Trip cost** is, likewise, also almost always included in studies into the mode choice (see Hensher and Rose (2007); Morikawa et al. (2002); Román et al. (2010)). This refers to the cost of making the trip to reach the destination. This attribute will be taken into account, but similar to the travel time attribute a division will be shown. For example, travelling by aeroplane the day before results in an additional hotel night and requires a transfer from the airport to the city centre both resulting in additional expenses. Similarly to travel time, the expected parameter sign is negative.
- **Departure / Arrival time**, Bhat (1998); de Jong et al. (2003) both write that including the time of day choice in a model improves the fit. The time aspect will be taken into account by using contexts, in which respondents are faced with different arrival times. These arrival and departure times will be shown to the respondent in the choice task. This context variable is taken into account as an interaction effect on the constant. It is expected that an earlier arrival time will have a positive effect on the choice for the night train.
- **Frequency** often mentioned as one of the most important factors for the usage of public transport (and related to waiting time). Román et al. (2010) also uses the factor for analysing the demand for a high-speed corridor between Barcelona and Madrid. However, as this study already uses a set arrival time and concerns long-distance travel, the frequency is of less importance. Furthermore, the frequency of a night train is rarely more than once a day. Therefore it is decided to not take this attribute into account.
- **Reliability** is included in the study by Román et al. (2010). It can indeed be important for the mode choice and delays in flights and trains are not uncommon. However, as the frequency of the night train is low and a percentage score is hard to interpret for decision-makers, it is chosen to exclude this attribute from the study.
- **Number of transfers**, Morikawa et al. (2002) specifically includes the number of transfers for the trip. Both night train and aeroplane are assumed to depart from the same location, it is assumed that each trip starts there. Therefore a transfer between access and main mode is not required. However, the aeroplane arrives at the airport and always requires a transfer to the city centre, therefore the main mode to egress mode transfer is relevant. Furthermore, only direct (non-transfer) trips for the main mode will be considered. Information about the need to transfer will be given in the context, but not varied in the choice experiment.
- **Comfort** is included in a study by Román et al. (2010). They found that comfort does play a role and even that there is an interaction between travel time and comfort (although the t-test value proved to be 1.8). As introduced, the night train has varying service levels. Therefore, it is interesting to see to what extent the comfort level influences the stated preference for the train, so it will be included. However, as the comfort level for each individual differs, this will be included as **Perceived comfort**, which will be assessed separately using the aforementioned comfort rating experiment. More information about which determinants of this 'perceived comfort' can be found in Section 4.5. It is expected that a higher night train comfort level will have a positive effect on the utility.
- **Car related attributes** several studies include the car as a travel mode, and in addition to presenting (some of) the attributes above include specific car-related attributes such as daily parking cost (Hensher & Rose, 2007) or toll costs. However, as only shared modes are used for this study, car specific attributes are not taken into account.

4.4. Results of focus group meetings

The literature review in the previous section provided an overview of often included variables in similar studies. However, no studies have been done that specifically focuses on the night train and mode choice. To address this problem focus groups were organised. The goal of these focus groups was to see if specific attributes that play a role for night train mode choice were missed. Additionally, as it is unknown what contributes to night train ’perceived comfort’, another goal was to identify determinants of this perceived comfort.

In total 3 focus groups were organised, with in total 19 participants. The participants of the focus groups could roughly be divided into three categories, one group consisted of students enrolled in a transport-related study at TU Delft, another group consisted of students enrolled in other programmes of TU Delft and the last group was organised at Royal HaskoningDHV in which members of the transport advisory group attended.

Participants were asked several questions. First, they were asked about the advantages and disadvantages of the night train. Followed by a question about what attributes they would consider when making a choice (from a broader choice set). Next, two questions were asked one about the comfort of the night train and another about the convenience of the night train.

Participants first wrote down their answers to these questions on paper and afterwards engaged in a discussion. After the focus groups, their written answers were reviewed and ’frequencies’ of the written down answers were noted, these can be found in Appendix C. Reviewing the results of the focus groups did not reveal additional completely new attributes, not mentioned in the previous section, that should be taken into account for mode choice. In general two thinking patterns could be identified, people that would absolutely not choose the night train, either because they think it is way too expensive or the longer journey time puts them off. Other people responded that they think it is a relaxed and beneficial way of travelling as they still need to sleep, can be moved during sleeping, and arrive rested in the morning. Important notions mentioned were the possibility to save on hotel costs, arrive in the destinations city centre and having more space to move during the journey. Concerns included the low frequency (’you travel at that time or you do not travel’), the unfamiliarity with booking international (night) train tickets, and concerns about sleep quality.

Comfort is determined by the quality of the accommodation, the ’inside environment’ which consists of attributes such as vibrations, sound level, temperature and lighting control. Other items contributing to comfort are the possibility to get food and drinks, sanitary facilities and privacy. The last one includes both with how many people the compartment is shared as well as if the compartment is lockable. It was also mentioned that women might put more weight on sanitary facilities. Another benefit mentioned compared to the aircraft was the additional space to move in the train. Surprising was that entertainment was thought of to be positive for yourself, but negative if other people in your compartment would use it without headphones.

Participants were also asked about what determines the convenience of a travel mode. Convenience was defined as ’everything outside of the train’, whereas comfort was about ’the inside of the train’. Useful insights were gained on how convenience is determined. Participants indicated that night trains have less restrictive baggage regulations, have less waiting time and avoid the burden of airport security. Furthermore, they said the night train is likely to stop in more locations to bring you closer to your destination or board closer to your home. Next, they indicated that the booking and information system on how to get a ticket is very important. And at last, that a comparison website where you can make a fair comparison between travelling by train and plane would be useful. One remark is made, some participants seemed a little confused between ’comfort’ and ’convenience’ as they clearly mixed up the items or wrote down the same twice.

4.5. Determinants of 'Perceived comfort'

Perceived comfort can be determined by various attributes. The importance of these attributes vary for each individual, e.g. one might give a lot of importance to sanitary facilities, while others mostly care about the inside environment of the transport mode, leading to different scores. Input from the focus groups is used in determining which attributes to include in the comfort rating experiment.

- **Accommodation type** this attribute is related to the physical accommodation during the journey. Román et al. (2010) also uses this attribute, but calls it 'comfort', as in that study it only relates to the width of the seats and the legroom. As introduced in Section 2.2, night trains often offer three types of accommodation: seats, couchettes and sleeper carriages, which will be used in the rating experiment. It is expected a 'nicer' accommodation results in a higher 'perceived comfort', therefore the parameter is expected to be positive.

- **Privacy** seems to be an important notion as was pointed out by the focus group. This relates to the number of people in your compartment, which can vary from 6 (in seats or couchette class) to 2 (in sleeper class), and the possibility to lock your compartment. This will be taken into account for the study. For the number of people, the sign is expected to be negative (as it means less privacy), for the possibility to lock your compartment it is expected the parameter is positive (as a lock means more privacy/security).

- **Catering**, the possibility to get some food or drinks during your journey in a night train was mentioned often in the focus groups. Therefore the effect of offering different onboard facilities for food and drinks will be examined in the comfort rating experiment. Similarly to the accommodation type attribute, it is expected that the parameter sign is positive.

- **Stops**, the number of stops during the night was not mentioned that often by respondents taking part in the focus group. However, because the number of stops will influence the sleep quality (and therefore comfort), it will be included in the comfort rating experiment. It is expected the parameter sign is negative.

- **Inside environment**, respondents taking part in the focus group highlighted several smaller notions that can be summarised into 'inside environment'. This includes the possibility to switch off lights for sleeping, or adjusting the temperature through heating or air-conditioning. Another remark made was about the sound level and vibrations of both modes. Although these characteristics do not vary among service levels, the information will be given about the 'inside environment', which can influence the constant for travelling by night train.

- **Facilities** were mentioned in the focus groups to be of importance. This, for example, includes a power socket and provisioning of WiFi. The possibility to shower was also mentioned. In the rating experiment, the possibility to shower will be included. It is expected that this contributes positively to the user experience. Some other information such as the provisioning of WiFi will be mentioned in the general information, but not varied in the experiment.

- **Staff** some people voiced that onboard staff influences perceived comfort. However, as this is very subjective (more staff onboard the train, compared to the aeroplane, does not necessarily mean better service and without information about the staff training is hard to interpret for people). Therefore it is decided to not include this as an attribute in the comfort rating experiment.

4.6. Socio-demographics and trip characteristics

Including socio-demographics into the survey and model gives more insight into the composition of the sample of respondents. Furthermore, using the collected socio-demographics insights can be gained upon different user groups and market segments are identified (part of the marketing perspective). M. Ben-Akiva and Bierlaire (1999) writes that including socio-demographics is needed to explain the heterogeneity of individuals when estimating choice models and therefore can improve model fit. Socio-demographics can directly influence the utility and/or can influence the utility indirectly via interactions with variables. The following socio-demographics were found to be often included in literature.

- **Age** is a common socio-demographic to collect, and therefore is included in the study (Buehler, 2011; Hensher & Rose, 2007; Limtanakool, Dijst, & Schwanen, 2006b; Paulssen et al., 2014; Román et al., 2010; Vredin Johansson, Heldt, & Johansson, 2006). It is expected that older people will choose more often for the night train (it is expected that the longer travel time is of less importance).

- **Income** is also an often collected socio-demographic. It differs to some extent in which way it is collected, some studies ask about dispensable income, while other studies collect the gross income (Buehler, 2011; Hensher & Rose, 2007; Limtanakool et al., 2006b; Paulssen et al., 2014; Román et al., 2010; Vredin Johansson et al., 2006). As it is indeed likely that the income level will influence the long-distance mode choice, it will be considered in this study.

- **Gender** is likewise often included in mode choice studies (Buehler, 2011; Hensher & Rose, 2007; Limtanakool et al., 2006b; Paulssen et al., 2014; Román et al., 2010; Vredin Johansson et al., 2006). Also in this study information about gender will be included. Additionally, this information can be used to test and verify if females are indeed more cautious about travelling at night (as was found by Vredin Johansson et al. (2006)).

- **Education level** is also included in models. Limtanakool et al. (2006b); Vredin Johansson et al. (2006) both include it in their study. Limtanakool et al. (2006b) states that there was previously no literature on the influence of education level on medium- and long-haul trips, but found the effect to be significant. As it appears to be an important determinant for mode choice, it will be included in this study. It is expected people with a higher education level, would choose the night train more often.

- **Household size and composition** is sometimes included in studies (see Buehler (2011); Hensher and Rose (2007); Limtanakool et al. (2006b); Vredin Johansson et al. (2006)). However, as to limit the amount of data to be collected and the study is about individual trips, household size and composition is not taken account in this study.

- **Work status**, Hensher and Rose (2007) does collect this socio-demographic data in their study for checking the sample representation, it does not incorporate it in model specification. This study does not focus on daily commuting trips, questioning the usefulness of the variable. However, the variable may help to check sample composition and provide useful insights from a marketing perspective. Therefore, it is chosen to collect this information.

- **Car availability** is included in studies by Buehler (2011); Limtanakool et al. (2006b). However, as it was decided to only include shared transport, this is ignored from the study.

In addition to socio-demographics, other characteristics can be included as well. This includes the trip purpose, trip distance and the current usage of travel modes.

- **Trip purpose** is included in a study by Buehler (2011); Román et al. (2010). For this study it is thought of to be a quite important determinant of mode choice, therefore both leisure and business purposes are included in the study. This will be done by including it in the trip context that will be shown to respondents. The purpose can be included directly or via interactions. It is expected that travellers with a business purpose are less willing to use the night train or morning plane.

- **Trip distance** is sometimes included in a study, for example by Burgdorf, Eisenkopf, and Knorr (2018) who did a study into the mode choice for long-distance trips in Germany and varied the distance from 300 to 900 kilometres. However, given the introduced context of a trip of about 12-14 hours, similar to distances such as Amsterdam and Vienna/Milan, it will not be included in this study.

- **Preferred main travel mode** is included in a study by Hensher and Rose (2007). It is interesting to see if the willingness to use the night train depends on the experience one has with travelling by train. It is expected that people with more experience would express a higher preference for the night train. This information will be collected in the study.

4.7. Framework

The different variables that are mentioned in the previous paragraphs are included in the theoretical framework. A graphic depiction of this framework, which will be used to set up the Discrete Choice Model, is shown in Figure 4.2. In the figure, square boxes refer to observable variables, such as socio-demographics, variables regarding the travel alternatives and the stated mode choice. The oval boxes refer to the unobservable variables, which are the utility levels of the alternatives. Dotted boxes refer to variables used in literature, but excluded in this study. Black lines reflect the main variable effect on utility and comfort rating. The green line reflects socio-demographics/background variable effects and interactions on the comfort rating. Red lines indicate the effect of background/trip characteristics on the utility of the different alternatives and blue lines the effect of interactions on the utility of alternatives.

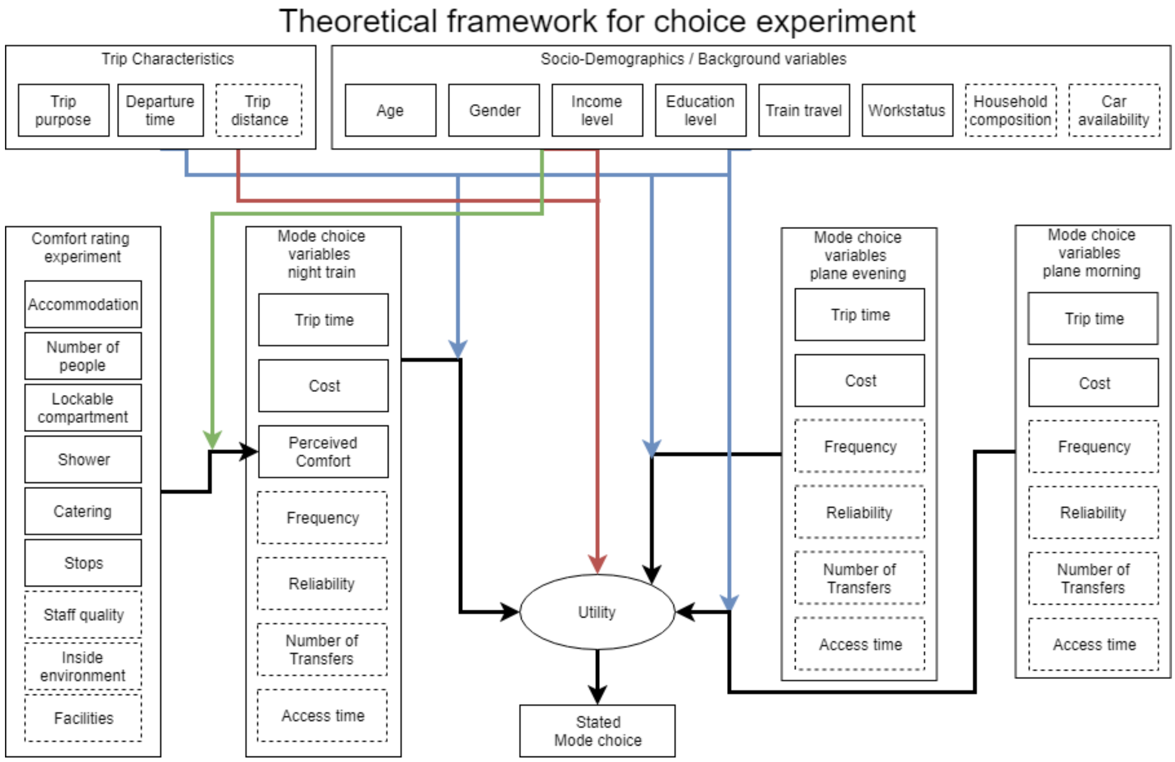


Figure 4.2: Graphic overview of the theoretic framework

4.8. Summary

Main takeaways

- Only the mode choice step of the traditional '4-step model' will be considered, the 'econometric' perspective is mainly used for analysing the data.
- Focus group meetings did not reveal any new, previously unknown, attributes that should be taken into account for the night train in mode choice.
- Varied attributes for the mode choice experiment are travel time, travel cost and 'perceived comfort'. Arrival time is considered in the context.
- Focus group meetings did reveal that for 'perceived comfort': the number of passengers in the compartment, accommodation type, possibility to lock the compartment, showers, catering facilities and the number of stops are considered for the night train comfort level and are therefore included in the comfort rating experiment.

This section introduced the theoretic framework for the model. This study uses the mode choice step from the traditional '4-step model', and includes the time dimension into the context that will be included in the survey. The 'econometric' perspective will be used to analyse travel behaviour. Additionally, parts of the 'marketing' perspective will be used to see if user groups can be identified. For this, information on several socio-demographics will be collected. Literature research was done to find out which variables should be included in the mode choice model. It is chosen to include: trip time, perceived comfort and trip costs. The arrival/departure time is included in the survey context. Focus groups did not reveal any new night train specific variables that should be included in the mode choice. They did help in determining attributes for 'perceived comfort'. Included in the rating experiment will be accommodation, the number of people in the compartment, the possibility to lock your compartment, catering facilities, the possibility to shower and the number of stops during the night.

5 Survey design

The chapter discusses the survey design process. First, Section 5.1 will introduce the survey context that will be used to make sure respondents limit the usage of own assumptions. The alternatives that are included in the mode choice experiment and the process of travelling is described in Section 5.2.1. The attribute levels for both the comfort rating experiment and the mode choice experiment are introduced in Sections 5.2.2 and 5.2.3. The generation of the choice sets for both experiments is discussed in Section 5.3. The survey construction, which discusses the set-up of the survey, is presented in Section 5.4. This chapter concludes with a concise summary in Section 5.5.

5.1. Survey context

A clear and complete context prevents respondents from making additional assumptions about 'missing' information to make their choice, resulting in data that is of less quality. Therefore, several decisions are made regarding this context that will be presented to respondents.

The research questions state that the study is about finding out what the willingness of people is to travel by night train, as an alternative for flying, for long-distance trips in Europe. For this reason, it is considered important that respondents are familiar with long-distance trips in Europe. Therefore they are asked if they have travelled outside the Benelux-countries in 2018, and for which purpose this trip was made. If respondents made multiple trips, they are asked to recall their last trip. Additionally, they are asked if they travelled alone or with someone else. These questions are used to compose a travel context that is familiar to the respondent.

As written in the introduction, this study focuses on long-distance trips. For night-trains, this is about 12-14 hours and reflects (using conventional rail) distances of 1000-1200 km. In the survey, this will be illustrated as trips between Amsterdam - Vienna/Milan.

It is expected that the arrival time at the destination influences the mode choice. Therefore, respondents will be presented with 2 different contexts (08:30 and 10:30). To further illustrate this, respondents are told that they hold tickets for a highlight (such as a museum, city tour etc.) or have a business appointment in the city centre (depending on earlier travel purpose question).

The main interest of this study is to see if respondents are willing to use the night train. This means that the research focuses on the main mode chosen for a trip. Therefore the access leg to the train station is ignored; respondents are told they can arrive on time at either the train station or the airport. Additionally, respondents are told the night train also stops at Utrecht and Arnhem. Inherently related to aeroplane travel is that it requires a transfer between the destination airport and the city centre, the night train arrives in the city centre and therefore does not need a transfer.

At last, to further illustrate the picture, respondents are told they are travelling with hand luggage only. Avoiding the considerations and uncertainties regarding handling large suitcases. This is the same for both leisure and business purposes.

5.2. Specification of the choice experiments

This section will introduce the alternatives in more detail (see Section 5.2.1), specify the attribute levels for both the comfort rating experiment in Section 5.2.2 and the mode choice experiment in Section 5.2.3.

5.2.1. Alternatives

To assess the Willingness to Use night trains, as an alternative for flying, the modes that need to be included are train travel and flying. It is chosen to use 3 different alternatives in the main mode experiment: night train, flying in the early morning and flying in the evening the day before. Other alternatives, like long-distance coaches, are excluded from the study as they are not considered to be a competitive alternative to aeroplane travel (comfort wise and the trip time is even longer than the train). This last notion was also mentioned during the focus groups. The following list provides more detail about the alternatives, the travel process is also illustrated in Figure 5.1:

- **Night train**, the night train departs from Schiphol train station. You only need to be there a short amount of time in advance. After boarding, the train drives to the final destination. Possibly stopping multiple times along the way, even in the middle of the night. After arriving the next morning at the central train station, the traveller can continue directly to its final destination.
- **Morning plane**, this alternative begins at Schiphol airport. It involves arriving in the early morning at the airport. After arrival, the usual airport routine (check-in, passport control, security check) has to be completed and the traveller has to wait until the departure time of the aircraft. After landing at the destination airport, the traveller has to disembark and find its way to the airport-city centre transfer. The transfer takes additional time and also arrives at the central train station. After the arrival at the central train station, the traveller can continue to its final destination.
- **Evening plane**, the major difference between this alternative and the previous one is the departure time. Instead of departing very early in the morning, the traveller departs the day before in the evening. The travel process itself is remarkably similar, the only difference is that the traveller needs to spend a night in a hotel. The next day, the traveller is always on time at its final destination. This last alternative is a base alternative and therefore has fixed attribute levels.

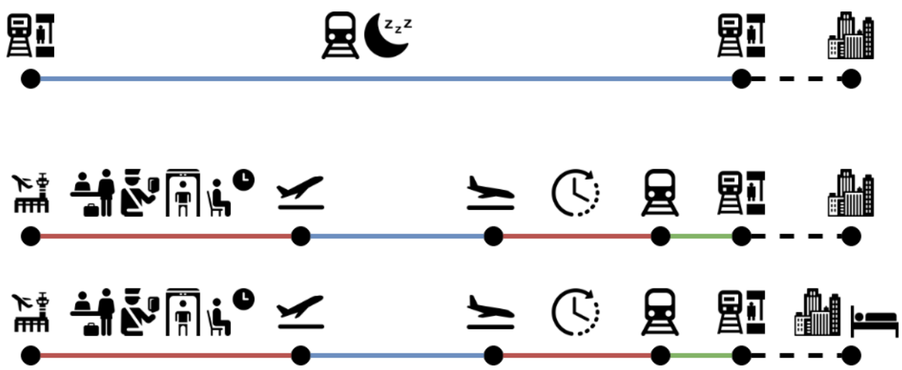


Figure 5.1: The process of travelling with different alternatives. Top: night train, middle: morning plane, bottom: evening plane.

5.2.2. Attribute levels comfort rating experiment

The attributes accommodation, the number of people in the compartment, the possibility to lock the compartment, catering facilities, the possibility to shower and the number of stops during the night were defined to be determinants of comfort. The comfort rating experiment will only be carried out for the night train alternative. This is because the main interest of the thesis is about establishing the comfort level of the night train, and the determinants thereof. There is less interest in the differences between perceived comfort of different flights, aeroplanes and airlines. The following list presents the reasoning for the chosen attribute levels, Table 5.1 summarises the attribute levels:

- **Accommodation:** the night train has three different kinds of accommodation (see Section 2.2): seats, couchettes and sleeper cabins. These categories will form the different levels of this attribute.
- **Number of people in the compartment**: this will be varied in the experiment between 2, 4 and 6 people. It is chosen to not tell respondents if they need to assume that the train is sold out (meaning all seat/beds are booked) or not. They are allowed to make their assumption on this fact.
- **Lockable compartment**: this will be varied between 'yes' and 'no'. If the compartment is lockable it means only the respondent and his fellow travellers in that compartment can enter the room.
- **Catering facilities:** three different kinds of catering facilities are included in the experiment. These are: no facilities (no possibility to buy food or drinks on board the train), a kiosk in the train offering some snacks or takeaway/small meals and a restaurant car. The last one refers to a dining car in which it is possible to have dinner/breakfast and sit down to eat it.
- **Possibility to shower:** this will be varied between 'yes' and 'no'.
- **Number of stops during the night:** refers to the number of stops the train makes during 00:00 and 06:00. Inspection of current timetables indicates this can be up to 6. Therefore levels 0, 3 and 6 are included in the experiment.

Table 5.1: The attribute levels for the comfort rating experiment.

Attributes	**Levels**		
Accommodation	Seat	Couchette	Sleeper
Number of people in compartment	2	4	6
Lockable compartment	Yes	No	
Catering facilities	None	Kiosk in train	Restaurant car
Possibility to shower	Yes	No	
Number of stops during night	0	3	6

5.2.3. Attribute levels mode choice experiment

The attributes to be included in the mode choice experiment were introduced in Section 4.3. These were trip time, trip cost and perceived comfort. The arrival/departure time is taken into account using context. The following list presents a reasoning for the chosen attribute levels, Table 5.2 summarises the attribute levels.

- **Trip time:** a distinction should be made between travel modes. First, the attribute levels of the train are explained. The distance between city pairs Amsterdam and Milan or Vienna by rail is respectively 1090 and 1175 kilometres (Rome2rio, n.d.). For calculations, this is averaged to 1132.5 kilometres. Using timetables from the ÖBB it is calculated that currently the average speed of a night train, including stops, is well below 100 km/h (between 70-80 km/h). Therefore, as improvements in service are possible, levels are based on 3 average speeds: 80, 90 and 100 km/h. This results in trip times of 11:30 (hh:mm), 12:45 and 14:00. Additionally, some waiting time is added. Before the trip, this is taken as 10 minutes, and after arriving as 5 minutes. The time after arriving reflects the alighting time at the train station. The time before is longer, to reflect that people arrive more time in advance, because of the consequences of missing the train. In total this gives trip times for the night train of 11:45, 13:00 and 14:15.

Trip times for travelling by aeroplane are based on current schedules. Investigation of airline schedules shows the flight time between Amsterdam and Vienna or Milan is between 01:40 (hh:mm) and 01:55. Two levels with 30 minutes difference are used. This is chosen as 01:45 and 02:15. Furthermore, the usual requirement to arrive at the airport 2 hours in advance of your flight is taken into account. Additionally, from the time of arriving at the gate to making your way towards the city centre transfer is assumed to take 45 minutes. At last, the airport city centre transfer itself is taken as 30 minutes (an average, Milan Malpensa airport towards Milan takes almost an hour, whereas Vienna Airport to Vienna takes 15 minutes). This means from aircraft touchdown to the central train station takes 1 hour and 15 minutes. In the end, this results in two levels for trip time: 05:00 and 05:30. The reference alternative is not varied and the trip time is fixed at 05:00.

- **Trip costs** as shown in Section 2.2, tickets for seats in the night train start at just €29,- and this goes up to €99,- for a bed in the deluxe sleeper cabin. However, prices can vary depending on season or demand. Therefore a wider price range is considered. It is opted to use four different levels: €40, €80, €120, €160. Furthermore, as yield management is highly complicated and sometimes strange discounts can be applied, no restrictions are set. Meaning a trip with a high comfort level and low price could appear in the choice set.

  For the aeroplane alternatives, prices for one-way flight tickets were investigated. Dates far in advance were inspected and prices were observed as low as €40 ranging to around €100. Equally as for the train, to reflect yield management, a wider range of three levels is chosen, these are €50, €100 and €150. The base alternative is fixed at €100. Also, €10 is taken into account for the airport city centre transfer. In the base alternative, costs for a hotel night are included. Average daily rates are €100,- (for Milan) and €140,- (for Vienna), a fixed average of €120,- is taken into account for hotel costs (PWC, 2016, 2018). In total, this results in 3 levels for the morning plane alternative: €60, €110 and €160. The evening plane alternative is fixed at €230.

- **Perceived comfort** is only included for the night train alternative. This will be displayed to respondents using a 5-star rating scale. Three levels are used, which reflect a very uncomfortable, average and very comfortable level, so 1, 3 and 5 stars. However, to prevent respondents from making own assumptions about aeroplane comfort, they are told to be travelling economy class. This assumption is valid as flying business class within Europe is in most companies not permitted.

Table 5.2: The attribute levels for the mode choice experiment.

Attributes	**Levels**			
Travel time night train [hh:mm]	11:45	13:00	14:15	
Travel time morning plane [hh:mm]	05:00	05:30		
Travel time evening plane [hh:mm]	05:00			
Travel cost night train [€]	40	80	120	160
Travel cost morning plane [€]	60	110	160	
Travel cost evening plane [€]	230			
Travel comfort night train [Stars]	1	3	5	

5.3. Generation of choice sets

Generating the choice sets that will make up the survey involves several choices that need to be made. The software tool Ngene is used to generate the choice sets (experimental design). The used Ngene syntax is included in Appendix B. The Ngene manual offers extensive guidelines for defining the computer syntax, which is used in making the decisions outlined below (ChoiceMetrics, 2018). No information is available on priors to generate efficient designs, therefore it is chosen to use the more 'conventional' designs. Of these 'conventional designs', a full factorial design is unpractical as it would lead to way too many choice sets. Therefore, a fractional factorial design is used. An orthogonal design is preferred as it searches for a design that is attribute level balanced (the attribute levels appear the same number of times in the choice sets) and in which no correlations exists between attributes.

The comfort rating experiment is only carried out for the night train. Therefore, a non-labelled experiment is chosen. This means the choice sets can be generated sequentially. Ngene is able to find an orthogonal design with 36 rows. Answering that many questions is too much for respondents, it would lead to fatigue. Therefore, it is chosen to make use of blocking. A new design is generated using 6 blocks, so every respondent gets 6 questions to rate their night train comfort level. Because some of the attributes are categorical like the catering facilities and accommodation, effects coding is used. Dummy coding could have been used as well, but when having multiple attributes that require dummy or effects coding, effects coding is the preferred approach (ChoiceMetrics, 2018).

For the main mode choice experiment, there are 3 different alternatives. There is one alternative with fixed attribute levels, that acts as a base alternative. The remaining two alternatives have varying attribute levels, therefore only combinations of these two alternatives need to be generated using the specialised computer software. The two alternatives reflect different transport modes, have alternative specific attributes and a varying number of attribute levels. Therefore it is needed to apply a labelled design, which means the choice sets are generated simultaneously. Furthermore, the comfort attribute is expected to be non-linear, therefore this attribute is effects coded. Additionally, an interaction effect is added between comfort and travel time. This is included as it is expected that the comfort level influences the perception of travel time, as was also found by Román et al. (2010). The software is instructed to find an orthogonal design with 36 rows. Again, to limit the number of choice sets a respondent needs to answer, the number of blocks is set to 6. Keeping the number of blocks the same between both experiments limits the difficulty in developing the survey.

5.4. Survey construction

5.4.1. Choice sets and questions

The actual survey is constructed using the generated choice sets. The survey is made with specialised software, SurveyGizmo. Every row in the experimental design is converted into a choice situation. For the comfort rating experiment, the choice set consists of one night train alternative. The question type is a star rating scale. Respondents are asked about their 'perceived comfort' level, in which 1 star means a 'very uncomfortable' comfort level and 5 stars a 'very comfortable' comfort level. An example of a question can be seen in Figure 5.2.

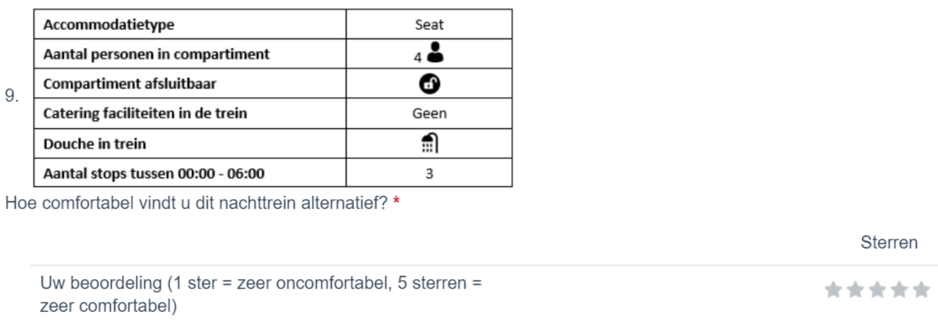


Figure 5.2: An example of a question for the comfort rating experiment, showing the choice set.

For the main mode choice experiment, the generated experimental design is converted into choice situations. The fixed reference alternative is added to the choice set. This means respondents are faced with the 3 different alternatives, of which 2 will vary in each choice set. Respondents are first asked to choose between the night train and morning plane alternatives. This is followed by a second question, asking them if they would choose this alternative or the evening plane alternative. An example of a question can be seen in Figure 5.3. As introduced, the survey uses 2 different context settings. Each context has 6 choice sets. Therefore, for the main mode choice experiment, respondents are faced with 12 choice sets.

47.

Kenmerken	Nachttrein	Vliegtuig	Vliegtuig + Hotel
Tijd aanwezig op station/luchthaven	17:45	02:30	17:30
Aankomsttijd op treinstation bestemming	08:00	08:00	22:30
Uw verwachte comfort niveau	★	Economy Class	Economy Class
Totale reistijd, waarvan:	14:15	05:30	05:00
- Tijd voor vertrek	00:10	02:00	02:00
- In trein/vliegtuig	14:00	02:15	01:45
- Uitstappen/overstap	00:05	00:45	00:45
- Luchthaventransfer	-	00:30	00:30
Totale kosten, waarvan:	€160	€110	€230
- Ticket	€160	€100	€100
- Luchthaventransfer	-	€10	€10
- Hotel	-	-	€120

Kiest u de nachttrein of reist u per vliegtuig in de vroege ochtend? *

- ◯ Nachttrein
- ◯ Vliegtuig in de vroege ochtend

48. Kiest u het zojuist gekozen alternatief of reist u per vliegtuig de avond ervoor, plus verblijf in hotel? *

- ◯ Het zojuist gekozen alternatief
- ◯ Vliegtuig de avond ervoor, plus verblijf in hotel

Figure 5.3: An example of a question for the main mode experiment, showing the choice set.

5.4.2. Other parts of the survey

The survey starts with several selection questions. If the respondent has not travelled outside the Benelux-countries (which means also not in Europe) the survey is terminated. Provided that the respondent does qualify for the survey, they will be asked some introductory questions. These questions are mainly used to build the travel context for the respondent. Furthermore, there are questions about how many times the respondent has flown or used an international train, and how long he/she thinks the journey to a city such as Milan/Vienna would take.

After the choice experiments have been completed, some additional questions will be asked. These are questions regarding socio-demographics and background. Respondents are asked about their gender, age, educational level, occupation, postal code and disposable income level. Finally, respondents are asked if they would like to participate in the voucher lottery. If they would like to, they are asked for their email address.

5.4.3. Survey testing

The first draft of the survey was tested with a handful of test respondents, which included non-transportation students. Their feedback consisted of several items, their most important comments were: a clearer division between the context settings, the usage of icons in the choice sets, choosing a single language (Dutch), and several suggestions about textual changes. This feedback was taken into account for designing the final survey. An illustration of the final survey structure can be seen in Figure 5.4.

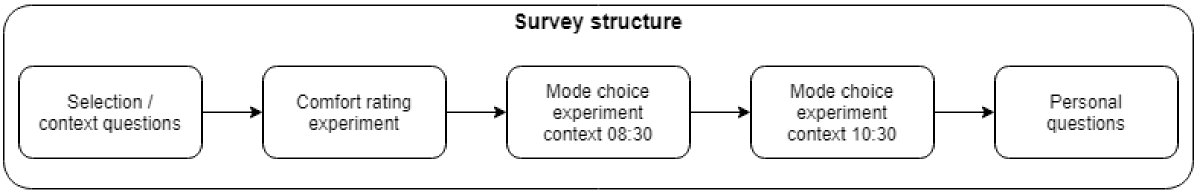


Figure 5.4: The structure of the survey respondents will face.

5.5. Summary

Main takeaways

- The survey context is defined as making a trip for 12-14 hours, similarly to Amsterdam - Vienna or Milan, for business or leisure with a preferred arrival time at either 08:30 and 10:30. The respondent is travelling alone or with someone else, and has hand luggage only.
- The choice set consists of the night train, morning plane and evening plane (+ stay in a hotel) alternatives.
- The survey consists of two main parts, a comfort rating experiment and a mode choice experiment.
- Respondents are asked 6 comfort rating questions and 12 mode choice questions.

This chapter detailed the process of designing the survey. First, the context that respondents need to take into account when answering the questions is elaborated on. It is specified as a trip between 12 and 14 hours, which is illustrated to be equal to Amsterdam - Vienna/Milan. People are travelling either for business or leisure, alone or with someone else. They are told to be travelling with hand luggage only. The mode choice will be asked in two different arrival time contexts, one at 08:30 and at 10:30. This is further portrayed by explaining that the respondent holds tickets for a touristic attraction or has a business appointment.

Furthermore, the choice experiments were specified: three alternatives are included in the choice set: night train, morning plane and evening plane (+ stay in a hotel), for each of them a travel timeline was presented. Attribute levels for both the comfort rating and main mode choice experiment were specified. The decisions that were made to generate the experimental designs using specialised software were highlighted, orthogonal fractional factorial designs are used. To limit the number of questions a respondent needs to answer, a blocked design is used. It is chosen to use 6 blocks.

The survey consists of 2 different experiments, the comfort rating and main mode choice experiment. For the comfort rating experiments people are faced with 6 questions, and for the main mode choice experiment also 6 for each context (12 in total). Additionally, some (selection) questions before and after the experiments are included. The survey was completed by a group of test respondents and feedback implemented.

6

Data analysis

This chapter elaborates on the procedure of data gathering in Section 6.1. Several socio-demographics and other characteristics of the sample are shown in Section 6.2. An overview of the recorded answers to several questions is presented in Section 6.3. This chapter concludes with a summary in Section 6.4.

6.1. Data gathering

A crucial step for answering the research questions is gathering enough completed responses. The research question is if people are willing to use the night train, as an alternative for flying, for European travel. For that reason, it was determined that the intended population was people who have recent (in 2018) experience with travelling outside the Benelux-countries (Belgium, Netherlands, Luxembourg). To ensure the sample came from this population, a selection question was included at the beginning of the survey. If a respondent did not have recent experience travelling outside Benelux-countries the survey was terminated. Another relevant question for data gathering is defining the minimum number of responses needed before analysis is meaningful. Orme (2010) presents a 'rule of thumb' to calculate the minimum number of respondents needed for Discrete Choice Modelling, shown in Equation 6.1.

$$\frac{n \cdot t \cdot a}{c} \geq 500 \tag{6.1}$$

In this equation n represents the minimum number of respondents needed, t the number of choice tasks, a the number of alternatives and c the maximum number of levels for any of the attributes included in the experiment. Substituting the known terms and rearranging the formula allows for calculating the minimum number of respondents, see Equation 6.2.

$$n \geq \frac{500 \cdot c}{t \cdot a} = \frac{500 \cdot 4}{36 \cdot 2} = 27.8 \approx 28 \tag{6.2}$$

This shows the minimum number of respondents is 28 when every respondent is faced with 36 choice tasks. This is simply too much for a respondent and therefore the experimental design was blocked in 6 blocks (see Section 5.3). Because of this reason, the minimum number of respondents that needs to be gathered is $28 \cdot 6 = 168$.

Data was collected in three ways: by distributing the survey through the customer panel of NS, by distributing flyers at Schiphol Airport and by sharing the survey link on social media platforms. The survey was open for responses from the May 27th until June 9th, 2019. In total 804 completed responses were collected. Cooperation with NS was sought. They were highly interested in the topic of this research and agreed to distribute the survey in their customer panel. In total 5000 people in the NS panel were approached Monday 27th of May to take part in the survey, 666 of them agreed to take part and completed the survey. Furthermore, a link to the survey was distributed among friends and connections through social media (Facebook, Twitter, LinkedIn, Yammer (company network)). This resulted in an additional 114 completed responses. Finally, Tuesday June 4th flyers (1000 pieces) were handed out at Schiphol Plaza. Although handing out the flyers was quite effortless, the return rate was very low; it resulted in only 23 completed responses. While that last fact is a pity, the number of collected complete responses is far larger than the minimum required.

6.2. Characteristics of sample

To get knowledge about the composition of the sample, several background variables were collected. The results can be seen in Table 6.1. The population was defined as Dutch people who have recently travelled outside Benelux-countries. However, it is almost impossible to check if the collected sample is a representative proxy of this population, because there are no numbers (publicly) available about the background variables. To give some idea about how the sample composition compares, Table 6.1 compares the characteristics from the sample to CBS and MPN statistics. However, it should be noted that those statistics are based on a different population, therefore an extended comparison and explanation for differences is not meaningful.

As could be seen in Table 6.1 the sample contains people who have (ample) experience with travelling by train. This is expected due to a large number of responses from the NS panel and the fact that my network mostly contains students who can travel by train for free. The latter group also shows up in the lowest disposable income category. Furthermore, a large share of the sample has completed higher education (HBO or higher). Remarkably, a lot of older people took the time to complete the survey, almost no people under 20 responded (or were approached).

Concluding, the sample representativity of a population consisting of Dutch people travelling abroad is questionable. This is because most respondents are regular train users. However, this is not considered to be a big problem in this study. It is not unreasonable to assume that people who are willing to use the night train instead of flying are experienced train travellers. It could be considered that a primary market comprises mainly regular train users. When reflecting on the results it should be considered that the sample consisted of mostly regular train travellers.

Table 6.1: Table showing sample characteristics

Background variable	Category	Absolute number	Relative	Comparison CBS
Gender	Male	448	56.1%	49.6%
	Female	351	43.9%	50.4%
	Total	799[1]		
Education	Preschool	0	0%	10.3%
	VMBO, HAVO, VWO, MBO1	23	2.9%	21.3%
	HAVO, VWO, MBO2-4	128	16.4%	38.7%
	HBO, WO Bachelor	311	39.8%	18.9%
	HBO, WO Master, or higher	320	40.9%	10.8%
	Total	786[1]		
Age	0 - 19	9	1.1%	22.2%
	20 - 39	198	24.6%	24.8%
	40 - 64	347	43.2%	34.2%
	65 - 79	244	30.3%	14.3%
	80+	6	0.7%	4.5%
	Total	804		
Disposable income	<10.000	80	11.9%	4.4%
	10.000 - 20.000	65	9.6%	16.3%
	20.000 - 30.000	122	18.1%	22.0%
	30.000 - 40.000	167	24.7%	16.5%
	40.000 - 50.000	86	12.7%	13.1%
	50.000 - 100.000	133	19.7%	24.7%
	100.000 - 200.000	18	2.7%	2.6%
	>200.000	4	0.6%	0.4%
	Total	675[1]		
Employment status[2]	Student/Scholar	94	11.9%	
	Parttime worker	149	18.9%	
	Fulltime worker	303	38.4%	
	Retired	226	28.6%	
	Jobless	17	2.2%	
	Total	789[1]		
Train travel[3]	Never	2	0.2%	26.1%
	<1 day a year	8	1.0%	16.9%
	1 - 5 days a year	38	4.7%	22.1%
	6 - 11 days a year	142	17.7%	13.0%
	1 - 3 days a month	221	28.7%	10.2%
	1 - 3 days a week	205	25.5%	5.5%
	4 or more days a week	178	22.1%	6.1%
	Total	804		

[1]: the total number is different from 804 because some respondents opted not to provide this information.
[2]: no meaningful comparison possible.
[3]: compared to MPN 2014 data.

6.3. Travel characteristics

The survey also included several questions related to recent travel experiences. Respondents were asked how often they travelled by international train (defined as Intercity Brussels, Intercity Berlin, Thalys, ICE and Eurostar) and by plane in the past 2 years. Results can be seen in Figure 6.1. Almost 40% of the sample does not have any experience of travelling by international train. The largest share of the sample (30.8%) travelled by aeroplane 2-3 times during the past 2 years.

Given the large sample of train travellers, it is tested if frequent train travellers travel more by international train. Frequent train travellers are defined as people who use the train one or more days a week. Results show that frequent train travellers do have more experience with travelling by international train (mean value 1.54 vs 1.09. However, these numbers are ordinal measurement scale, therefore not reflect an absolute number of trips). An independent samples t-test revealed that there is a significant difference between both groups (the mean international train travel experience is 0.45

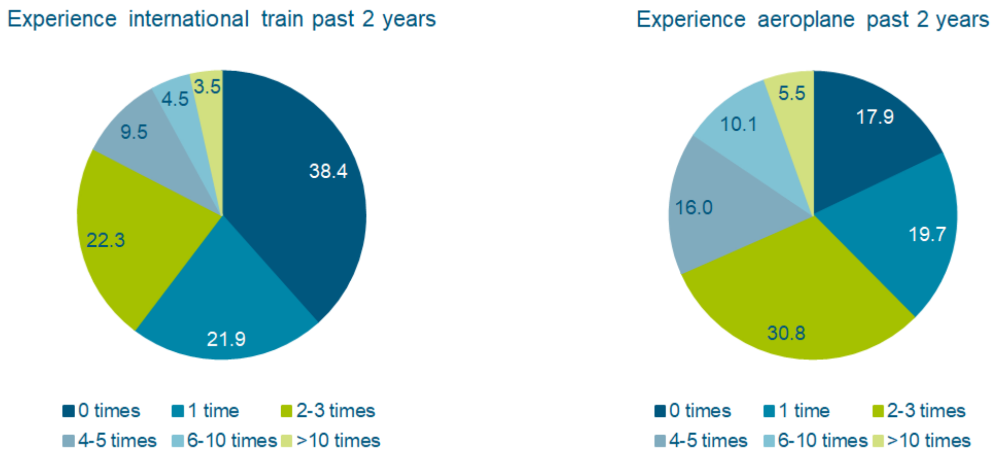


Figure 6.1: Travel experience of respondents by international train and aeroplane [%].

'categories' higher). However, another t-test revealed that the frequent train travellers also have more experience with aeroplane travel (2.18 vs 1.78 on an ordinal scale). Therefore, this group, compared to non-frequent train travellers, travels more in general. Furthermore, even the frequent train travel group travels more by aeroplane than by international train. Therefore, the effect of this bias on the results could be limited. It will be explored if being a frequent train traveller influences the perceived comfort and utility levels of the different alternatives for mode choice.

Furthermore, before giving any information or indication as to how long a journey would take from arrival at the airport or train station until arrival at the city centre, respondents were asked to estimate the total time in hours. The resulting histograms could be seen in Figure 6.2. The histogram for the night train shows some drops/gaps. This is likely the results of being unfamiliar with the travel alternative, therefore people used mostly rounded numbers for estimation. On average, the estimations were quite accurate. For the night train, the average travel time is estimated as 12 hours and 6 minutes, which is inside the range of included attribute levels. For the plane, travel time is estimated as 5 hours and 15 minutes, again right in the middle of the included attribute levels.

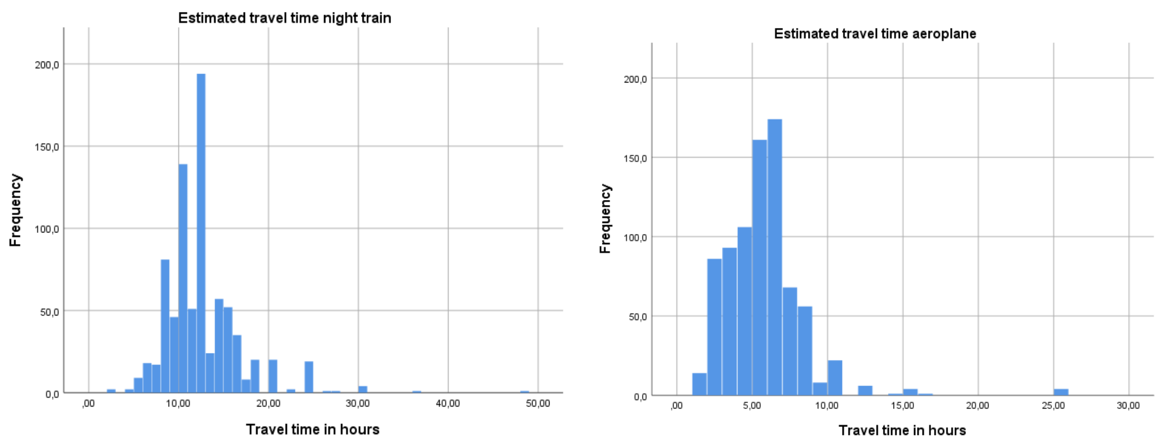


Figure 6.2: Histograms of travel time estimations.

Finally, the mode choices the respondents made are analysed. The percentages chosen for each mode for each separate choice task are shown in Figure 6.3. On average, the night train is chosen around 55%. However, there are some exceptions: in choice task 11 the share of the night train is lowest for both contexts (23.6% and 17.6%). Looking back at the composition of the choice set this is a logical result, the train cost is highest, comfort low and the plane morning costs are lowest. This means the night train is almost dominated by the other alternatives. On the other side, choice task 29 resulted in a very high share for the night train for both contexts (87.2% and 86.4%). Again, this is not surprising as this choice task had a low travel time, 5-star comfort, lowest cost for the night train, and the average cost for the plane morning alternative. In this choice task, the night train dominated the other alternatives.

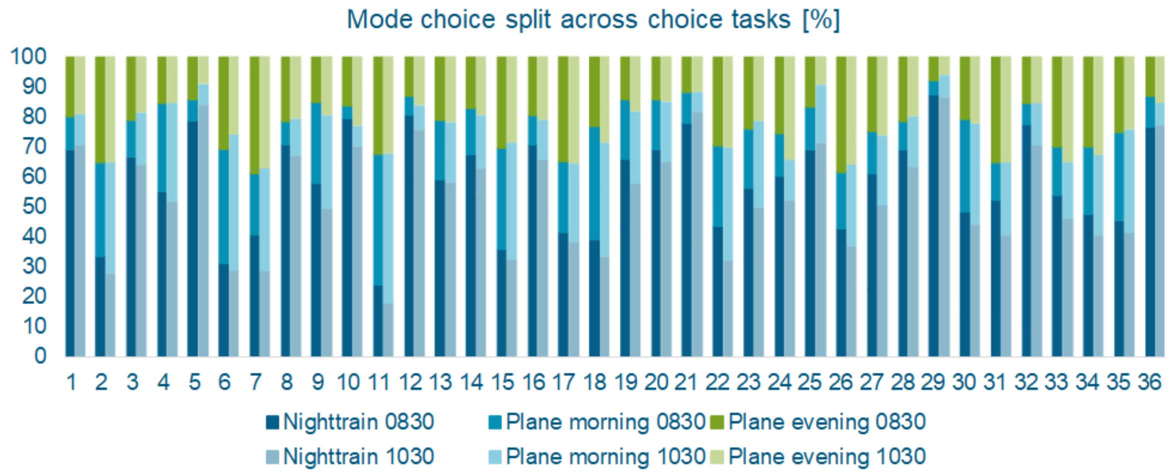


Figure 6.3: Mode choice split across choice tasks for both preferred arrival time contexts

6.4. Summary

Main takeaways

- In total 804 responses were collected. Respondents were approached using the NS panel, at Schiphol Plaza and through the authors own network.
- The majority of the sample is highly educated and uses the train at least once a week.
- The estimated travel time for both aeroplane and night train are realistic.
- On average, across all choice sets, the night train is chosen in 55% of the choice tasks.

This chapter provided an overview of the collected data. In total 804 completed responses have been collected from the customer panel of NS, at Schiphol Plaza and friends and family. Characteristics show that the majority of the respondents is highly educated and also a large percentage of people is retired. Furthermore, almost 50% of the sample regularly uses the train in daily life. It is impossible to check if the sample is a good proxy of the intended population (Dutch people who travelled outside the Benelux-countries in 2018) due to a lack of information on this population. However, as it likely that mostly regular train users are key market for the night train, the composition of the sample is not considered to be a big problem. Nevertheless, it will be further explored if the fact of being a frequent train traveller influences the mode choice and 'perceived comfort'. Estimations for the travel time by aeroplane and night train proved to be in the middle of the included attribute levels. It has been shown that respondents during the Stated-Preference experiment, on average, opted for the night train travel alternative in 55% of the time.

7
Results

This chapter presents the model outcomes for the different experiments. The model developed for the comfort rating experiment is discussed in Section 7.1. Section 7.2 points out the model development process of the mode choice model, including the estimation outcomes of MNL, NL and ML models and compares which of those is the 'best' model for the collected data. Section 7.3 discusses the combination of both experiments. The latent class model to derive insights into market segments is presented in Section 7.4. Finally, a summary is provided in Section 7.5.

7.1. Comfort rating experiment: regression

7.1.1. Model estimation

The 'perceived comfort' variable in this study is both included as a dependent and as an independent variable. As introduced, in the rating experiment the 'perceived comfort' is the dependent variable. The experiment included several attributes, for recollection: accommodation, number of people, the possibility to lock the compartment, catering facilities, the possibility to shower and number of stops during the night. Furthermore, several socio-demographics of the sample were collected.

The research question to be answered is what comfort determinants are for a night train trip. This will be answered by setting up a regression model. For the selection of a suitable regression model, it is important to consider the measurement level of the target variable. In general, the target variable, a number rating, could be considered to be a continuous scale. To limit the number of response options in the experiment, a simplification was made. People were asked to rate their 'perceived comfort' level of a case on a 1 to 5-star scale, which can be considered to be an ordinal measurement scale. However, since the target variable still represents a continuous variable, a simplification is made and the data is, at first, analysed using a standard linear regression model. The regression models are estimated using IBM SPSS Statistics 25.0.

Before the regression model can be estimated, some of the independent variables (attributes) have to be coded. This is because they are categorical variables, which is the case for accommodation and catering facilities. It would be incorrect to incorporate these variables as linear effects, the utility of a 'sleeper' accommodation is not necessarily twice the utility of a 'couchette' accommodation. Therefore, effects coding was applied. This resulted in two new parameters for each categorical variable. The coding of the main attributes and other variables can be seen in Table 7.1.

Additionally, the collected socio-demographics can be included in the model. Some of these were measured on a very detailed scale, e.g. as income or train travel frequency. Most of these are ordinal scale variables. Coding those variables using effects coding would result in a major increase in the number of parameters. Therefore, it is decided to recode the variables in 'simple' bins, for example 'high education' which could be yes or no. The new coded variables make use of effects coding (to not mix dummy and effects coding) and can be seen in Table 7.1. Please note that people that opted to withheld their answer, are included in the 'Other/Else' categories.

Table 7.1: Coding of the included attributes in the comfort rating experiment

Attribute	Level	Coding	Effects coded	
Main attributes			Parameters	
Accommodation			Sleeper	Couchette
	Sleeper	0	1	0
	Couchette	1	0	1
	Seat	2	-1	-1
Catering facilities			Restaurant	Kiosk
	Restaurant car	0	1	0
	Kiosk	1	0	1
	None	2	-1	-1
Shower			Shower	
	Yes	1	1	
	No	0	-1	
Lock			Lock	
	Yes	1	1	
	No	0	-1	
Number of Stops			Stops	
	0, 3, 6		Real values	
Number of People			People	
	2, 4, 6		Real values	
Socio-Demographics			Parameters	
Gender			Gen	
	Female	1	1	
	Male & Other	0	-1	
Education			HighEdu	
	HBO Bachelor or higher	1	1	
	Else	0	-1	
Income			HighIncome	
	High disposable income (>€40.000)	1	1	
	Else	0	-1	
Workstatus			Student	Retired
	Student	0	1	0
	Retired	1	0	1
	Else	2	-1	-1
Train Travel			FreqTrain	
	Once or multiple days a week	1	1	
	Else	0	-1	
Age			Age	
	Complete range		Real values	

7.1.2. Expectations

Several expectations regarding the parameters to be estimated can be formulated. Main attributes are the primary determinants of the comfort rating. It is expected that most main parameters have a positive sign. This is because most are coded in a 'positive' way, e.g. if there is the possibility to shower it is coded 1, otherwise -1. Exceptions are the attributes 'Number of people' and 'Number of stops', which are expected to have a negative sign, as (logically) more people or stops along the way are expected to reduce the rating for 'perceived comfort'.

Including socio-demographics and other background variables may help to determine if certain groups, e.g. people with high income, rate the comfort of the night train higher or lower. Previous studies have shown that travelling at night causes more anxiety for women than for men (Lynch & Atkins, 1988; Stradling, Carreno, Rye, & Noble, 2007). This effect might be found in the mode choice experiment. However, it is similarly interesting if this effect is also present in the rating experiment. If present, the parameter for women could be negative, meaning women have a lower base rating of the 'perceived comfort' of the night train. Furthermore, regular train users (more than one day a week) are expected to give a higher rating, because they are familiar with train travelling and it is their regular mode choice. Regarding income and education levels it is expected that the parameters are negative. Meaning that people who have a high disposable income and/or high education level give a lower rating, this is because they might have experience with more luxurious ways of travelling.

Finally, regarding expectations for interaction effects, a study into mode choice from Switzerland showed that elderly people prefer more comfort over lower fares (Ohnmacht & Scherer, 2010). In the rating experiment, this effect could be present as a general background variable (age parameter) and/or as an interaction effect with the accommodation type. Furthermore, as stated before, it might be that people with more income or higher education are more used to private or luxurious transport and therefore have a stronger dislike for the 'Number of people' attribute, this could show up in an interaction effect. A same kind of reasoning can be applied to the combination of high income and the accommodation type, therefore expecting a positive parameter. At last, it was mentioned in the focus groups that bathroom facilities might be more important for women. Therefore, this is tested as an interaction effect.

7.1.3. Linear regression results

In this subsection, the outcomes of the linear regression are reported. Several models are estimated, the results are visible in Table 7.2. Parameter estimates, as well as t-value and significance, is reported. First, a model with only the main attributes was estimated. Additionally, to see if the expectations hold, a model with all the socio-demographics and interactions was estimated. In the resulting model, not all parameters were proven to be significant at 95% level. Therefore, insignificant parameters were step-wise removed and the model was re-estimated, also called backwards elimination. This process was repeated until all remaining parameters were significant.

Table 7.2: Parameter estimates, t and significance values from linear regression.

Model:	Main effects			Background & Interactions			Final model		
Parameter:	Value	t	Sig	Value	t	Sig	Value	t	Sig
Main attributes									
(Constant)	3.999	90.949	<0.001	4.036	43.054	<0.001	4.086	61.321	<0.001
Sleeper	0.380	17.716	<0.001	0.710	10.330	<0.001	0.722	11.083	<0.001
Couchette	0.133	6.147	<0.001	0.309	4.394	<0.001	0.285	4.276	<0.001
People	-0.261	-28.376	<0.001	-0.244	-20.285	<0.001	-0.253	-26.966	<0.001
Lock	0.160	10.629	<0.001	0.163	10.938	<0.001	0.163	10.962	<0.001
Restaurant	0.149	6.769	<0.001	0.147	6.775	<0.001	0.148	6.810	<0.001
Kiosk	0.115	5.376	<0.001	0.116	5.481	<0.001	0.116	5.532	<0.001
Shower	0.128	8.287	<0.001	0.128	8.383	<0.001	0.127	8.374	<0.001
Stops	-0.070	-11.429	<0.001	-0.066	-8.988	<0.001	-0.070	-11.560	<0.001
Socio-Demographics									
Gen				0.009	0.556	0.578			
HighIncome				-0.125	-2.867	0.004	-0.110	-6.553	<0.001
HighEdu				0.084	1.589	0.112			
Student				0.043	1.418	0.156			
Retired				0.007	0.314	0.753			
FreqTrain				-0.033	-1.998	0.046	-0.034	-2.104	0.035
Age				-0.002	-1.381	0.167	-0.006	-2.845	0.004
Interactions									
SleeperAge				-0.006	-5.315	<0.001	-0.003	-5.505	<0.001
CouchetteAge				-0.003	-2.607	0.009	-0.003	-2.456	0.014
IncomePeople				0.005	0.533	0.594			
EduPeople				-0.024	-2.196	0.028	-0.012	-2.934	0.003
GenderShower				0.013	0.877	0.381			
EduStops				-0.008	-1.131	0.258			
IncomeSleeper				-0.019	-0.835	0.404			
IncomeCouchette				0.029	1.233	0.217			
R^2	0.263			0.287			0.286		
$Adj.R^2$	0.262			0.283			0.283		

Comparing the R^2 values for the different models, it is shown that the final model performs better than the initial model and explains 28.6% of the variance in the dependent variable.

Standard linear regression assumed that the dependent variable is of an interval measurement scale. While it was argued that this might be a reasonable assumption, it is explored if applying ordinal logistic regression results in completely different parameter estimates. The estimated parameters have the same signs, although one additional parameter 'FreqTrain' is not significant at 95% and therefore excluded. The interested reader can find the full results in Appendix D.

Comparing both models, it is noted that the ordinal model has one parameter fewer, the interpretation of the linear model is more straightforward and the underlying assumption for linear regression is not unreasonable. Additionally, the research questions can be answered based on the linear regression model. For these reasons, it is decided to consider the results from linear regression in the remainder of this study.

7.1.4. Parameter interpretation

Considering the results of the linear regression one can make several observations. First, all the parameters have the expected sign. All main parameters are highly significant. The parameter with the largest positive effect is 'Sleeper', which represents the sleeper type accommodation. Another observation is that people would like to get food or drinks on the train, but only make a very small distinction between a 'kiosk' or 'restaurant car'. The difference in rating points is only 0.032. This means that upgrading catering facilities to a restaurant car would yield a far smaller marginal increase in comfort rating than upgrading from no catering to a kiosk. Strongest negative effect is the number of people in the compartment. People have a strong dislike for sharing the accommodation, even more so if they are highly educated. A 2 people compartment for someone highly educated leads to a disutility of $2 \cdot (-0.258 - 0.012) = -0.54$ and 6 person compartment to a disutility of $6 \cdot -0.27 = -1.62$. The difference (1.08 rating points) means that even a non-stop night train with showers, a restaurant car, lockable compartments, but with shared 6 person accommodation, gets a lower perceived comfort rating than a basic night train that stops 6 times, has no catering facilities, no showers, no locks but does have more private compartments for only 2 people.

Furthermore, regarding the previously formulated expectations, it can be noted that not all of them hold. It was found that 'gender' did not have a significant influence on the overall comfort rating and valuation of the possibility to shower. The expectation that people with a high disposable income rate the night train lower does hold. In contrast to expectation, being a regular train user negatively influences the comfort rating. However, pointing towards a reason for this negative effect would be speculating as this data was not collected. Finally, the effect of age on the comfort rating is highlighted. The same effect as in the Swiss study could not be found. It appears that with increasing age respondents give a lower rating to the overall night train comfort level ('Age' parameter), but also put less weight on the accommodation type as both 'SleeperAge' and 'CouchetteAge' interactions are negative. This means that the older the person, the less extreme they feel about the type of accommodation. To illustrate the extent of this last effect, the values for the 'Sleeper' and 'Couchette' parameters with increasing age are presented in Table 7.3. Note that the differences between 'Sleeper' and 'Couchette' do not diminish with age.

Table 7.3: Table showing the parameter values and differences between accommodation types for varying ages.

Parameter	Accommodation	Parameter value				Utility differences between levels			
Age [years]		25	45	65	80	25	45	65	80
Sleeper	Sleeper	0.647	0.587	0.527	0.482	0.437	0.437	0.437	0.437
Couchette	Couchette	0.21	0.15	0.09	0.045	1.067	0.887	0.707	0.572
	Seat	-0.857	-0.737	-0.617	-0.527				

Examples of possible night train observations for different 'perceived comfort' levels are shown in Table 7.4. The predicted score is based on a 40-year old, highly educated, non-frequent train traveller, who has a high disposable income. Note that the predicted score does not reach a score of 5.0. This is due to the above mentioned socio-demographic and interaction effects. Furthermore, these configurations are only some examples, many more combinations are possible, which would result in a similar score. For this same individual, the current night train service from Dusseldorf towards Vienna would score 4.0 stars in 'Sleeper - 2 beds' class, 2.9 stars in 'Couchette - 4 beds' class and 1.3 stars in 'Seat' class.

Table 7.4: Examples of night train configurations for different 'perceived comfort' levels.

Train comfort/ Attribute	Accommodation	Number of People	Lock compartment	Catering	Shower	Stops	Predicted score
1 star	Seat	6	No	Kiosk	No	6	1.1
2 stars	Seat	4	No	Kiosk	Yes	3	2.1
3 stars	Couchette	4	Yes	Kiosk	Yes	6	3.0
4 stars	Couchette	2	Yes	Kiosk	Yes	0	3.9
5 stars	Sleeper	2	Yes	Restaurant car	Yes	0	4.3

7.2. Mode choice experiment: discrete choice modelling

7.2.1. Model estimation

The collected data is used to estimate a discrete choice model. Several different models are estimated, with a varying number of parameters. Answering the research questions involves researching which main parameters are significant, but also extending the model to incorporate context, socio-demographic and interaction parameters and explore the effect of those variables on mode choice.

Similarly to data preparation for the regression models, ordinal and categorical variables are effects coded. This is done for several socio-demographics and background variables. The 'perceived comfort' was assumed to be linear in the rating experiment. For this reason, it is chosen to include the 'perceived comfort' attribute as a linear variable with a quadratic component. The used coding of parameters is visible in Table 7.5. The collected data is first prepared in SPSS Statistics 25.0, the choice models are estimated using PandasBiogeme (Bierlaire, 2018).

Table 7.5: Coding of the included attributes in the mode choice experiment

Attribute	Level	Coding		
Main attributes			Parameters	
Travel time night train			TTTrain	
	11:45	1.175	Real values	
	13:00	1.3		
	14:25	1.425		
Travel cost night train			TCTrain	
	40	0.4	Real values	
	80	0.8		
	120	1.2		
	160	1.6		
Travel time plane			TTPlane	
	5	1	Real values	
	5:30	1.1		
Travel cost plane			TCPlane	
	60	0.6	Real values	
	110	1.1		
	160	1.6		
Perceived comfort			Comfort	Comfort_Q
	1 - 5 stars	[1-5]	Linear & quadratic coded	
Socio-Demographics			Parameters	
Gender			Gender	
	Female	1	1	
	Male & Other	0	-1	
Education			HighEdu	
	HBO Bachelor or higher	1	1	
	Else	0	-1	
Income			HighIncome	
	High disposable income (>€40.000)	1	1	
	Else	0	-1	
Workstatus			Student	Retired
	Student	0	1	0
	Retired	1	0	1
	Else	2	-1	-1
Train Travel			FreqTrain	
	Multiple days a week	1	1	
	Else	0	-1	
Travel purpose			Purpose	
	Business	1	1	
	Leisure	0	-1	
Arrival Time			ArrivalTime	
	Arrival at 08:30	0	-1	
	Arrival at 10:30	1	1	

7.2.2. Expectations

It is expected that the parameters regarding travel cost and travel time have a negative sign. This is in line with intuition and also found by previous mode choice studies. The total utility effect of both linear and quadratic comfort components combined is expected to be positive. Because a higher comfort level adds to the utility level, this should lead to an increase in the probability of choosing the more comfortable mode.

Lynch and Atkins (1988); Stradling et al. (2007) both found that women are more reluctant to travel by public transport at night. Therefore it will be tested if this effect can also be found in this study for long-distance travel. Furthermore, it is expected that people with a higher education level choose the train more often, as found in some cases by the study of Limtanakool et al. (2006a). It is interesting to see what the influence is of being a regular train user in daily life, it is expected to be positive. Furthermore, it is expected that people with a higher disposable income will choose faster modes, and with age, people are more inclined to use the train (Limtanakool et al., 2006a). Additionally, the travel purpose, business or leisure, will be included in the mode choice, as it is an important determinant (Buehler, 2011; Román et al., 2010). The business purpose is expected to have a disutility for morning plane or night train.

Furthermore, it is expected that the preferred arrival time in the morning does influence the mode choice. It is expected that people have a higher preference for the night train when the arrival is early in the morning. Additionally, the study by Román et al. (2010) found that with a higher comfort level the travel time parameter has a higher (less negative) value. Therefore, this effect will also be tested. At last, the effect of the purpose on the travel cost parameter will be explored. It is expected that people who travel for business, pay less attention to the travel cost.

7.2.3. Model estimation results

Several discrete choice models with a varying number of effects are estimated, these are main effects, socio-demographics, background variables and other interactions. A MNL model with only main effects acts as a base model. Furthermore, a Nested Logit model is as an intermediate step estimated to see if there are nests present. This was the case for the modes that both arrive in the morning. At last, also a Panel Mixed Logit model with 200 draws is estimated, to accommodate for the panel structure in the dataset. The nest in this model is taken into account as an error component. The resulting utility specifications used for this model are shown in Equations 7.1 to 7.3. The plane in the evening mode is the base alternative and therefore the 'ASC_Planeref' is fixed to 0. The results of the parameter estimations, as well as t and significance values, are shown in Table 7.6. The interested reader can find the PandasBiogeme code and more estimation results in Appendix E

$$\begin{aligned} V_{EP} = {} & ASC_{NT} + \beta_{Com} \cdot Com + \beta_{Com\_Q} \cdot Com^2 + \beta_{TTTrain} \cdot TTTrain + (\beta_{TCTrain} + \beta_{CP}) \cdot TCTrain \\ & + \beta_{AT\_Train} \cdot AT + \beta_{HE\_Train} \cdot HE + \beta_{HI} \cdot HI + \beta_{Pp\_train} \cdot Pp + \beta_S \cdot S + \beta_{Age} \cdot Age + \sigma_{Morning} \end{aligned} \tag{7.1}$$

$$\begin{aligned} V_{MP} = {} & ASC_{MP} + \beta_{TTPlane} \cdot TTPlane + \beta_{TCPlane} \cdot TCPlane + \beta_{AT\_MP} \cdot AT \\ & + \beta_{Gender} \cdot Gender + \beta_{HE\_Plane} \cdot HE + \beta_{Pp\_plane} \cdot Pp + \beta_{FT\_Plane} \cdot FT + \sigma_{Morning} \end{aligned} \tag{7.2}$$

$$V_{EP} = ASC_{EP} = 0 \tag{7.3}$$

In these formulas, Com = comfort rating, CP = interaction with travel purpose and travel cost for the night train, AT = arrival time context, HE = highly educated, HI = high income, Pp = travel purpose, S = student, FT = frequent train traveller.

Table 7.6: Parameter estimations for main attributes, socio-demographics and interactions for the various models.

Model:	MNL Base			NL Final			Panel ML Final		
Parameter:	Value	t	Sig	Value	t	Sig	Value	t	Sig
ASC_Train	1.68	5.66	<0.001	0.919	3.19	0.001	2.06	5	<0.001
ASC_Plane	1.85	3.39	<0.001	1.65	3.27	0.001	3.1	4.8	<0.001
Beta_Com	0.518	7.49	<0.001	0.491	7.15	<0.001	0.648	8.09	<0.001
Beta_Com_Q	-0.0294	-2.58	0.01	-0.0283	-2.65	0.008	-0.0322	-2.43	0.015
Beta_TCPlane	-0.972	-15	<0.001	-0.9	-11.7	<0.001	-1.12	-15.9	<0.001
Beta_TCTrain	-0.774	-15.8	<0.001	-0.667	-11.5	<0.001	-0.939	-13.7	<0.001
Beta_TTPlane	-1.72	-1.67	0.094	0.829	-1.81	0.071	-1.27	-2.26	0.024
Beta_TTTrain	-0.93	-4.8	<0.001	-0.888	-4.38	<0.001	-1.21	-4.84	<0.001
Beta_ArrivalTime_Train				-0.0547	-2.04	0.041	-0.0776	-2.1	0.036
Beta_ArrivalTime_Plane				0.0782	2.52	0.012	0.0869	2.11	0.035
Mu_morning				1.18	15.1	<0.001			
Sigma_morning_std							3.32	21.5	<0.001
Beta_Gender_Plane				0.219	6.76	<0.001	0.121	4.08	<0.001
Beta_Gender_Train				0.133	4.68	<0.001			
Beta_HighEdu_Plane				-0.442	-10.8	<0.001	-0.874	-5.07	<0.001
Beta_HighEdu_Train				-0.205	-5.45	<0.001	-0.578	-3.37	<0.001
Beta_HighIncome_Plane				-0.134	-3.81	<0.001			
Beta_HighIncome_Train				-0.22	-7.36	<0.001	-0.117	-3.36	<0.001
Beta_Purpose_Plane				-0.447	-12.3	<0.001	-1.11	-6.92	<0.001
Beta_Purpose_Train				-0.672	-10.8	<0.001	-1.35	-7.81	<0.001
Beta_Student_Plane				0.331	6.35	<0.001			
Beta_Student_Train				0.159	2.98	0.003	-0.209	-3.98	<0.001
Beta_FreqTrain_Plane				-0.57	-4.51	<0.001	-0.729	-4.91	<0.001
Beta_Age_Train				0.0134	8.52	<0.001	0.017	8.15	<0.001
Beta_CostPurpose				0.167	3.08	0.002	0.184	2.7	0.007

Reflecting on the results it can be noted that all the main parameters have the expected sign and are significant at 95% level. The exception is 'Beta_TTPlane' in the MNL and NL models, which is not significant. When accommodating for the panel structure of the dataset, this parameter also turns out to be significant at 95% and has the expected sign. Compared to the NL Final model, taking into account the panel effect does turn several variables insignificant, those were removed using 'backward elimination'.

The different arrival time contexts were solely taken into account as an 'interaction on a constant', meaning that the arrival time influences the base preference for either the night train or the morning plane compared to the evening plane base alternative. At last, a nest (see parameter 'Mu_morning') was found to be significant, meaning that both modes that arrive on the same day share unobserved attributes. This was modelled with an additional error term in the Panel ML model. A more thorough interpretation of the parameter estimation will be provided in Section 7.2.6.

7.2.4. Comparison

A comparison between the different estimated models is shown in Table 7.7. One can already notice that applying a Panel Mixed Logit model results in a large jump in the ρ^2, and likewise large increase in final loglikelihood. Nevertheless, the likelihood ratio test (LRS) is used to verify if a model with more parameters fits the data better.

Table 7.7: Model fit information for the estimated models

Model:	#Parameters	Rho-square	Rho-square-bar	Initial LL	Final LL
MNL Base	8	0.155	0.154	-10599.41	-8955.28
NL Final	24	0.205	0.203	-10599.41	-8421.31
Panel ML Final 200 draws	21	0.357	0.355	-10599.41	-6815.64

Application of the likelihood test only makes sense when one can obtain a more complex model B by restraining model A. This is, for example, the case with the NL Final versus the MNL base model. These calculations are presented in Equations 7.4.

$$LRS = -2 \cdot (LL_{MNLFinal} - LL_{NLFinal}) = -2 \cdot (-8955.28 + 8421.31) = 1067.94 \tag{7.4}$$

One can see that the LRS outcome for the added nest and socio-demographics is a large number of 1067.94. The critical χ^2 value for a parameter difference of 16 and significance level at 95% is 26.30 and at 99% significance 32.00. This means that the chance that the 'NL Final' model fits the data better (higher LL) due to coincidence is less than 1%. Therefore it can be concluded that the 'NL Final' model definitely is the better model.

Comparing the NL Final and the Panel ML Final model involves applying a different test. Because the ML model is not nested under the NL Final model, the Ben Akiva & Swait test should be applied (M. Ben-Akiva & Swait, 1986). The formula can be seen in Equation 7.5. In this formula, p is the chance that while model A fits the data better, model B is the better model for the population. N is the number of observations, in our case 9648, j is the number of alternatives in the choice set (our case: 3) and the needed LL scores are shown in Table 7.7. Substituting these values in the test is shown in Equation 7.6.

$$p = NormSDist(-\sqrt{2 \cdot N \cdot \ln(j) \cdot ((LL(B) - LL(A))/LL(0))}) \tag{7.5}$$

$$\begin{aligned} p = NormSDist(-\sqrt{2 \cdot 9648 \cdot \ln(3) \cdot ((-8421.31 + 6815.64)/-10599.41)}) \\ = NormSDist(-56.69) = 0 \end{aligned} \tag{7.6}$$

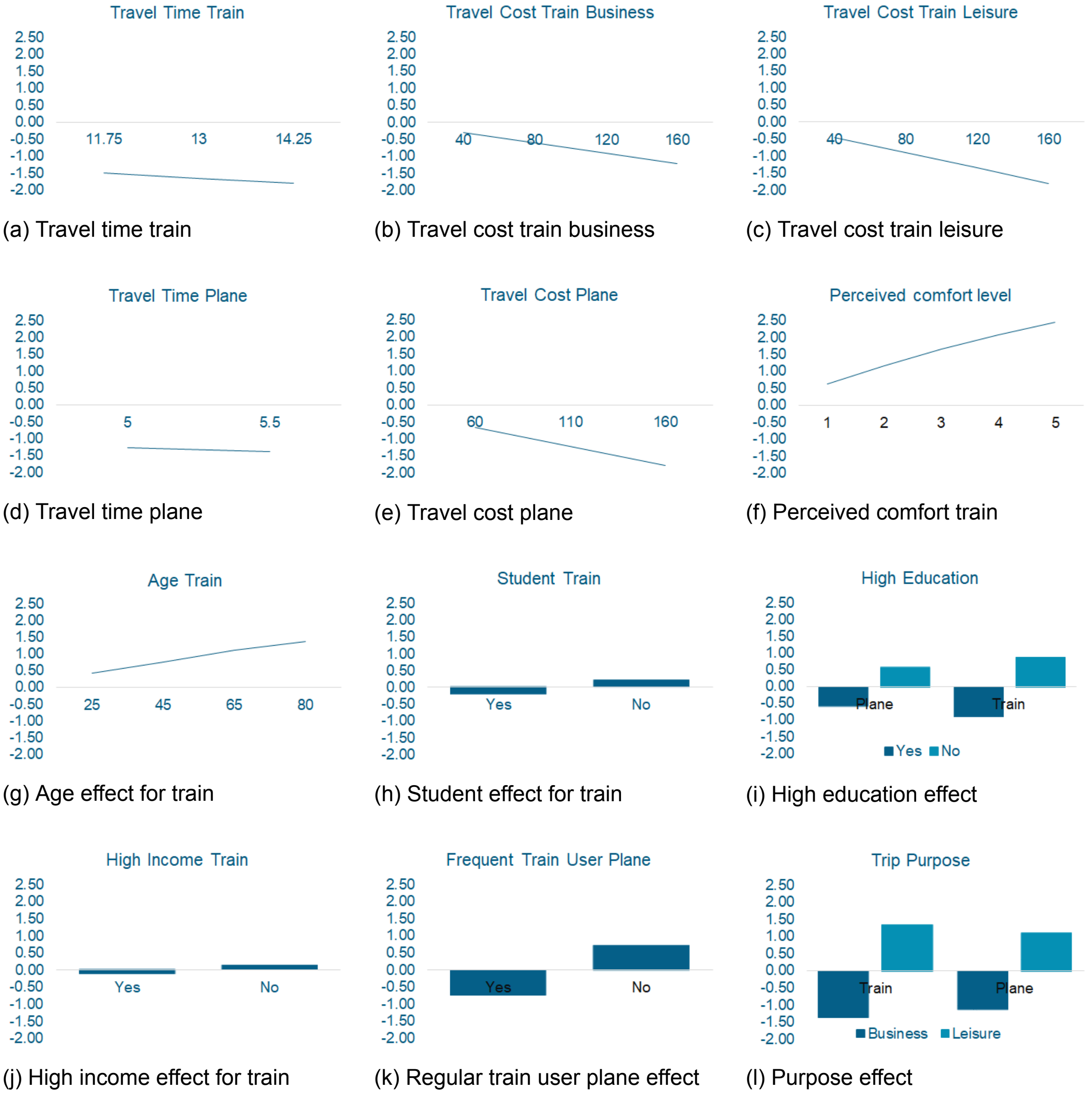
The outcome of the Ben-Akiva & Swait test is 0, meaning that, as expected, the final Panel Mixed Logit model is indeed the best fitting model. This model explains 35.7% of the initial variance. It is decided to use this model for the remainder of the study.

7.2.5. Utility contribution

Since the final model and parameter estimations have been determined, the overall contribution to the utility of the different modes can be calculated. The utility range and contribution of main attributes is presented in Table 7.8, additionally it is visualised and shown also for some other variables in Figure 7.1.

Table 7.8: Utility range, minimum and maximum contributions of utility for main parameters

Parameter	Range	Min. Utility Contribution	Max. Utility Contribution
TC Train	40 - 160	-1.797	-0.302
TC Plane	60 - 160	-1.792	-0.449
TT Train	11:45 - 14:15	-1.810	-1.492
TT Plane	05:00 - 05:30	-1.397	-1.270
Comfort	1 - 5 stars	0.616	2.435



(a) Travel time train
(b) Travel cost train business
(c) Travel cost train leisure
(d) Travel time plane
(e) Travel cost plane
(f) Perceived comfort train
(g) Age effect for train
(h) Student effect for train
(i) High education effect
(j) High income effect for train
(k) Regular train user plane effect
(l) Purpose effect

Figure 7.1: Utility contributions of the different attributes

Observing the figures one can see that the largest negative effect for the night train is caused by the travel cost for the leisure purpose, and largest positive effect of the main attributes is the comfort 5-star level. Of the socio-demographics, age has the highest positive effect on the train utility. The largest negative effect of the socio-demographics and background variables is caused by the business travel purpose for the night train, as well as for the morning plane.

7.2.6. Parameter interpretation

Interpreting the estimations of the parameters allows providing useful insights about choice behaviour. The ASCs for the night train, as well as the morning plane alternative, are positive and quite substantial. However, interpretation is a little more complicated. The ASC is the utility when all the attributes would be zero, including cost and travel time. That would obviously not result in a 'valid' travel alternative. Furthermore, the reference alternative is not a 'non-travel alternative', it has attribute values, especially the cost in the reference alternative is high compared to the other alternatives. Therefore, it is logical that the other two alternatives have positive ASCs.

The effect of comfort is indeed quadratic. The utility difference between a 1 star and 3-star comfort level is 1.039 utility points. In other terms, if this is divided by the travel time parameter, downgrading from a 3-star comfort level to a 1-star level is equivalent to an additional 8.2 hours of travel time. Similarly, decreasing from 5 to 3-star comfort has a utility difference of 0.781 points and therefore 6.1 hours of travel time. Furthermore, one can observe that the utility contributions of the travel time for night train and morning plane are quite similar, yet the magnitude of the attribute levels differs a lot. This implies that people do not put equal weights on these travel minutes, i.e. one minute onboard the night train weighs less than one minute onboard an aeroplane in the morning.

Regarding socio-demographics and background variables, one can observe that the trip purpose has a large effect on utility. Taking the night train for business purposes results in a penalty equal to about €358 in monetary terms (compared to leisure purpose). For the morning plane alternative, this penalty is about €294. This means business travellers have a strong preference for taking the plane the day before and staying in a hotel at their destination. Furthermore, the effect of frequent train users is highlighted. If one is a frequent train traveller, one associates a disutility of about 1.458 points with the morning plane alternative. This is compared to non-frequent train travellers and the night train and evening plane modes. This number is about equal to 11.5 hours of travel time by night train, meaning that frequent train travellers travel either by night train or the day before by plane in the evening. Non-frequent train travellers are more willing to use the aeroplane in the early morning.

A comment is made about the 'Beta_ArrivalTime' parameters. This is the parameter that represents the different arrival time contexts. This context effect was only included as an interaction effect on the constant. The resulting values are quite small. They represent that for arriving at the 08:30, the night train utility is improved by about 1.4 hours worth of travel time, the plane in the morning receives a penalty of about 0.6 hours in terms of travel time.

Finally, reflecting on the previously formulated expectations, it has to be concluded that some effects could not be found significantly in the collected data. This is summarised in Table 7.9. Expectations that do not hold are highlighted, note that providing an exact reason why the expectation was violated is not possible, as that (qualitative) data was not collected.

Women, compared to men, are more willing to use the morning plane alternative (positive contribution). The effect found by Román et al. (2010) regarding a lower time parameter when comfort is higher could not be found, it was not significant and had a counter-intuitive sign. People with a higher education level do have a disutility for the night train. Therefore it is concluded the effect found in previous studies that higher educated people prefer the train could not be found. While the estimated sign for that hypothesis that regular train users prefer the night train is in line with expectation, it is not significant. However, a similar term for the morning plane was significant. Frequent train users do have a strong dislike of travelling by aeroplane in the early morning. This results in the finding that those travellers do choose more often for the night train. Therefore, it can be said that a bias in the sample is indeed present.

Table 7.9: Overview of expectations

Expectation	Expected sign	Estimated sign	Significant	Expectation hold
More travel time, more disutility	-	-	Yes	Yes
More travel cost, more disutility	-	-	Yes	Yes
Comfort combined positive	+	+	Yes	Yes
Women reluctant to take night train	-	+	No	No
Higher educated preference for night train	+	-	Yes	No
Regular train user prefers night train	+	+	No	No
Higher disposable income preference for faster modes	-	-	Yes	Yes
Increased age, more likely to choose night train	+	+	Yes	Yes
Business purpose disutility for night train and morning plane	-	-	Yes	Yes
Arrival time morning, higher preference for night train	+	+	Yes	Yes
Higher comfort, smaller travel time parameter	+	-	No	No
Business purpose, smaller cost parameter	+	+	Yes	Yes

7.3. Model combination

Both models for mode choice and comfort rating have been estimated, now the results can be combined to get some useful insights. In the comfort rating experiment, 'perceived comfort' was the dependent variable, while in the mode choice experiment this was one of the independent variables. In the same experiment, it was proved that several socio-demographics, such as age, and interactions with age and accommodation level also determined the comfort rating. An interaction with age and comfort level was not significant in the mode choice experiment. This makes the combined model less complicated, as the age-related effects are covered in the comfort rating experiment.

Both models can be coupled via the Willingness-to-Pay (WtP) for the different comfort levels. The mode choice experiment determined night train travel cost parameters for both leisure and business purposes and the utility contributions for the 1, 3 and 5-star comfort levels. This can be used to calculate the WtP to upgrade to a higher comfort level. Because of the assumption that the comfort attribute is continuous, the WtP can also be calculated for 2 and 4-star comfort levels. Table 7.10 lists the results.

Table 7.10: The Willingness to Pay for an increase in the comfort level, for both travel purposes

Comfort level	Utility Contribution	Utility Difference	WtP Business [€]	WtP Leisure [€]
1 star	0.6158			
2 stars	1.1672	0.551	73.03	49.10
3 stars	1.6542	0.487	64.50	43.37
4 stars	2.0768	0.423	55.97	37.63
5 stars	2.435	0.358	47.44	31.90

The results show that a business traveller would be willing to pay about €137 to improve the comfort level from 1 to 3 stars and €103 from 3 to 5 stars. These values are additive, meaning 1 to 5 stars WtP would be €240. Logically, as the night train travel cost parameter for the leisure purpose is larger, the WtP for an increase in the comfort level is less for the leisure purpose. Another important observation is that the utility difference and therefore WtP between 1-3 and 3-5 stars is not equal. This is the result of the quadratic component. People are willing to spend more money on getting the comfort from a 'very uncomfortable' to a 'medium comfort' level than on a 'medium comfort' to 'very comfortable' level. This is in line with what intuition would expect, i.e. people pay to get out of a very bad experience, but are less interested in upgrading to a 'premium' experience, especially for leisure purposes.

The two experiments can be connected by using the (absolute) linear regression coefficients of the comfort rating experiment and combining them with the previously calculated WtP values. Due to effects coding of categorical variables, it is important to realise that the difference in utility is twice the linear regression coefficients. For example: having no shower has a disutility of -0.127, and the availability of a shower an utility of 0.127. The added value of a shower is the difference, i.e. $2 * 0.127 = 0.254$ utility points. Therefore, the WtP for the shower, when one considers a 1-star comfort level and business purpose equals $0.254 \cdot 73.03 = €18.55$.

The same calculation for a change between accommodation types is more complicated. This is because the accommodation parameter values also depend on age. This was shown in Table 7.3. Together with the other comfort rating attributes the WtP of all of them can be calculated. Table 7.11 presents the results for both leisure and business travel purposes and a given comfort level for the train trip.

Table 7.11: WtP values, for both leisure and business purposes, for an increase in a comfort level attribute, given an comfort level (CL) of the night train.

Purpose				Business [€]				Leisure [€]			
Attribute	Level	Socio-Demographic	Difference	CL 1 ->2	CL 2 ->3	CL 3 ->4	CL 4 ->5	CL 1 ->2	CL 2 ->3	CL 3 ->4	CL 4 ->5
Base WtP				73.03	64.50	55.97	47.44	49.10	43.37	37.63	31.90
Accommodation	Seat -> Couchette	25 years	1.067	77.92	68.82	59.72	50.62	52.39	46.28	40.15	34.04
		45 years	0.887	64.78	57.21	49.65	42.08	43.55	38.47	33.38	28.30
		65 years	0.707	51.63	45.60	39.57	29.75	34.71	30.66	26.60	22.55
		80 years	0.572	41.77	36.89	32.01	27.14	28.09	24.81	21.52	18.25
	Couchette -> Sleeper	25 years	0.437	31.91	28.19	24.46	20.73	21.46	18.95	16.44	13.94
		45 years	0.437	31.91	28.19	24.46	20.73	21.46	18.95	16.44	13.94
		65 years	0.437	31.91	28.19	24.46	20.73	21.46	18.95	16.44	13.94
		80 years	0.437	31.91	28.19	24.46	20.73	21.46	18.95	16.44	13.94
People	2 levels	Highly Educated	0.265	19.35	17.09	14.83	12.57	13.01	11.49	9.97	8.45
		Not Highly Educated	0.241	17.60	15.54	13.49	11.43	11.83	10.45	9.07	7.69
Lock	No ->Yes		0.326	23.81	21.03	18.25	15.47	16.01	14.14	12.27	10.40
Shower	No ->Yes		0.254	18.55	16.38	14.22	12.05	12.47	11.01	9.56	8.10
Catering	None ->Kiosk		0.380	27.75	24.51	21.27	18.03	18.66	16.48	14.30	12.12
	Kiosk -> Restaurant car		0.032	2.34	2.064	1.79	1.52	1.57	1.39	1.20	1.02
Stops	2 levels		0.070	5.11	4.52	3.92	3.32	3.44	3.04	2.63	2.23

One should note that the linear attributes 'stops' and 'number of people' both have 3 different levels, so calculating the Willingness to Pay to reduce two levels just involves multiplying the found value by 2, e.g. the WtP for reducing the number of people from 6 to 2 for someone who is travelling for leisure, highly educated and a given night train comfort level of 3 stars equals $2 \cdot 9.97 = €19.94$.

However, one should take care when interpreting these values. While the values may give an indication of the order of magnitude for the maximum price that could be asked, they are also sensitive to changes in parameter values. For this reason, it may be of more interest to look at the comparisons between the comfort attributes.

Looking at the comparisons between the comfort attributes, one can note that people would be willing to pay the most for upgrades in accommodation levels. The previously found effect that with increasing age people see less added value in upgrading to a more comfortable accommodation type, explains the declining WtP values with age. Furthermore, people put the most value in the possibility to lock the travel compartment and the possibility to get food/drinks on the train. However, little value is seen in an upgrade in the dining experience (from 'kiosk' to a 'restaurant car') and a reduction in the number of stops during the night. This last finding is remarkable, it might have been difficult for respondents to assess how the number of stops impacted their travel experience.

7.4. Latent class choice model: identifying segments

Identifying segments in the dataset allows for getting useful insights in the mode choice behaviour from a marketing perspective. As argued, the marketing perspective provides answers to questions such as which groups are most likely to choose the night train. To do so, a latent class choice model is estimated. For estimating the latent class choice models LatentGOLD Choice software is used.

All previously included main attributes, plus the quadratic component for the comfort, are incorporated in the latent class model. To identify the classes in the dataset, several covariates are added. These are: previous international train or plane travel experience, gender, age, education, income, work status, travel purpose and train usage. Travel experience, income and train usage are incorporated using a linear effect, to others are taken into account as nominal variables using effects coding. Respondents that opted to withheld one of their answers for these questions were excluded by the program, resulting in 662 respondents.

Several latent class models, with a varying number of classes, were estimated. The estimation statistics such as the log-likelihood, BIC and R^2 score can be seen in Table 7.12. The lowest BIC score is that of the 7-class model. The difference in BIC scores with the other estimated models is large, and therefore it is chosen to work with a 7-class model.

Investigating the parameter estimates revealed that some parameters were insignificant. The program indicated that the parameter for the travel time of the morning plane is constant across all 7 classes, for this reason, it was fixed. However, the resulting parameter has a positive sign. At first sight, this seems invalid, because time is associated with disutility. Letting go of the fixation, reveals that some classes return a positive parameter, resulting in an overall positive parameter estimate when the parameter is fixed. Because the plane cost is one of the key determinants of morning plane utility, it is decided to keep it in the model. It is noted that the potential error is limited, as for all groups the relative importance of the morning plane travel time does not exceed 5%. Several other effects were insignificant for some classes, illustrating that some classes do not pay attention to some attributes. These parameters were removed using backward elimination, maximum one parameter was removed per class per iteration. The final model has a BIC score of 9941, which is higher than the initial model, meaning the insignificant parameters lead to over-fitting. The rho-squared score is .63, meaning that compared to a null model, the final model can explain 63 per cent of the choices made. For full results see Appendix F.

Table 7.12: Model fit statistics of the different latent class models.

Model	Loglikelihood	BIC	Rho-square
1-class	-7344.97	14741.90	0.184
2-class	-6119.40	12466.13	0.377
3-class	-5147.65	10698.00	0.515
4-class	-4858.23	10294.53	0.566
5-class	-4601.44	9956.34	0.613
6-class	-4492.69	9914.19	0.633
7-class	-4389.12	9882.43	0.652
8-class	-4314.20	9907.96	0.654
7-class final	-4506.11	9941.05	0.625

Respondents have a probability of 79.1% of belonging to the first three classes. The first two classes are about equally large, while the third class is +-18%. The remainder is split among the other 4 classes. The classes are described in the following bullets and characterised in Figure 7.2.

- Class 1: ‘Comfort minded people’. The probability of belonging to this class is 32.3%. People in this class put most attention to the comfort level of the night train. The relative importance (RI) of the comfort attributes is more than 60%. People in this group choose for the night train in 52% of the cases, followed by the morning plane (39%). One is most likely to be between 20 and 39 years old (42%), has a low income (36%), works full time (38%) and likely has not travelled by international train before (43%).
- Class 2: ‘Night train lovers’. One has a probability of belonging to this class of 29.1%. People in this class pay most attention to the cost of the plane ticket (RI: 38%), followed by night train comfort. Due to this fact, the night train is the favoured travel option. It is chosen almost every time (98%). People in this class are most likely to be older than 65 (49%), are retired (44%), have an average income and are highly educated (university master, 41%).
- Class 3: ‘Morning plane disfavoured’. 17.8% of the sample belongs to this class. The relative cost of the morning plane is most important (RI: 63%) for this group. Because this is a disutility, that means this mode is little used (3%). People in this group are more likely to be male (67%), have a university master (63%), are middle-aged (48% between 40 and 64), work fulltime (47%) and have a high income (33% more than €50.000).
- Class 4: ‘Time-sensitive people’. This group represents only 5.5% of the sample. People in this class only pay attention to travel time and arrival time. The arrival time context variable for the night train is most important (RI: 90%). For the morning plane option, they only consider the travel time and that has low importance (RI: 3%). Therefore, that is the favourite travel option (89%). In this group, there is a high probability one never travelled by international train (69%).
- Class 5: ‘Evening plane favoured’. Probability of being in this class is 8.8%. People in this class pay a lot of attention to the cost of the morning plane (RI: 76%), resulting in a large disutility. The travel time of the night train (RI: 7%) and comfort (RI: 16%) play a smaller role. This results in a preference for the evening plane (96%). There is equal probability (50%) to travel for business or leisure, compared to other classes in which the leisure purpose varied between 61-88%. People are most likely to be male (78%), earn a high income (44%) and hold a university master’s degree (56%).
- Class 6: ‘Plane users’. 5.4% of the sample belongs to this group. People in this group mostly choose aeroplane options (79%). They are cost-sensitive towards the morning plane (RI: 51%) and pay attention to night train comfort (RI: 36%). In that sense, this group is similar to class 3. However, the difference is in the covariates. People in this class are young (44% between 20-39), and most likely to earn a low income (35%),
- Class 7: ‘Frequent flyers’. The probability of belonging to this class is only 1.2%. People in this class do not opt for the night train (only 2%). Their choice behaviour is only described by the arrival time variables. 42% chooses the evening plane, 56% the morning plane. People in this class are middle-aged (62% between 40 and 64), hold a university degree (49%) and are frequent flyers (38% more than 6 times).

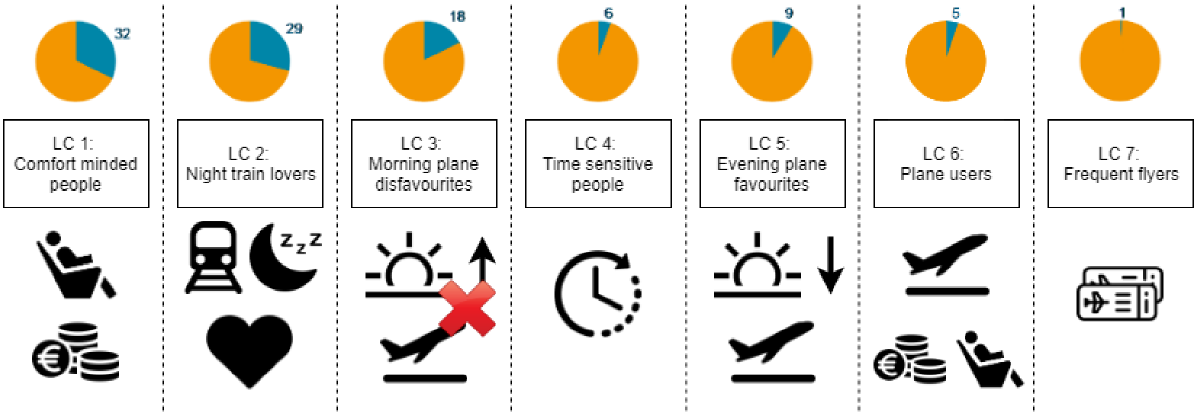


Figure 7.2: Characterisation of the different latent classes

Based on choice behaviour it is possible to identify the classes that are most likely to use the night train. Those are the 'comfort-minded people', 'night train lovers' and 'morning plane disfavoured'. To a lesser extent 'plane users' should also be considered, because they are price-sensitive and pay attention to comfort. The descriptions of the classes provide useful insights into which groups should be targeted when a night train is introduced. Due to the lack of flexibility in the used program, the interpretation or simulation of LC model is quite burdensome. This is why the latent class model was used solely to identify groups in the dataset.

7.5. Summary

Main takeaways

- The number of people in the train compartment is the major determinant of night train comfort.
- A basic night train with private accommodation for 2 people scores higher than a night train with luxury facilities such as showers, a restaurant car etc., but with shared accommodation for 6 people.
- The price attribute is the largest contributor to disutility, comfort level the largest positive contributor to the utility.
- The trip purpose is the most important background variable for mode choice. Business travellers have a significant penalty for both the night train and morning plane.
- Being a frequent train traveller results in a negative contribution to the morning plane alternative.
- The marginal Willingness to Pay for an increase in night train comfort level decreases with higher initial comfort levels.
- Several latent classes were identified. Most likely segments to use the night train are: ’comfort-minded people’, ’night train lovers’ and ’morning plane disfavoured’.

This chapter analysed the data and set-up a comfort rating model and a mode choice model. A linear regression model is chosen for the comfort rating experiment. It showed that the number of people for the shared accommodation is a major determinant of the ’perceived comfort’ rating for the night train. A basic night train with accommodation for 2 people scores better than a night train with shared accommodation for 6 people with all kinds of amenities (such as showers, restaurant car etc). Furthermore, it showed that, from a traveller perspective, upgrading catering facilities from a ’kiosk’ to ’restaurant car’ returns only a small higher comfort rating. Gender did not have a significant influence on the comfort rating, whereas age did. People tend to give less weight to the accommodation type when getting older.

For the mode choice experiment, several discrete choice models were estimated. In the end, it turned out that the Panel Mixed Logit model, including some socio-demographics, background variables and interactions fits the data best. That model explains about 35% of the initial variance. Of the main attributes, the price has the largest contribution to disutility. Furthermore, it showed that the trip purpose is very important for mode choice. People who travel for business are much less inclined to opt for the night train. It further showed that with age, people generally prefer the night train and that if you are a frequent train user you are less likely to choose the morning plane than the evening plane and night train alternatives.

Combining both models it was noted that the Willingness to Pay for an increase in night train comfort level depends on the trip purpose, and sits between €32 and about €73. Logically, people are more willing to pay to get out of a very uncomfortable experience than to upgrade to a premium experience. Of the comfort attributes, people are most willing to pay for an upgrade of their accommodation type and to reduce the number of people in the compartment. Reducing the number of stops during midnight, or upgrading existing catering facilities is not something people would spend a lot of money on.

A latent class model was estimated to see if segments for the mode choice could be identified. The latent class model with 7 classes performed best. Three classes were identified that are most likely to use the night train, these are ’comfort-minded people’, ’night train lovers’ and ’morning plane disfavoured’. To a lesser extent also a fourth group ’plane users’ was identified. The descriptions of these classes in terms of covariates provides useful insights from a marketing perspective.

8 Application

This chapter explores the Willingness to Use the night train for variations in attribute levels in Section 8.1. Furthermore, some scenarios are formulated and the effect of those on the potential night train market share is explored in Section 8.2. Section 8.3 provides a concise summary of this chapter.

8.1. Exploring the willingness to use the night train

The models have been estimated and utility contribution has been explored. This information can now be used to determine the probabilities of choosing the night train for various combinations of attributes. For exploring this range of the Willingness to Use the night train, every time one attribute is varied, while the others remain fixed. The fixed attributes are set on their middle level, or when not possible, to the lower end (i.e. the flight time = 5 hours and night train price = €80). The socio-demographics and background variables from the collected respondents are used, which means for each attribute variation in total 9648 predictions are made. The resulting predictions are averaged and reported in Table 8.1.

However, before interpreting the resulting values, the assumptions that were made for the Stated-Preference study are summarised. Only when these assumptions are considered the following numbers are meaningful. The trip duration is about 12-14 hours for the night train, this results (using conventional train technology) in about 1200 km. Travellers need to be at their European destination the next morning, either for leisure or business. They travel with hand luggage only, can be on-time at the departure point and they choose between flying or taking the night train, excluding other travel modes.
Table 8.1 shows that the comfort level has the biggest influence on the mode choice for the night train. The difference in the probability of choosing the night train ranges from 43% to a high 74%. Based on the utility contribution this is a logical finding. When the comfort level of the night train is decreased, the predicted market share of the other two modes does not increase equally. About 33% goes to the evening plane alternative and the remaining 66% is predicted to choose the morning plane alternative.

Other large changes in mode choice probabilities occur when varying the price of the night train tickets (about 22% difference in night train market share) and the ticket price for the early morning flight (about 13% difference in morning plane market share). In the first case about the same shift to the other two modes as previously found is present, while in the last case almost all mode switchers choose the night train, while only 2% is predicted to opt for the evening plane. At last, the effect of changes in the travel time for both plane and night train is relatively small.

These findings could lead to the conclusion that investing major amounts of money into new infrastructure to significantly reduce the night train journey time would only lead to a relatively small increase in the night train market share, and therefore not be a smart decision. However, note that the night train travel time and distances were based on conventional trains, not high-speed trains. Investing in new infrastructure would allow for operating those trains and reaching other destinations in the same

Table 8.1: Probabilities (Prob.) of choosing the night train (NT), morning plane (MP) or evening plane (EP) alternatives, for varying attribute levels. The 90% confidence intervals (CI) are also shown.

Attribute	Level	Prob. NT [%]	90% CI NT [%]		Prob. MP [%]	90% CI MP [%]		Prob. EP [%]	90% CI EP [%]	
Travel time train [hh:mm]	11:45	64.40	59.66	68.62	14.57	11.64	18.46	21.03	18.23	24.00
	13:00	61.87	56.60	66.66	16.18	12.75	20.16	21.95	18.82	25.38
	14:15	59.24	53.92	63.74	17.90	14.44	22.18	22.86	20.00	26.00
Travel cost train [€]	40	68.45	63.34	73.11	11.97	9.17	15.34	19.58	16.44	22.91
	80	61.87	56.60	66.66	16.18	12.75	20.16	21.95	18.82	25.38
	120	54.43	49.57	59.35	21.27	17.29	25.49	24.30	21.32	27.57
	160	46.36	40.59	52.04	27.06	22.19	32.16	26.59	22.99	30.25
Comfort level	1 star	42.57	37.61	48.07	29.75	25.00	34.69	27.68	23.92	30.89
	2 stars	53.13	47.60	58.45	22.05	17.75	26.72	24.82	21.48	28.13
	3 stars	61.87	56.60	66.66	16.18	12.75	20.16	21.95	18.82	25.38
	4 stars	68.54	63.98	72.81	11.97	9.27	15.14	19.49	16.56	22.71
	5 stars	73.52	69.32	77.24	9.09	6.91	11.84	17.39	14.67	20.27
Travel time plane [hh:mm]	05:00	61.87	56.60	66.66	16.18	12.75	20.16	21.95	18.82	25.38
	05:30	63.20	57.85	67.78	14.64	11.73	18.64	22.16	18.92	25.65
Travel cost plane [€]	60	54.77	49.08	60.01	24.42	20.13	29.49	20.82	17.99	23.90
	110	61.87	56.60	66.66	16.18	12.75	20.16	21.95	18.82	25.38
	160	67.09	61.98	71.24	10.23	7.86	13.62	22.68	19.31	26.09

time-frame. Instead, if considering only the current distance range, it would be better to invest in new comfortable trains, while keeping the night train ticket price competitive. It could also be considered to raise the base price of flight tickets by taxation.

8.2. Scenario application

Several scenarios are formulated for predicting the different mode shares. These scenarios are more detailed in the list below, and an overview is presented in Table 8.2.

1. **Reference:** this scenario represents the base scenario. The scenario is partly based on information gathered in an interview with Mrs Aalbers from the Ministry of Infrastructure and Water Management. During the interview, it became clear that the ministry sees no easy opportunities to significantly reduce the travel time of a potential night train service, as the infrastructure budget is fixed until 2030. One should also consider the night train would only run on Dutch territory for a limited amount of time. Furthermore, the ministry has little influence on the type of trains and the ticket price of night trains. (Mrs. Aalbers, personal communication, June 11, 2019). The ticket price is set to €80, this is based on real prices 2-3 months in advance for 'Couchette' accommodation, which was shown to result in average comfort levels. The travel time for the night train is chosen as 14:15 hours. The plane attributes are set to 5 hours and €110 ticket price.

2. **Upgraded trains:** the night trains are upgraded in this scenario. This means compared to the reference scenario the comfort level does increase to 4 stars. To accommodate for the costs of upgrading the trains, ticket prices are increased by €40 to €120. The aeroplane attributes remain unchanged.

3. **Fast, luxury trains and upgraded infrastructure:** this scenario represents a situation in which brand-new fast trains are bought, and a significant amount of money becomes available to upgrade the infrastructure. This allows an average speed of 100km/h for the night train service. To compensate for the investment cost in new trains and infrastructure, the night train ticket price is raised to €160. The aeroplane attributes remain unchanged.

4. **Current trains, optimised timetable:** compared to the reference scenario, the timetable is optimised. The result is that the average speed can increase to 90 km/h. This results in shorter travel time for the night train of 13 hours. All the other attributes remain unchanged.

5. **Luxury trains, optimised timetable:** compared to the previous scenario, these night trains have a higher comfort level. Likewise, the night train ticket price set to €120.

6. **Low-cost competitor:** this scenario positions the night train as a competitor to low-cost flights and bus services. This means the ticket price is really low (€40), but the comfort level is as well. No investments to the infrastructure are made and, as a low-cost flight competitor, the plane ticket price is set to €60.

Table 8.2: Different scenarios with the according attribute levels for the mode choice experiment and market shares for the night train (NT)

Scenario:	1. Reference	2. Upgraded trains	3. Fast, luxury trains, upgraded infrastructure	4. Current trains, optimised timetable	5. Luxury trains, optimised timetable	6. Low-cost competitor
Travel time train [hh:mm]	14:15	14:15	11:45	13:00	13:00	14:15
Travel cost train	€80	€120	€160	€80	€120	€40
'Perceived comfort'	3 stars	4 stars	5 stars	3 stars	5 stars	1 star
Travel time plane [hh:mm]	05:00	05:00	05:00	05:00	05:00	05:00
Travel cost plane	€110	€110	€110	€110	€110	€60
Market share NT						
Purpose business	46	48	53	48	56	27
Purpose leisure	63	63	67	66	71	43
Frequent train traveller	70	70	73	72	76	58
Non-frequent train traveller	59	59	63	62	68	39

Simulations are run to predict the market shares, under the same assumptions as previously mentioned, for the different scenarios. Similarly to the previous section, the socio-demographics from the sample are used as some of the social-demographics showed to be important contributors to the utility level. It is chosen to highlight the effect of being a frequent train traveller and the travel purpose on the predicted market shares. The simulation results can be seen in the lower part of Table 8.2. Additional information about the confidence intervals of the predicted market shares can be found in Appendix E.4.2.

Considering the results of the simulation one can see it is predicted that, under the assumptions, the night train could have a substantial market share. This is not unexpected, as the analysis of the gathered data in Chapter 6 already showed that respondents opted for the night train in about 55% of all choice sets.

As expected from the utility contribution, the travel purpose has a strong effect on the potential market share. The difference in night train market shares for the purposes is about 20%. The 'evening plane' alternative has a relatively stable 40% market share when travelling for business. For leisure, this alternative scores about half of that. Sometimes it is thought that the night train should compete with low-cost carriers. However, the low comfort level in this scenario leads to a substantial lower market share. The share drops between 20 and 25%, especially for business purpose the share drops to below 30%.

Frequent train travellers have a negative effect towards taking a plane in the early morning (see the utility contribution). They either choose the night train or the plane in the evening. Likewise this shows in the market share, for most scenarios the difference is about 8-11%, with exception of the low-cost scenario, in this case the market share difference is almost 20%. Results are as expected, the frequent train traveller chooses the night train more often.

Taking a more thorough look at the scenarios, it is noted that the most promising scenario is the one in which new luxury trains are introduced and the timetable is optimised. Surprisingly, upgrading the infrastructure is the runner up, the increased price to recover investment costs in infrastructure results in a lower market share. Considering this last scenario in more detail, it should be realised that the infrastructure is not upgraded easily. Usually this is a long and tedious process, taking several years. Therefore, when investigating if the night train is worth introducing, it is more realistic to consider the current infrastructure.

Two scenarios reduce the total travel time by optimising the timetable. Question remains what this means. Investigating current night train schedules, it is noticed that the current service between Dusseldorf and Vienna stops for almost an hour in the middle of the night. This time is included to split the train: after Nurnberg one part continues to Vienna, the other to Innsbruck. In theory it should be possible to reduce this couple/decouple time, either by reducing the buffer times or choosing for Electric Multiple Units when investing in new rolling stock. Furthermore, often the night train has to share the tracks with freighter trains in the middle of the night. This means the speed needs to lowered to be compatible with freighter services (Bird et al., 2017). Optimising the time table by reducing the effect of these two issues could lead to the reduction in travel time. However, this does need multi-country coordination to prevent that the night train has to wait at a border stop.

Considering an introduction time-line of these scenarios, some could be placed in chronological order. Introducing the night train, as-is, would result in the reference scenario. When the first initial experiences are evaluated and cross-country negotiations processing, and all partners recognise the service is a valuable addition to mobility, the time-tables could be optimised (scenario 4). If services are successful, investments could be made in the new rolling stock. Resulting in the final predicated market share of 71% for leisure purposes. This could be one of the time-lines in which a night train service could be introduced and improved. The market shares are shown in Figure 8.1.

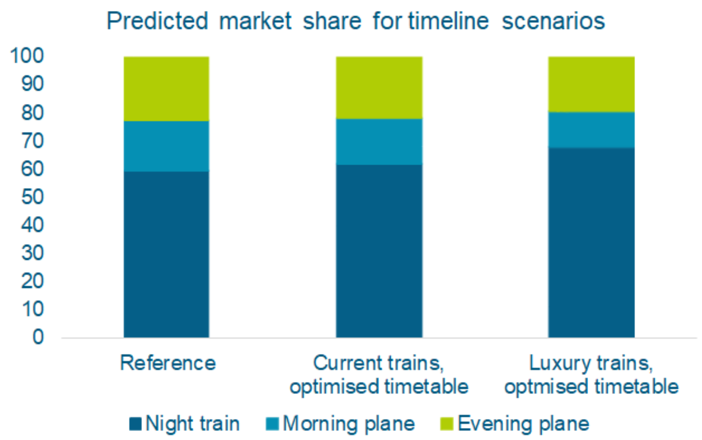


Figure 8.1: Mode market shares for the different time-line scenarios

Question remains if an indication could be given about the size of the market. Literature suggests one should be very careful with demand forecasting on Stated-Preference data, see for example Bates (1988); M. E. Ben-Akiva, Morikawa, and Shiroishi (1989); Fujii and Gärling (2003). Nevertheless, it is highly interesting for assessing a business case what passenger numbers could be expected. Once again it is stressed that care was taken in survey development, but that the numbers are only valid under those assumptions.

Additionally, the following steps and assumptions are made to arrive at passenger number predictions. In the introduction it was written that the number of aeroplane passengers between Amsterdam and Milan/Vienna was about 750.000 each. When assuming return trips, this results in 375.000 one-way passengers. To both these destinations around 11 to 14 flights depart daily, of which 2 or 3 in the early morning. This results in a percentage of about 20% if demand is uniform distributed (and the night train only captures those early morning passengers). Both leisure and business purposes are considered. Data analysis showed that 22% of the respondents travelled for business. This is used to calculated a 'combined potential market share'. Table 8.3 shows the potential number of night train passengers one could expect for the identified time-line.

Table 8.3: Potential numbers of passengers for a night train service.

Scenario	NT Combined Market share [%]	Passengers yearly [-]	Passenger daily [-]
1. Reference	59.23	44,423	122
4. Current trains, optimised timetable	61.89	46,418	127
5. Luxury trains, optimised timetable	67.67	50,753	139

Table 8.3 shows that it is predicted the night train could attract about 120 daily one-way passengers in the reference scenario, and with improvements for the identified time-line up to 139. In terms of train occupancy, the night train between Dusseldorf and Vienna has a maximum capacity of 246 passengers in the current composition and operates daily (vagonWEB, 2019). This would result in occupancy ranging between 49-57%, meaning the night train is not filled to capacity. Researching the operating cost of a night train is outside the scope of this research. However, generally speaking, a higher occupancy rate would be preferable. This can be achieved by taking into account that, when the line to Vienna gets extended towards the Netherlands, the night train could still pick-up passengers at the currently serviced German stops. Furthermore, one could also change the composition of the night train, which would be fairly easy considering it a combination of a power car and separate coaches.

8.3. Summary

Main takeaways

- The night train comfort level has the largest influence on potential market share, between 1 star and 5 star comfort is market share difference of 30%.
- Business travellers often (40%) opt for flying the day before and staying in a hotel.
- When introduced as-is, the night train could have a potential market share of 59%, resulting in about 122 daily passengers.
- Positioning the night train as alternative for low-cost flights, with similarly low comfort levels, leads to a significant drop (20%) in predicated market share.

This chapter applied the estimated mode choice model to predict the potential market share of the night train. A sensitivity analysis showed that the biggest influence on the night train market share is the comfort level attribute, between lowest and highest comfort level is a difference of almost 30%. Other large changes in the night train market share are caused by the travel cost of both night train and morning plane.

Scenario analysis showed that when introducing the night train as-is, the market share could be around 59%. There is a significant difference between leisure and business purposes. Business travellers often (40%) opt for travelling the day before and staying in a hotel. The night train market share for business travel purposes is about 14-18% lower.

A time-line was identified. After introduction the timetable could be optimised by reducing coupling/decoupling times and better coordinating freighter and passenger service trains. Doing so reduces the night time travel time, resulting in a slight increase in predicted market shares. Introducing new luxury night trains, increases the comfort level and as a result the predicated market share increases to about 71%. Positioning the night train as low-cost airline competitor, with similarly low comfort levels, is not favourable for the market share, it leads to a drop of about 20%.

Assuming year round-service and uniform demand results in daily passenger numbers between 122 and 139 (one-way) to either Vienna or Milan. This does mean that occupancy rates of the train will be low (around 50%). One should note though that the train composition can be easily changed and the train also serves the German market.

9 Conclusion

This chapter forms the conclusion of this study. Section 9.1 provides the key findings, followed by implications in Section 9.2 and limitations in Section 9.3. Finally, suggestions for further research are provided in Section 9.4.

9.1. Key findings

Travelling by night train has several advantages. One can relax and sleep during travelling, depart in the evening and arrive rested the next morning at the destination. Furthermore, taking the train at night offers to possibility to cover long distances and save a lot of emissions compared to travelling by aeroplane. Therefore, the night train might be a key step towards a more sustainable mobility. However, it is currently unknown what the willingness is of Dutch people to travel by night train, as an alternative for flying, for long-distance European travel.

To answer this question, this study uses a Stated-Preference approach. 804 respondents took part in the study. The survey consisted of 2 different experiments. First, respondents were asked to rate their 'perceived comfort' of different night train configurations in a comfort rating experiment. After which respondents were faced with a mode choice experiment that included night train, morning plane and evening plane alternatives. Data was analysed using Discrete Choice Modelling. An adapted form of the Hierarchical Information Integration theory was used to combine both modes. A latent class choice model was estimated to identify market segments.

The night train offers to possibility to sleep or relax during the journey, resulting in a higher comfort level. This comfort level might differ between people and contributes to the utility in mode choice. However, up to now, it was unknown what contributes to this 'perceived comfort' level of the night train. Analysis of the comfort rating experiment shows that the 'perceived comfort' rating on a 1 to 5-star scale is heavily influenced by the number of people in the shared accommodation. A non-stop night train with shared accommodation for 6 people with lockable compartments, showers, a restaurant car etc. would score lower than a basic night train that lacks all these 'luxury facilities', but does offer 2-person shared accommodation. With increasing age, people put less weight on the accommodation type (i.e. seat, couchette or sleeper). Current ÖBB night train services would score 3 stars in couchette class.

Using focus groups it was determined that people do not take previously unknown attributes into account when considering the night train travel option. Travel time and cost, together with comfort for the night train were proved to be significant explanatory variables for mode choice. People value one minute of travel time onboard the night train less than one minute onboard an aeroplane in the early morning.

When people need to be at their destination early (08:30), they opt more often for the night train. This is compared to a situation in which their preferred arrival time is 10:30. The travel purpose is an important background variable. If someone travels for business, one is much more likely to travel the evening before and stay in a hotel than travel in the early morning or by night train. People who regularly travel by train in daily life, are more likely to either travel the day before by aeroplane or take the night train. For this group, flying in the early morning is less preferred. Furthermore, with increasing age one is more likely to choose the night train and if one has completed higher education there is a preference for travelling by aeroplane the day before.

A latent class choice model provided insights into which groups are most likely to use the night train. These groups were identified as: 'comfort-minded people', 'night train lovers' and 'morning plane disfavoured'. To a lesser extent also a fourth group, 'plane users', was identified to be interested in using the night train. The 'comfort-minded people' find the comfort factor of the night train highly important and are mostly young. The 'night train lovers' are people who are older than 65, retired and have a high education level. In the 'morning plane disfavoured' class, people are middle-aged, highly educated and earn a high income.

Under the assumptions made for this study, the potential market share of the night train is explored. It shows that the night train market share is most sensitive for the comfort level of the night train, followed by the price of the night train tickets. When the night train is introduced as-is, combined market share for both travel purposes could be just below 60%. Reducing the travel time by optimising the timetable, for example by better coordination between passenger night and freight services and reducing coupling/decoupling time, allows for a slight increase to above 60% market share. Finally, when service is successful, new investments in comfortable, luxury trains could be made. This would result in a further increase in the market share to about 67%. It is highlighted that positioning the night train as an alternative for low-cost flights, with similarly low comfort and low prices, is not favourable for the market share. In that case, the predicted share is almost 20-25% lower than in other scenarios.

Using the scenarios and assuming year-round service and uniform demand, daily one-way (Dutch) passengers could be between 120 and 140 to either Vienna or Milan. This would result in an occupancy rate of about 50% using the current night train configuration. While higher rates are preferred, one should consider that a night train to Milan or Vienna also serves the German market and the night train configuration is easily adapted.

As substantial number of people indicate that they would choose the night train under the given context. One can therefore conclude that people do indeed seem willing to use the night train, as an alternative for flying, for long-distance trips in Europe.

9.2. Implications

It is shown that introducing a night train service to replace aeroplane flights could be a worthwhile option. Leisure travellers seem to be most attracted to the night train travel option. It remains to be seen if it really reduces the number of airline passengers, and therefore the number of flights, between the city pairs that would be connected by the night train. In theory, the night train could also generate new demand instead of substituting aeroplane passengers. However, previous research suggests this effect is limited to a specific group of 'night train enthusiasts' and further research is outside the scope of this study (Bird et al., 2017). Some measures aimed at the aeroplane ticket price, such as the proposed departure tax or growing astonishment about tax-free kerosene for airlines could tip the scales in favour of the night train.

During the finalisation of this thesis, KiM Netherlands Institute for Transport Policy Analysis released a report about the potential of night train services to and from the Netherlands (Savelberg, 2019). The report states that the total number of trips in the long-term could be around 0.7-1.0 million trips. Compared to the results in this report, this, at first sight, seems high. However, the KiM report considers a broader number of destinations (8) and both trips to and from the Netherlands. Furthermore, they state they have no information about the demand side of the market. This study is the first to provide more information and they could be considered to be supplementary.

There are also some implications for the night train service itself. Simulations showed that spending a lot of money on upgrading the infrastructure might not be the best way to increase the potential market share of the night train. Positioning the night train as a low-cost alternative would only result in a market share drop. From a traveller perspective, it would be better to increase the comfort level. As travellers are most sensitive for the number of people in the compartment, this could best be done by introducing more private compartments, switching away from the 6-person compartments. Introducing compartments for solo travellers might also attract more business-oriented traffic. However, the cost perspective of these decisions is outside the scope of this study and would need further research.

Another implication is that if the night train gets introduced, it needs to be a visible travel option. While not taken into account in the choice experiments, during the focus group meetings it was reported that it is important to compare the travel options and have a streamlined booking process. Selling the tickets via a Global Distribution System allows for increasing the night train visibility on airfare comparison websites. Teaming up with travel organisations, such as TUI or Sunweb, could offer packages of city trips using the night train, making it easier for travellers to choose the night train instead of the 'standard' aeroplane option.

9.3. Limitations

This study does contain several limitations which should be taken into account when considering the results. This study considered a set context, i.e. a trip within Europe of about 12-14 hours, for either business or leisure, solo or with someone else and with two arrival time contexts, both in the early morning. This is considered to be a business case for the night train, but says nothing about the demand for such services if people would like to arrive around or after midday or travel to destinations outside this 12-14 hour scope. For trips outside this scope, further research is needed.

Another limitation for this study is the usage of the NS panel for gathering respondents. While this provided a lot of data for estimating the choice models, it also meant that people that took part are (very) regular train users in daily life, probably one of the most loyal train users as they are willing to spend their own time on answering surveys for NS. This means a bias could be present in the dataset. This showed in the finding that people who are regular train travellers have a dislike for the morning plane alternative, which resulted in a higher night train market shares. Similar significant effects regarding socio-demographics are found high education, high income, being a student and age.

As always with any Stated-Preference study, it remains to be seen if people opt for the night train in reality. As the night train travel option does not exist, it is hard for people to completely imagine the travel experience. This, for example, showed in the low Willingness to Pay for the reduction of stops during the middle of the night, did people realise this could wake them up? Their decisions could also be influenced by a negative experience, after which they are very reluctant to try the service again. Something that is neglected in the current study.

Usually, leisure travellers (which are more likely to use the night train) travel together and make decisions together. This group decision making is not modelled. Also, it might be that during this study, social desirability played a role while answering the survey: during the research period, public debates and elections took place. During these events, climate (change) was often the topic and keywords such as 'vliegschaamte' (shame of flying) were introduced in media. Another factor to take into account is that the offered choice set in the study is not complete, i.e. if people are in principle against flying or have a fear of flying, they would have chosen the night train. In reality, they might also opt for a day train or other transport mode.

9.4. Suggestions for further research

This study was the first study of its kind into the Willingness to Use night trains as an alternative for flying. This leaves several suggestions for further research. One of the suggestions would be to distribute the survey with a more varied public. This study significantly found that people with a higher education level are less likely to choose the night train, and regular train users and older people more likely to use the night train. Groups who were also highly present in the sample. This outcomes for this study are therefore valid for this group. When the study is distributed in a more general public, it would result in a potential market share that would be more in line with reality. Furthermore, for assessing the business case it would be a good idea to include questions that would try to reveal if people intend to use this service regularly or just once as a try-out.

Some people have a fear of flying or other reason to dislike the aeroplane. Including questions regarding these issues and using the results to adapt the choice set, solves the problem of those respondents only having one realistic option: the night train. Additionally, attitudinal statements (for example regarding the environment) could be included to investigate these kinds of effects on night train usage.

Further research is needed into the effect of the preferred arrival times outside the used context (i.e. arrival around/after midday). This would provide more information about potential night train usage. For an operator, this additional information would be useful for making decisions regarding the introduction of such services.

Using flexible software, such as PythonBiogeme or PandasBiogeme, it is possible to set-up more advanced latent class choice models. This would allow to further explore the segments in the dataset and related covariates. These models can then be used to investigate how the specific segments would react to different scenarios.

Several implications regarding the night train product from a traveller perspective were provided. However, it is unknown if these recommendations are cost-effective. Further research is needed into the cost/operator perspective of these decisions.

As said, this study used a fixed context, with travel times of up 14 hours. Using conventional train technology, this meant the study considered destinations such as Amsterdam - Milan or Vienna. Recently, China introduced new high-speed sleeper trains. Increasing the speed allows for linking cities further away during the night. City pairs such as Amsterdam - Barcelona or Lisbon might become a possibility using high-speed technology. As these cities are further away, travelling by aeroplane also takes longer and makes the night train a more competitive option. This could be an interesting topic for further research.

At last, a suggestion would be to repeat a similar study in a country where the night train is already introduced, such as Germany or Austria. The collected Stated-Preference data could then be compared to actual Revealed-Preference data to calibrate the choice model, allowing for more accurately predicting the potential market share of the night train.

References

Airbus. (2018). *Global Market Forecast 2018-2037.* Retrieved from https://www.airbus.com/aircraft/market/global-market-forecast.html

Álvarez, A. G. (2010, 1). Energy Consumption and Emissions of High-Speed Trains. *Transportation Research Record: Journal of the Transportation Research Board*, *2159*(1), 27–35. Retrieved from http://journals.sagepub.com/doi/10.3141/2159-04 doi: 10.3141/2159-04

Anable, J. (2005, 1). 'Complacent Car Addicts' or 'Aspiring Environmentalists'? Identifying travel behaviour segments using attitude theory. *Transport Policy*, *12*(1), 65–78. Retrieved from https://www.sciencedirect.com/science/article/pii/S0967070X0400054X doi: 10.1016/J.TRANPOL.2004.11.004

Arentze, T. A., & Molin, E. J. (2013, 12). Travelers' preferences in multimodal networks: Design and results of a comprehensive series of choice experiments. *Transportation Research Part A: Policy and Practice*, *58*, 15–28. Retrieved from https://www.sciencedirect.com/science/article/pii/S0965856413001857?via%3Dihub doi: 10.1016/J.TRA.2013.10.005

Bates, J. (1988). Econometric issues in stated preference analysis. *Journal of Transport Economics and Policy*, *22*(1), 59–69. Retrieved from https://www.scopus.com/record/display.uri?eid=2-s2.0-0001884340&origin=inward&txGid=a59808cd1125562feff8d46c31f2ff32

Ben-Akiva, M., & Bierlaire, M. (1999). Discrete Choice Methods and their Applications to Short Term Travel Decisions. In (pp. 5–33). Springer, Boston, MA. Retrieved from http://link.springer.com/10.1007/978-1-4615-5203-1_2 doi: 10.1007/978-1-4615-5203-1{_}2

Ben-Akiva, M., McFadden, D., & Train, K. (2019). Foundations of Stated Preference Elicitation: Consumer Behavior and Choice-based Conjoint Analysis. *Foundations and Trends® in Econometrics*, *10*(1-2), 1–144. Retrieved from http://www.nowpublishers.com/article/Details/ECO-036 doi: 10.1561/0800000036

Ben-Akiva, M., & Swait, J. (1986). The Akaike Likelihood Ratio Index. *Transportation Science*, *20*(2), 133–136. Retrieved from https://doi.org/10.1287/trsc.20.2.133 doi: 10.1287/trsc.20.2.133

Ben-Akiva, M. E., Morikawa, T., & Shiroishi, F. (1989). Analysis of the reliability of stated preference data in estimating mode choice models. In *Transport policy, management & technology towards 2001: Selected proceedings of the fifth world conference on transport research* (Vol. 4). Retrieved from https://trid.trb.org/view/338877

Bhat, C. R. (1998, 8). Analysis of travel mode and departure time choice for urban shopping trips. *Transportation Research Part B: Methodological*, *32*(6), 361–371. Retrieved from https://www-sciencedirect-com.tudelft.idm.oclc.org/science/article/pii/S0191261598000046 doi: 10.1016/S0191-2615(98)00004-6

Bierlaire, M. (2018). *PandasBiogeme: a short introduction* (Tech. Rep.). Transport and Mobility Laboratory, ENAC, EPFL. Retrieved from http://transp-or.epfl.ch/documents/technicalReports/Bier18.pdf

Bird, G., Collins, J., Da Settimo, N., Dunmore, D., Ellis, S., Khan, M., … Nistri, D. (2017). *Research for TRAN Committee – Passenger night trains in Europe: the end of the line?* (Tech. Rep.). Brussels: European Parliament, Policy Department for Structural and Cohesion Policies, Brussels. Retrieved from http://www.europarl.europa.eu/RegData/etudes/STUD/2017/601977/IPOL_STU(2017)601977_EN.pdf

Boeing. (2018). *Commercial Market Outlook 2018-2037.* Retrieved from https://www.boeing.com/commercial/market/commercial-market-outlook/

Börjesson, M., & Eliasson, J. (2014, 1). Experiences from the Swedish Value of Time study. *Transportation Research Part A: Policy and Practice*, *59*, 144–158. Retrieved from https://www.sciencedirect.com/science/article/pii/S0965856413002103#f0005 doi: 10.1016/J.TRA.2013.10.022

Bos, I. D. M., Van der Heijden, R. E. C. M., Molin, E. J. E., & Timmermans, H. J. P. (2004, 9). The Choice of Park and Ride Facilities: An Analysis Using a Context-Dependent Hierarchical Choice Experiment. *Environment and Planning A: Economy and Space*, *36*(9), 1673–1686. Retrieved from `http://journals.sagepub.com/doi/10.1068/a36138` doi: 10.1068/a36138

Brownstone, D., & Small, K. A. (2005, 5). Valuing time and reliability: assessing the evidence from road pricing demonstrations. *Transportation Research Part A: Policy and Practice*, *39*(4), 279–293. Retrieved from `https://linkinghub.elsevier.com/retrieve/pii/S096585640400103X` doi: 10.1016/j.tra.2004.11.001

Buehler, R. (2011). Determinants of transport mode choice: a comparison of Germany and the USA. *Journal of Transport Geography*. doi: 10.1016/j.jtrangeo.2010.07.005

Burgdorf, C., Eisenkopf, A., & Knorr, A. (2018, 9). User acceptance of long distance bus services in Germany. *Research in Transportation Economics*, *69*, 270–283. Retrieved from `https://www.sciencedirect.com/science/article/abs/pii/S0739885917303256` doi: 10.1016/J.RETREC.2018.07.023

Centraal Bureau voor de Statistiek. (2019). *Bijna 80 miljoen passagiers op Nederlandse luchthavens.* Retrieved from `https://www.cbs.nl/nl-nl/nieuws/2019/06/bijna-80-miljoen-passagiers-op-nederlandse-luchthavens`

ChoiceMetrics. (2018). *Ngene 1.2 User Manual & Reference Guide* (Tech. Rep.). Retrieved from `www.choice-metrics.com`

Chorus, C. (2017). *SEN1221: Statistical analysis of choice behaviour - lecture slides.* Delft: TU Delft Brightspace.

Chorus, C. G., Rose, J. M., & Hensher, D. A. (2013, 2). Regret Minimization or Utility Maximization: It Depends on the Attribute. *Environment and Planning B: Planning and Design*, *40*(1), 154–169. Retrieved from `http://journals.sagepub.com/doi/10.1068/b38092` doi: 10.1068/b38092

DB International, & International Union of Railways (UIC). (2013a). *Night Trains 2.0 - New opportunities by HSR? - Study* (Tech. Rep.). Berlin. Retrieved from `https://www.shop-etf.com/en/uic-study-night-trains-brochure-new-opportunities-by-hsr-full-report`

DB International, & International Union of Railways (UIC). (2013b). *Night Trains 2.0 - New opportunities by HSR? - Summary* (Tech. Rep.). Berlin. Retrieved from `https://www.shop-etf.com/en/uic-study-night-trains-2-0-final-presentation-new-opportunities-by-hsr-executive-summary`

de Jong, G., Daly, A., Pieters, M., Vellay, C., Bradley, M., & Hofman, F. (2003, 5). A model for time of day and mode choice using error components logit. *Transportation Research Part E: Logistics and Transportation Review*, *39*(3), 245–268. Retrieved from `https://www-sciencedirect-com.tudelft.idm.oclc.org/science/article/pii/S1366554502000376` doi: 10.1016/S1366-5545(02)00037-6

Donners, B. (2018). *Trein is goed alternatief voor korte vluchten vanaf Schiphol* (Tech. Rep.). Amersfoort: Royal HaskoningDHV. Retrieved from `https://www.royalhaskoningdhv.com/nl-nl/nederland/nieuws/nieuwsberichten/trein-is-goed-alternatief-voor-korte-vluchten-vanaf-schiphol/8122`

European Environment Agency. (2018). *Aviation and shipping — impacts on Europe's environment* (Tech. Rep.). Luxembourg: Publications Office of the European Union. Retrieved from `https://www.eea.europa.eu/publications/term-report-2017` doi: 10.2800/4907

Eurostat. (2017). *Air passenger transport between the main airports of the Netherlands and their main partner airports (routes data) (avia_par_nl).* Retrieved from `https://ec.europa.eu/eurostat/data/database`

Fujii, S., & Gärling, T. (2003, 5). Application of attitude theory for improved predictive accuracy of stated preference methods in travel demand analysis. *Transportation Research Part A: Policy and Practice*, *37*(4), 389–402. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0965856402000320#BIB12` doi: 10.1016/S0965-8564(02)00032-0

Hensher, D. A. (2010, 7). Hypothetical bias, choice experiments and willingness to pay. *Transportation Research Part B: Methodological*, *44*(6), 735–752. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0191261509001477#bib13` doi: 10.1016/J.TRB.2009.12.012

Hensher, D. A., & Rose, J. M. (2007). Development of commuter and non-commuter mode choice models for the assessment of new public transport infrastructure projects: A case study. *Transportation Research Part A: Policy and Practice*. doi: 10.1016/j.tra.2006.09.006

Johnson, R. M., & Orme, B. K. (1996). *How Many Questions Should You Ask in Choice-Based Conjoint Studies?* (Tech. Rep.). Retrieved from www.sawtoothsoftware.com

Kroesen, M. (2018). *SEN1721: Travel Behaviour Research lecture 1 - 2.* Delft: TU Delft Brightspace.

Kröger, S. (2018). *Motie van het lid Kröger over aansluiting op het Europese nachtnet | Tweede Kamer der Staten-Generaal.* Retrieved from https://www.tweedekamer.nl/kamerstukken/moties/detail?id=2018Z02981&did=2018D06032

Lechner, R. (2018). *ÖBB-Bilanz 2017: Mit vollem Einsatz auf Erfolgskurs.* Retrieved from https://presse.oebb.at/de/presseinformationen/oebb-bilanz-2017-mit-vollem-einsatz-auf-erfolgskurs

Li, Y.-W. (2015). Evaluating the Urban Commute Experience: A Time Perception Approach. *Journal of Public Transportation*, *6*(4), 41–67. Retrieved from https://scholarcommons.usf.edu/cgi/viewcontent.cgi?article=1390&context=jpt doi: 10.5038/2375-0901.6.4.3

Limtanakool, N., Dijst, M., & Schwanen, T. (2006a, 9). The influence of socioeconomic characteristics, land use and travel time considerations on mode choice for medium- and longer-distance trips. *Journal of Transport Geography*, *14*(5), 327–341. Retrieved from https://www.sciencedirect.com/science/article/pii/S0966692305000438 doi: 10.1016/J.JTRANGEO.2005.06.004

Limtanakool, N., Dijst, M., & Schwanen, T. (2006b, 9). The influence of socioeconomic characteristics, land use and travel time considerations on mode choice for medium- and longer-distance trips. *Journal of Transport Geography*, *14*(5), 327–341. Retrieved from https://www.sciencedirect.com/science/article/pii/S0966692305000438#bib40 doi: 10.1016/J.JTRANGEO.2005.06.004

Louviere, J. J. (1984). Hierarchical Information Integration: a New Method For the Design and Analysis of Complex Multlattribute Judgment Problems. *Advances in Consumer Research*, *11*, 148–155. Retrieved from http://acrwebsite.org/volumes/6233/volumes/v11/NA-11

Lynch, G., & Atkins, S. (1988). The influence of personal security fears on women's travel patterns. *Transportation*, *15*(3), 257–277. Retrieved from http://link.springer.com/10.1007/BF00837584 doi: 10.1007/BF00837584

McFadden, D. (1972). *Conditional logit analysis of qualitative choice behavior.* Berkeley, California. Retrieved from https://eml.berkeley.edu/reprints/mcfadden/zarembka.pdf

Mokhtarian, P. L., & Salomon, I. (2001, 9). How derived is the demand for travel? Some conceptual and measurement considerations. *Transportation Research Part A: Policy and Practice*, *35*(8), 695–719. Retrieved from https://www.sciencedirect.com/science/article/pii/S0965856400000136 doi: 10.1016/S0965-8564(00)00013-6

Molin, E., & van Gelder, M. (2008, 1). Freeway Access to Public Transport. *Transportation Research Record: Journal of the Transportation Research Board*, *2076*(1), 106–113. Retrieved from http://journals.sagepub.com/doi/10.3141/2076-12 doi: 10.3141/2076-12

Morikawa, T., Ben-Akiva, M., & Mcfadden, D. (2002). Advances in Econometrics Discrete choice models incorporating revealed preferences and psychometric data. *Advances in Econometrics*. Retrieved from https://doi.org/10.1016/S0731-9053(02)16003-8 doi: 10.1016/S0731-9053(02)16003-8

Nederlandse Spoorwegen. (2016). *Bye Bye City Night Line.* Retrieved from https://nieuws.ns.nl/bye-bye-city-night-line/

Noland, R. B., & Small, K. A. (1995). Travel demand forecasting, travel behavior analysis, time-sensitive transportation, and traffic assignment methods. *Transportation Research Record*(1493), 150–158. Retrieved from https://trid.trb.org/view/452605

Norojono, O., & Young, W. (2003, 4). A Stated preference freight mode choice model. *Transportation Planning and Technology*, *26*(2), 1–1. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/715020600 doi: 10.1080/715020600

OBB. (2018a). *Destinations.* Retrieved from https://www.nightjet.com/en/reiseziele.html

OBB. (2018b). *Nightjet of the future.* Retrieved from https://www.nightjet.com/en/ausstattung/nightjetzukunft.html

ÖBB. (2019). *Nightjet - Fahrplan 2019.* Retrieved from https://www.nightjet.com/dam/jcr:

51e3c07e-c91e-46d6-beb4-0dce62d78347/nightjet-oesterreich-2019.pdf
Ohnmacht, T., & Scherer, M. (2010, 1). More Comfort, Shorter Travel Time, or Low Fares? *Transportation Research Record: Journal of the Transportation Research Board*, *2143*(1), 100–107. Retrieved from http://journals.sagepub.com/doi/10.3141/2143-13 doi: 10.3141/2143-13
Orme, B. (2010). Sample Size Issues for Conjoint Analysis. In *Getting started with conjoint analysis: Strategies for product design and pricing research* (2nd ed., pp. 57–66). Madison, Wisconsin: Research Publishers LCC. Retrieved from https://www.sawtoothsoftware.com/download/techpap/samplesz.pdf
Ortúzar S., J. d. D., & Willumsen, L. G. (2011). *Modelling transport* (4th ed.). Wiley-Blackwell. Retrieved from https://www.wiley.com/en-nl/Modelling+Transport,+4th+Edition-p-9780470760390
OVPro. (2016). *Slaaptrein City Night Line houdt op te bestaan | OVPro.nl.* Retrieved from https://www.ovpro.nl/trein/2016/12/09/slaaptrein-city-night-line-houdt-op-te-bestaan/?gdpr=accept
Ozbay, K., & Yanmaz-Tuzel, O. (2008, 5). Valuation of travel time and departure time choice in the presence of time-of-day pricing. *Transportation Research Part A: Policy and Practice*, *42*(4), 577–590. Retrieved from https://www.sciencedirect.com/science/article/pii/S0965856408000049 doi: 10.1016/J.TRA.2007.12.002
Paulssen, M., Temme, D., Vij, A., & Walker, J. L. (2014, 7). Values, attitudes and travel behavior: a hierarchical latent variable mixed logit model of travel mode choice. *Transportation*, *41*(4), 873–888. Retrieved from http://link.springer.com/10.1007/s11116-013-9504-3 doi: 10.1007/s11116-013-9504-3
Pel, A. (2017). *CIE5811: Advanced Transportation Modelling lecture 1.* Delft: TU Delft Brightspace.
PWC. (2016). *Staying power, European cities hotel forecast for 2016 and 2017* (Tech. Rep.). Retrieved from https://www.pwc.com/gx/en/hospitality-leisure/pdf/european-cities-hotel-forecast-2016-2017.pdf
PWC. (2018). *Best placed to grow? European cities hotel forecast for 2018-2019* (Tech. Rep.). Retrieved from https://www.pwc.fr/fr/assets/files/pdf/2018/04/pwc-european-cities-hotel-forecast-for-2018-and-2019.pdf
Román, C., Espino, R., & Martín, J. C. (2010, 1). Analyzing competition between the high speed train and alternative modes. The case of the Madrid-Zaragoza-Barcelona corridor. *Journal of Choice Modelling*, *3*(1), 84–108. Retrieved from https://www.sciencedirect.com/science/article/pii/S1755534513700307 doi: 10.1016/S1755-5345(13)70030-7
Rome2rio. (n.d.). *Rome2rio: discover how to get anywhere.* Retrieved from https://www.rome2rio.com/
Savelberg, F. (2019). *Slapend onderweg: potentieel van de internationale nachttrein van en naar Nederland* (Tech. Rep.). Den Haag: Kennisinstituut voor Mobiliteitsbeleid. Retrieved from https://www.kimnet.nl/publicaties/rapporten/2019/06/20/slapend-onderweg-potentieel-van-de-internationale-nachttrein-van-en-naar-nederland
Schiphol Group. (2018). *Our most important traffic and transport figures.* Retrieved from https://www.schiphol.nl/en/schiphol-group/page/traffic-review/
Small, K. A. (2012, 12). Valuation of travel time. *Economics of Transportation*, *1*(1-2), 2–14. Retrieved from https://www.sciencedirect.com/science/article/pii/S2212012212000093 doi: 10.1016/J.ECOTRA.2012.09.002
Stradling, S., Carreno, M., Rye, T., & Noble, A. (2007, 7). Passenger perceptions and the ideal urban bus journey experience. *Transport Policy*, *14*(4), 283–292. Retrieved from https://www.sciencedirect.com/science/article/pii/S0967070X07000182 doi: 10.1016/J.TRANPOL.2007.02.003
United Nations. (2018). *What is the Paris Agreement? | UNFCCC.* Retrieved from https://unfccc.int/process-and-meetings/the-paris-agreement/what-is-the-paris-agreement
vagonWEB. (2019). *NightJet 40421 train composition.* Retrieved from https://www.vagonweb.cz/razeni/vlak.php?zeme=OeBB&kategorie=NJ&cislo=40421&nazev=&rok=2019
Vredin Johansson, M., Heldt, T., & Johansson, P. (2006, 7). The effects of attitudes and per-

sonality traits on mode choice. *Transportation Research Part A: Policy and Practice*, *40*(6), 507–525. Retrieved from https://www.sciencedirect.com/science/article/pii/S0965856405001473 doi: 10.1016/J.TRA.2005.09.001

Whitehead, J. C., Pattanayak, S. K., Van Houtven, G. L., & Gelso, B. R. (2008, 8). COMBINING REVEALED AND STATED PREFERENCE DATA TO ESTIMATE THE NONMARKET VALUE OF ECOLOGICAL SERVICES: AN ASSESSMENT OF THE STATE OF THE SCIENCE. *Journal of Economic Surveys*, *22*(5), 872–908. Retrieved from http://doi.wiley.com/10.1111/j.1467-6419.2008.00552.x doi: 10.1111/j.1467-6419.2008.00552.x

Zijlstra, T., & Huibregtse, O. (2018). *De Vliegende Hollander | Document (onderzoekspublicatie) | Kennisinstituut voor Mobiliteitsbeleid.* Retrieved from https://www.kimnet.nl/publicaties/rapporten/2018/03/22/de-vliegende-hollander

A

Scientific paper

This appendix provides the scientific paper that summarises the study.

Willingness to Use night trains for European long-distance travel

M.K. Heufke Kantelaar, BSc

Graduate student of Transport, Infrastructure & Logistics, Faculty of Civil Engineering & Geosciences, Delft University of Technology, Stevinweg 1, 2628CN, Delft, The Netherlands

Abstract

Limiting climate change forces us to switch to more sustainable mobility. Travelling by train is more sustainable. Night trains might be a solution, they offer several advantages such as a higher comfort level and can travel long-distances during the night. However, there currently is no knowledge about the Willingness to Use night trains. This paper is the first study into this Willingness to Use night trains, as an alternative for flying, for long-distance European travel. To do so, this paper makes use of two Stated-Preference experiments. A comfort rating experiment, in which the comfort rating is the dependent variable. In this experiment, it is explored to which extent night train characteristics influence the 'perceived comfort' rating. In the mode choice experiment, it is turn investigated how this comfort rating is traded-off against more traditional mode choice attributes such as trip time and trip costs. The paper presents the results of linear regression and a Panel Mixed Logit model, estimated on 804 collected responses in the Netherlands. The results of both models are combined to derive Willingness-to-Pay values for an improvement in one of the comfort attributes, given an initial comfort level. Furthermore, several segments were identified using a latent class choice model. At last, the Willingness to Use the night train is explored for several scenarios. It is shown that when the night train is introduced as-is, it is predicted to have a market share of about 60%. Positioning the night train as a low-cost competitor results in a significant drop in market share.

Keywords: Night train, Perceived comfort, Long-distance mode choice, Panel Mixed Logit model, Hierarchical Information Integration

1. Introduction

Travelling by aeroplane is a major contributor to human-induced greenhouse gas emissions. To reduce the impact of climate change, a switch to more sustainable mobility is needed. Travelling by train is much more friendly for the environment, it can save up to five times the CO_2 emissions compared to an aeroplane (Álvarez, 2010). However, travelling by train is usually far slower making the number of possible destinations limited. This poses a question if people are willing to use night trains for European long-distance travel.

Scientifically, there is little to no knowledge about the Willingness to Use night trains for European long-distance trips. To the best of the authors' knowledge, no study has been done directly addressing this issue. However, mode choice, in general, is a widely studied topic in travel behaviour. Stated-Preference studies into mode choice typically include a variety of attributes, such as trip time, access/egress time, trip cost, waiting time, departure/arrival times, frequency and the number of transfers (e.g. Bhat (1998); Morikawa et al. (2002); de Jong et al. (2003);

Hensher and Rose (2007); Román et al. (2010); Paulssen et al. (2014)). The comfort of the travel mode is not often taken into account. Román et al. (2010) used it in a mode choice study comparing aeroplane to High-Speed Train alternatives for the Madrid - Barcelona corridor, in which it is found to be a significant effect.

Based on these findings, it poses the question of how this relates to mode choice including the night train. One can argue that the comfort level of the night train is higher compared to an aeroplane. Comfort, in general, is expressed as a star rating (e.g. hotels). However, the level of comfort is very subjective, it differs between persons. This 'perceived comfort' level can be included in a traditional choice experiment. This would allow examining trade-offs with other included attributes, such as trip time or costs. Extending this to gather knowledge about the Willingness-to-Pay (WtP) for an increase in this 'perceived comfort' level is possible.

However, only including the 'perceived comfort' attribute in a mode choice experiment does not provide any information about how travellers arrive at this comfort rating. Therefore, it is proposed to construct an additional experiment, a comfort rating experiment, in which respondents are asked to rate different night train configurations. These night train configurations differ in various attributes such as the accommodation type, number of people in the compartment or the possibility to shower. The results of this experiment provide insight into which measures could be taken to improve the perception of the night train comfort level. The proposed approach is based on the Hierarchical Information Integration theory, originally introduced by Louviere (1984). This will be introduced in the next section.

Furthermore, it is highly interesting from a marketing perspective to identify certain segments who are most likely to use the night train for their travel. To answer this question, a latent class choice model is also estimated.

Summarising, this paper contributes to the scientific literature by being the first to study the Willingness to Use night trains, as an alternative for flying, for European long-distance travel. This will be studied with a mode choice experiment using Stated-Preference data. Additionally, using a rating experiment the determinants of the night trains' 'perceived comfort' will be studied. Combining both models using an adapted form of the HII approach, allows for getting further insights such as the Willingness to Pay. This approach is applied and model results reported on data collected from a sample consisting of 804 (mostly) regular train travellers recruited in the Netherlands. At last, insights into segments that are most likely to use the night train is provided by the application of a latent class choice model.

The remainder of this paper is structured as follows. First, the methodology is explained in more detail. This is followed by a presentation and discussion of model estimation results. At first, for the comfort rating experiment, consisting of a regression model. Next, results for the estimated Panel Mixed Logit mode choice model are discussed and segments identified using a latent class choice model. This will be followed by scenario analysis in the model application. Finally, conclusions are drawn, as well as implications for society, limitations of the study and possibilities for future research are discussed.

2. Methodology

As mentioned in the Introduction, the applied main approach involves first determining the 'perceived comfort' rating of various combinations of night train attributes. It is assumed people trade-off this comfort attribute with other attributes, such as trip cost and trip time, in a mode choice experiment. This approach is inspired by the Hierarchical Information Integration theory, which is briefly discussed next.

2.1. *Hierarchical Information Integration theory*

The Hierarchical Information Integration theory was first introduced by Louviere (1984). It is meant as an approach for studying decisions in which many attributes might play a role. It assumes that decision-makers first group together attributes in sets, forming constructs. These constructs are each individually evaluated by the decision-makers. After that, these individual impressions for the decision-constructs are reviewed together and used to evaluate the alternatives in the choice set, resulting in a preference for one of them. The 'perceived comfort' level can be one of those decision constructs.

The traditional approach involves designing two different experiments, a sub-experiment and a bridging experiment. In the sub-experiment, it is explored how the attributes defining the target variable are traded-off. The bridging experiment is used to explore to what extent the decision-constructs itself trade-off against each other.

Variations of this traditional HII approach exist. Bos et al. (2004); Molin and van Gelder (2008) both adapt the bridging experiment in this HII framework. That approach directly included the decision-construct evaluation into the main choice experiment, showing it next to other attributes such as travel cost and travel time. The two experiments can be linked together if the same scale is used for attributes values. A similar approach is also applied in this study, although there is only one decision construct: 'perceived comfort'. This means the term bridging experiment is strictly speaking not correct, therefore it will be referred to as the mode choice experiment. The 'perceived comfort' is the dependent variable in the comfort rating experiment and one of the independent variables in the mode choice experiment. These experiments will be introduced next.

2.2. *Comfort rating experiment*

The comfort rating experiment aims to explore to which extent comfort determinants influence the comfort rating of a night train. To do so, potential passengers will be asked how they would rate their 'perceived comfort' for several night train configurations. This comfort rating may be influenced by several determinants.

However, currently, there is little to no knowledge about what determines the comfort level of a night train. Studying the current night train service levels offered by Austrian operator ÖBB, provided several attributes. Besides, focus groups were organised. During these focus groups, respondents were asked about what determines their comfort level for the night train. Combining these results, it is decided to use six different attributes:

1. Accommodation type: this attribute reflects the type of accommodation aboard the night train. Mirroring the service levels by ÖBB attribute levels will be Sleeper, Couchette and Seat.

2. Number of people in the compartment: was often mentioned by people taking part in the focus groups. The included attributes will be taken as 2, 4 or 6 people.

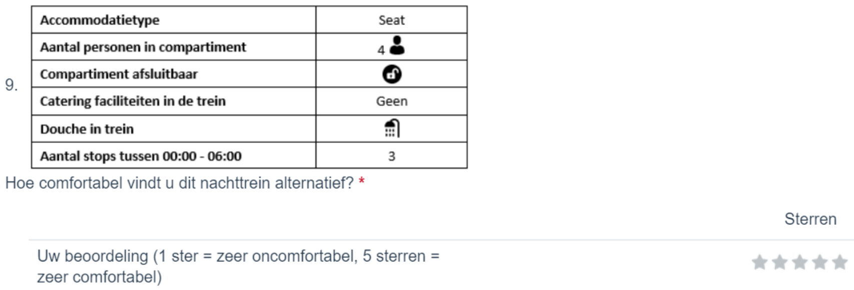
3. Possibility to lock compartment: relates to the possibility to physically lock the compartment for other travellers. Only the traveller and fellow travellers in the compartment can enter. This is varied between yes and no.

4. Catering facilities. This indicates possibilities for getting food or drinks on the train. Three different service levels are distinguished: none, kiosk and a restaurant car. In a kiosk, travellers can buy some snacks/sandwiches and other light meals. In a restaurant car, people have a place to sit down and have a more extensive diner.

5. Possibility to shower: this is varied between yes and no.

6. Number of stops during the night: this reflects the number of stops the night train makes between 00:00 and 06:00. Stopping, travellers boarding and alighting might disrupt sleep and therefore impact comfort levels. This is varied between 0, 3 and 6 stops.

The comfort rating experiment was designed using Ngene software. A fractional factorial design was constructed. This resulted in 36 different profiles. Because this would lead to respondent fatigue, it was chosen to block the design into 6 blocks. Therefore, each respondent was faced with 6 different night train compositions to rate. An example of a comfort rating question is illustrated in Figure 1.


9.

Accommodatietype	Seat
Aantal personen in compartiment	4
Compartiment afsluitbaar	
Catering faciliteiten in de trein	Geen
Douche in trein	
Aantal stops tussen 00:00 - 06:00	3

Hoe comfortabel vindt u dit nachttrein alternatief? *

Sterren

Uw beoordeling (1 ster = zeer oncomfortabel, 5 sterren = zeer comfortabel)

Figure 1: Example of a comfort rating question (in Dutch).

2.3. Mode choice experiment

The goal of the mode choice experiment is to examine how this 'perceived comfort' level is traded off against other attributes, such as travel cost or travel time in mode choice. This set-up is visualised in Figure 2. This study will look into trips of about 12-14 hours, using conventional train technology, this will be illustrated as connections to Amsterdam - Vienna or Milan. As the main goal of this research is to explore what the Willingness to Use the night train is, as an alternative for flying, it is decided to focus on these two modes. Therefore, the included modes in the mode choice experiment are Night Train (NT), Morning Plane (MP) and Evening Plane (EP). While both NT and MP alternatives arrive early in the morning, the EP alternative arrives the day before, which means a hotel stay is needed. Literature research was performed to identify often used attributes in mode choice. As no previous study on this topic has been conducted, participants of focus groups were asked what factors they would take into account when making a decision. This revealed no previously unknown attributes specifically for the night train that needed to be included. The resulting framework can be seen in Figure 3. The following list presents the chosen attributes and the included levels. The EP alternative acts as a base alternative and has fixed attribute levels.

1. Travel time night train: this attribute is based on the average distance (by rail) between Amsterdam and Vienna/Milan, and the average speed of the night train service. This could be 80, 90 or 100 km/h. Resulting in attribute levels of 11:45, 13:00 and 14:15 hours.

2. Travel time plane: reflects the total travel time from arrival at the airport until arrival at the destinations city centre. The attribute levels are 05:00 and 05:30 hours. The base alternative is fixed to 05:00 hours.

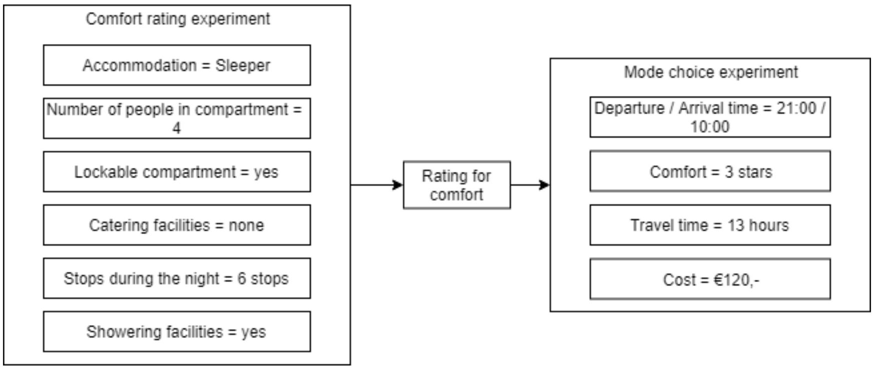


Figure 2: Conceptual overview of applied approach, showing both the comfort rating experiment and the main mode choice experiment. Note that in the figure only one alternative is shown

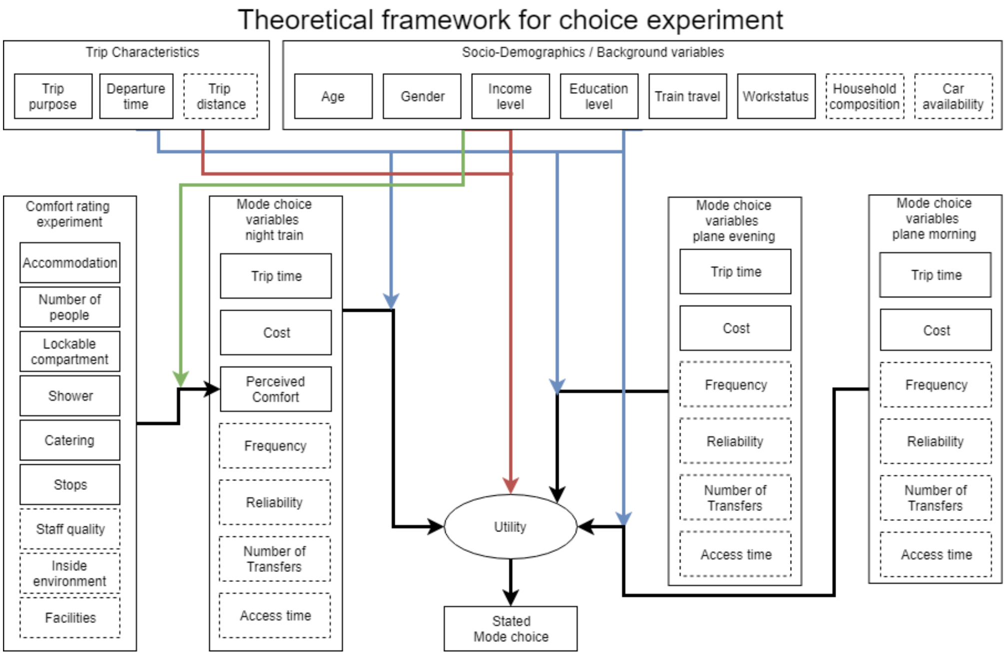


Figure 3: Graphic overview of the theoretic framework. Square boxes: observable variables, oval boxes: unobservable variables, dotted boxes: variables used in other studies, but excluded. Black lines: main effects, green line: socio-demographics/background variables and interactions on comfort rating, red lines: socio-demographics/background on mode choice, blue lines: interactions on mode choice.

3. Travel cost night train: the one-way trip cost for a night train ticket. Attribute levels are chosen to cover a wider range of prices: €40, €80, €120 and €160.

4. Travel cost plane: the one-way cost for the aeroplane alternative. This includes the airport-city centre transfer. Attribute levels: €60, €110, €160. For the Evening Plane alternative the middle price level is chosen and a hotel night has to be added. For the hotel, the average of €120 is taken (PWC, 2016, 2018). This results in a fixed price of €230.

5. Perceived comfort: this is the comfort level of the night train. This is now an independent variable. The attribute levels are chosen as 1, 3 or 5 stars. This reflects a very uncomfortable to a very comfortable comfort level.

The attribute 'perceived comfort' in the mode choice experiment is an independent variable. It is explained to respondents that it is 'their comfort' level. By doing so, they use of the previously completed comfort rating experiment as a reference frame.

Additionally, a survey context was provided to limit respondents from making own assumptions. Based on previous answers to the introductory questions of the survey, respondents were told to be travelling for leisure or business and whether they are travelling alone or with someone else. Also, they were told that they can be on-time on the departure location, and can ignore the access leg of the journey. It was also stated that respondents are travelling with hand-luggage only.

Including the arrival/departure time in a model is shown to improve the model (Bhat, 1998; de Jong et al., 2003). Therefore, it is chosen to include this in the context. To explore the arrival time effect on the Willingness to Use the night train, respondents are asked the same questions in two different settings. One with the desired arrival time at 08:30 and another where this is 10:30.

In total, an experiment with 36 choice sets was designed. Similarly to the comfort rating experiment, this was blocked in 6 blocks. Therefore, for each context-setting respondents had to answer 6 questions, multiplied by 2 arrival time contexts, results in a total of 12 questions. Figure 4 provides an example of a mode choice question.

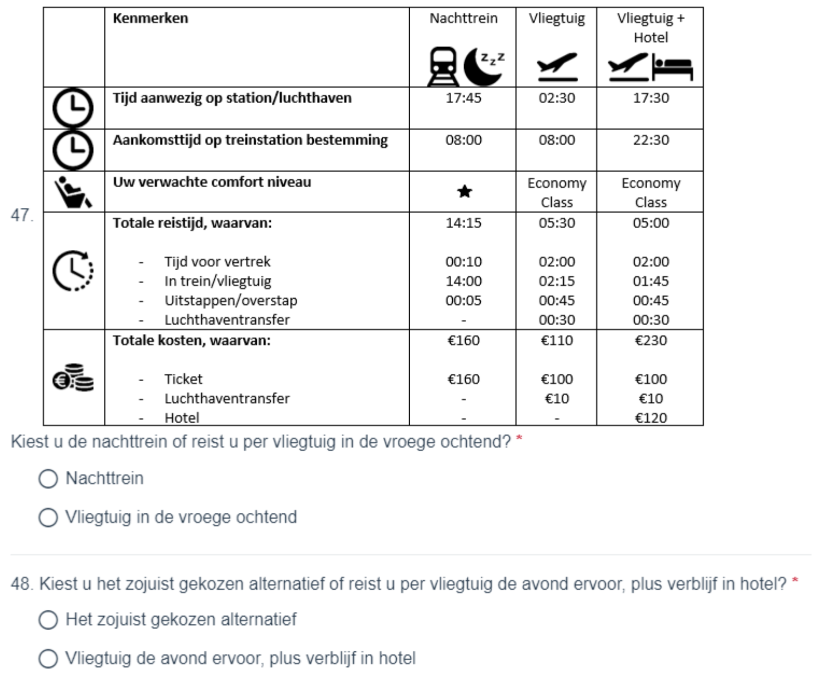
2.4. Background variables

After completion of the choice experiments, respondents were asked some questions regarding socio-demographics and other background variables. This to improve model fit and possibly identify target markets for the night train. Data that was collected provided information on gender, age, education level, employment status, travel purpose, as well as information on their frequency of international train and aeroplane travel.

2.5. Data collection and sample characteristics

The intended population was defined as 'Dutch people who had travelled outside Benelux-countries in 2018', this was chosen so respondents would be familiar with travelling abroad and could imagine making such as trip. Data was collected in three ways: by distributing the survey link to members of the Dutch Railways (NS) customer panel, by handing out flyers at Schiphol Airport and by distributing the link on several social media platforms. This resulted in a total of 804 collected answers.

Table 1 provides more information on the collected background variables for the sample. It is shown that the sample consists of a high percentage of people who are highly educated. Additional variables are included in the model to check if a bias is significant. Furthermore, a



47.

	Kenmerken	Nachttrein	Vliegtuig	Vliegtuig + Hotel
	Tijd aanwezig op station/luchthaven	17:45	02:30	17:30
	Aankomsttijd op treinstation bestemming	08:00	08:00	22:30
	Uw verwachte comfort niveau	★	Economy Class	Economy Class
	Totale reistijd, waarvan:	14:15	05:30	05:00
	- Tijd voor vertrek	00:10	02:00	02:00
	- In trein/vliegtuig	14:00	02:15	01:45
	- Uitstappen/overstap	00:05	00:45	00:45
	- Luchthaventransfer	-	00:30	00:30
	Totale kosten, waarvan:	€160	€110	€230
	- Ticket	€160	€100	€100
	- Luchthaventransfer	-	€10	€10
	- Hotel	-	-	€120

Kiest u de nachttrein of reist u per vliegtuig in de vroege ochtend? *

◯ Nachttrein

◯ Vliegtuig in de vroege ochtend

48. Kiest u het zojuist gekozen alternatief of reist u per vliegtuig de avond ervoor, plus verblijf in hotel? *

◯ Het zojuist gekozen alternatief

◯ Vliegtuig de avond ervoor, plus verblijf in hotel

Figure 4: Example of a mode choice question

Table 1: Distribution of background variables

Background variable	**Category**	**Percentage**	**Background variable**	**Category**	**Percentage**
Gender	Female	43.9%	Train travel	Never	0.2%
	Male	56.1%		<1 day a year	1.0%
Education level	Preschool	0%		1 - 5 days a year	4.7%
	VMBO, HAVO, VWO, MBO1	2.9%		6 - 11 days a year	17.7%
	HAVO, VWO, MBO2-4	16.4%		1 - 3 days a month	28.7%
	HBO, WO Bachelor	39.8%		1-3 days a week	25.5%
	HBO, WO Master or higher	40.9%		4 or more days a week	22.1%
Age	0-19	1.1%	Employment status	Student/Scholar	11.9%
	20-39	24.6%		Parttime worker	18.9%
	40-64	43.2%		Fulltime worker	38.4%
	60-79	30.3%		Retired	28.6%
	80+	0.7%		Jobless	2.2%
Disposable income	<10.000	11.9%	Disposable income		
	10.000 - 20.000	9.6%		40.000 - 50.000	12.7%
	20.000 - 30.000	18.1%		50.000 - 100.000	19.7%
	30.000 - 40.000	24.7%		>100.000	3.3%

lot of the respondents are regular train travellers. This is expected as by distributing the survey in the NS panel, loyal train users were approached. Additionally, by social media students were reached, who can travel by train for free in the Netherlands. Therefore, one must conclude that the sample representativity for Dutch people travelling abroad is questionable. The sample should be considered to be a convenience sample. However, this is not considered to be a (big) problem for the study, as it could be considered that regular train users comprise a primary market for the night train.

2.6. Model estimation

This subsection discusses the followed model estimation procedure. Two models were estimated, one for the comfort rating experiment and another for the mode choice experiment. As stated, in the comfort rating experiment the dependent variable is the 'perceived comfort'. Respondents were asked to rate their comfort on a 5-star rating scale. One could argue that this measurement scale is ordinal, which would mean an ordinal logit regression model should be estimated. However, it is chosen to make a simplification. In reality, people could also award scores on the full range between 1 and 5 stars, meaning the variable can also be interpreted continuously. That also provides the possibility for interpolation between the levels not measured. For these reasons, it is chosen to use a linear model for the comfort rating experiment.

In addition to the main attributes that were varied in the choice experiment, it is explored if including social-demographics and other background variables improve the model fit. This is indeed the case. The final model fit is $R^2 = 0.263$. For estimation, some of the attributes were effects coded, IBM SPSS 25.0 was used for estimation. The resulting linear formula for predicting the comfort rating of a night train configuration is:

$$
\begin{aligned}
CR = C \; &+ \; (\beta_{Sleeper} + \beta_{SleeperAge}) \cdot Sleeper \; + \; (\beta_{Couchette} + \beta_{CouchetteAge}) \cdot Couchette \\
&+ \; (\beta_{NoP} + \beta_{EduNoP}) \cdot NoP + \; \beta_{Lk} \cdot Lk \; + \; \beta_{Restaurant} \cdot Restaurant \; + \; \beta_{Kiosk} \cdot Kiosk \\
&+ \; \beta_{FT} \cdot FT \; + \; \beta_{Age} \cdot Age + \; \beta_{Sh} \cdot Sh + \; \beta_{NoS} \cdot NoS \; + \; \beta_{HI} \cdot HI
\end{aligned}
$$

In this formula: CR is the predicted comfort rating, C = regression constant, NoP = number of people in compartment, Lk = possibility to lock the compartment, Sh = possibility to shower, NoS = number of stops, HI = high income, FT = frequent train traveller. The estimated coefficients and t-values are shown in Table 2.

For the mode choice experiment, 3 different alternatives were included: night train (NT), morning plane (MP) and evening plane (EP). The EP is the base alternative, its utility is fixed to 0. Several logit models were estimated. As with the comfort rating experiment, also socio-demographic and background variables were included. Backward elimination was used to remove insignificant parameters. Table 3 shows the parameter estimations for a base MNL model as well as the final Panel Mixed Logit model. Models were estimated with PandasBiogeme (Bierlaire, 2018). The Panel Mixed Logit model converged with 200 draws from a normal distribution. The resulting equations:

$$
\begin{aligned}
U_{EP} = ASC_{NT} \; &+ \; \beta_{Com} \cdot Com + \beta_{Com\_Q} \cdot Com^2 + \beta_{TTTrain} \cdot TTTrain \\
&+ (\beta_{TCTrain} + \beta_{CP}) \cdot TCTrain + \beta_{AT\_Train} \cdot AT + \beta_{HE\_Train} \cdot HE \\
&+ \beta_{HI} \cdot HI + \beta_{Pp\_train} \cdot Pp + \beta_{S} \cdot S + \beta_{Age} \cdot Age + \sigma_{Morning} + \epsilon_{NT}
\end{aligned}
$$

$$
\begin{aligned}
U_{MP} = ASC_{MP} \; &+ \; \beta_{TTPlane} \cdot TTPlane + \beta_{TCPlane} \cdot TCPlane + \beta_{AT\_MP} \cdot AT \\
&+ \beta_{Gender} \cdot Gender + \beta_{HE\_Plane} \cdot HE + \beta_{Pp\_plane} \cdot Pp + \beta_{FT\_Plane} \cdot FT + \sigma_{Morning} + \epsilon_{MP}
\end{aligned}
$$

$$
U_{EP} = ASC_{EP} + \epsilon_{EP} = 0 + \epsilon_{EP}
$$

In this formulas, Com = comfort rating, CP = interaction with travel purpose and travel cost for the night train, AT = arrival time context, HE = highly educated, HI = high income, Pp = travel purpose, S = student, FT = frequent train traveller.

Note that the comfort rating attribute is modelled as a linear + quadratic component. Reflecting that with increasing comfort levels the utility of upgrading to a higher level drops (decreasing marginal utility). Furthermore, it was found that both the NT and MP alternatives share unobserved factors, resulting in a nest. This is modelled using an additional error component $\sigma_{Morning}$ in the Panel Mixed Logit model.

2.7. Model combination

Both models can be combined to calculate the Willingness to Pay (WtP) values for an increase in one of the comfort attributes, given an initial night train comfort rating. To do so, a 'base WtP' is calculated by dividing the marginal utility contribution of the comfort level by the travel cost parameter. Next, the WtP for an increase in the comfort attributes can be calculated by multiplying the (absolute) linear regression coefficients of the comfort rating experiment with the 'base WtP' number. Note that some of these comfort attributes were effects coded, so in that case, the utility contribution is twice the coefficient (the difference between -1 and 1 levels).

2.8. Latent class choice model

To get insights into the segments of the population that are most willing to use the night train, a latent class choice model was estimated. This model was estimated using LatentGold Choice software. The model consisting of 7 classes performed best. It has a BIC score of 9941.05 points and $R^2 = 0.625$, compared to a 1 class model with a BIC score of 14741.90 and $R^2 = 0.185$. Due to a lack of flexibility in the used program, the interpretation or simulation of LC model is quite burdensome. This is why the latent class model was used solely to identify segments in the dataset.

3. Results

This section presents and discusses the results of model estimations. The results of the comfort rating experiment are presented first, followed by the results of the mode choice experiment. It is shown how to models can be combined to derive WtP values for an improvement in comfort rating attributes and user groups most willing to use the night train are identified.

3.1. Comfort rating experiment

Table 2 presents the parameter estimations for the regression model. The dependent variable is the comfort rating, which is predicted by 6 different comfort attributes and several background variables. The estimations reflect the extent of the change in comfort rating when the attribute value is altered. Main observations:

- All main parameters have the expected sign. Meaning that attributes that are expected to contribute negatively to the comfort level, indeed have a negative sign and vice verse.
- Largest (absolute) parameter is for the 'Sleeper' parameter. The difference between 'Couchette' and 'Sleeper' is almost half a rating point. 'Seat' accommodation results in a penalty of almost a full rating point.
- The 'Number of people' has a strong effect on the comfort rating. It has an effect between -0.506 (2 people) and -1.518 (6 people). Reflecting that people have a strong dislike for sharing the accommodation.
- This 'Number of people' effect is higher if the person is highly educated (interaction), lower if one is not.

Table 2: Estimates for main attributes as well as included background variables.

Parameter	Estimate	T-value	Parameter	Estimate	T-value
Constant	4.086	61.321	High Income	-0.110	-6.553
Sleeper	0.722	11.083	Frequent train traveller	-0.034	-2.104
Couchette	0.285	4.276	Age	-0.006	-2.845
Number of People	-0.253	-26.966	Sleeper * Age	-0.003	-5.506
Possibility to lock compartment	0.163	10.962	Couchette * Age	-0.003	-2.456
Restaurant car	0.148	6.810	Education * Number of People	-0.012	-2.934
Kiosk	0.116	5.532			
Possibility to shower	0.127	8.374			
Number of stops	-0.070	-11.560			

- With increasing age people give, on average, a lower rating to the night train (it reduces the constant). Furthermore, age has a negative effect on both accommodation type parameters. Note that the difference between accommodation types 'Sleeper' and 'Couchette' stays equal. However, the marginal utility difference between 'Seat' and 'Couchette' declines with age.
- The difference between catering parameters is rather small. Meaning that people associate utility with the possibility to get food or drinks on the train, but there is little difference between 'kiosk' and a 'restaurant car'.
- The background variable 'Frequent train traveller' appears to have a counter-intuitive sign. It was expected that people who travel regularly by train, award the night train a higher score (because in daily life the train is their mode of choice). However, the estimation is negative, meaning frequent train travellers give a lower score. Pinpointing a reason for this effect is not possible as that data was not collected.
- Combining all this information leads to the conclusion that a basic night train (no amenities) with 2 person accommodation has a higher comfort rating than a night train with all kinds of luxury facilities, but with shared 6 person accommodation.

Table 3 presents the estimations for two logit models. One is a base MNL model with only main attributes, the other is extended Panel Mixed Logit model including several background variables. Main observations:

- All main parameters have the expected sign.
- Both alternative specific constants are positive, meaning that compared to the base alternate both MP and NT alternatives are more preferred.
- The comfort attribute has a positive linear component and negative quadratic component, reflecting a decreasing marginal utility with higher comfort levels.
- Note that the travel time and travel cost parameters for both modes are scaled. However, for the time the scaling differs between modes. Therefore these cannot be compared 1 to 1. Travel time for the train is scaled by a factor 10, for the plane a factor 5. Travel cost is scaled by a factor 100.
- Taking this into account, it is noted that a minute onboard the night train causes less disutility than onboard an aeroplane.
- The travel cost parameter for the night train interacts with travel purpose (CostPurpose). Meaning that if travelling for business, the cost parameter is higher (less negative).
- With increasing age, people associate a positive utility with the night train (compared to both aeroplane modes).

Table 3: Estimations for base MNL and Panel Mixed Logit model.

Model:	Base MNL		Panel ML				
Parameter:	Estimation	T-value	Estimation	T-value	Parameter	Estimation	T-value
ASC_{Train}	1.68	5.66	2.06	5	β_{Gender_Plane}	0.121	4.08
ASC_{Plane}	1.85	3.39	3.1	4.8	$\beta_{HighEdu\_Plane}$	-0.874	-5.07
β_{Com}	0.518	7.49	0.648	8.09	$\beta_{HighEdu\_Train}$	-0.578	-3.37
β_{Com_Q}	-0.0294	-2.58	-0.0322	-2.43	$\beta_{HighIncome\_Train}$	-0.117	-3.36
$\beta_{TCPlane}$	-0.972	-15	-1.12	-15.9	$\beta_{Purpose\_Plane}$	-1.11	-6.92
$\beta_{TCTrain}$	-0.774	-15.8	-0.939	-13.7	$\beta_{Purpose\_Train}$	-1.35	-7.81
$\beta_{TTPlane}$	-1.72	-1.67	-1.27	-2.26	$\beta_{Student\_Train}$	-0.209	-3.98
$\beta_{TTTrain}$	-0.93	-4.8	-1.21	-4.84	$\beta_{FreqTrain\_Plane}$	-0.729	-4.91
β_{AT}			-0.0776	-2.1	β_{Age_Train}	0.017	8.15
β_{AT_Plane}			0.0869	2.11	$\beta_{CostPurpose}$	0.184	2.7
					$\sigma_{Morning\_std}$	3.32	21.5
Parameters	8		21				
Final loglikelihood	-8955.28		-6815.64				
Rho squared	0.155		0.355				

- Travel purpose has a strong influence on utility for both NT and MP alternatives. In other words, if travelling for business, there is a preference for travelling the day before and staying in a hotel.
- People who are highly educated, have a preference for travelling by EP.
- Frequent train users have a disutility for the morning plane alternative. Resulting in a higher chance of choosing the night train.
- The AT parameters, reflecting the arrival time scenario (which was modelled as an interaction effect on the constant), show that there is a slight preference for the night train when the desired arrival time is 08:30, and a slight disutility for the morning plane.
- Other socio-demographics (such as being a student or the gender) do influence the utility of the modes, but the effect is limited.

3.2. Combination of models

Table 4 shows the WtP values for an increase in a comfort attribute, given an initial night train comfort level (CL). This is shown for both leisure and business travel purposes.

It is shown that the WtP decreases with higher initial comfort levels and also with age for the upgrade from 'Seat' to 'Couchette'. This is logically completely in line with the previously found result (a negative quadratic component and age effect). It shows that people have a low WtP for reducing the number of stops, and also for upgrading the catering facilities, when the possibility to get food/drinks on the train is already provided. A high WtP is achieved for reducing the number of people in the compartment and the possibility to lock the compartment.

3.3. Latent class choice model

A 7-class latent choice model is used to identify segments that are most willing to use the night train. The cumulative probability of belonging to the first three classes is 79.1%. The remainder is split among the other 4 classes. Using the included socio-demographics and background variables each segment can be described as follows:

Table 4: WtP values, for both leisure and business purposes, for an increase in a comfort level attribute, given an comfort level (CL) of the night train.

Purpose				Business [€]				Leisure [€]			
Attribute	Level	Socio-Demographic	Difference	CL 1 ->2	CL 2 ->3	CL 3 ->4	CL 4 ->5	CL 1 ->2	CL 2 ->3	CL 3 ->4	CL 4 ->5
Base WtP				73.03	64.50	55.97	47.44	49.10	43.37	37.63	31.90
Accommodation	Seat -> Couchette	25 years	1.067	77.92	68.82	59.72	50.62	52.39	46.28	40.15	34.04
		45 years	0.887	64.78	57.21	49.65	42.08	43.55	38.47	33.38	28.30
		65 years	0.707	51.63	45.60	39.57	29.75	34.71	30.66	26.60	22.55
		80 years	0.572	41.77	36.89	32.01	27.14	28.09	24.81	21.52	18.25
	Couchette -> Sleeper	25 years	0.437	31.91	28.19	24.46	20.73	21.46	18.95	16.44	13.94
		45 years	0.437	31.91	28.19	24.46	20.73	21.46	18.95	16.44	13.94
		65 years	0.437	31.91	28.19	24.46	20.73	21.46	18.95	16.44	13.94
		80 years	0.437	31.91	28.19	24.46	20.73	21.46	18.95	16.44	13.94
People	2 levels	Highly Educated	0.265	19.35	17.09	14.83	12.57	13.01	11.49	9.97	8.45
		Not Highly Educated	0.241	17.60	15.54	13.49	11.43	11.83	10.45	9.07	7.69
Lock	No ->Yes		0.326	23.81	21.03	18.25	15.47	16.01	14.14	12.27	10.40
Shower	No ->Yes		0.254	18.55	16.38	14.22	12.05	12.47	11.01	9.56	8.10
Catering	None ->Kiosk		0.380	27.75	24.51	21.27	18.03	18.66	16.48	14.30	12.12
	Kiosk -> Restaurant car		0.032	2.34	2.064	1.79	1.52	1.57	1.39	1.20	1.02
Stops	2 levels		0.070	5.11	4.52	3.92	3.32	3.44	3.04	2.63	2.23

1. 'Comfort minded people': the probability of belonging to this class is 32.3%. In this class most attention is put on the comfort level of the night train. The relative importance (RI) of the comfort attributes is more than 60%. People in this group choose for the night train in 52% of the cases, followed by the morning plane (39%). One is most likely to be between 20 and 39 years old (42%), have a low income (36%) and have not travelled by international train before (43%).
2. 'Night train lovers'. One has a probability of belonging to this class of 29.1%. Most attention is paid to the cost of the ticket (RI: 38%) and night train comfort. Due to this fact, the night train is chosen almost every time (98%). People in this class are most likely to be older than 65 (49%), have an average income and hold an university master (41%).
3. 'Morning plane disfavoured'. 17.8% of the sample belongs to this class. Cost of the morning plane is most important (RI: 63%). Because this is a disutility, that means this mode is little used (3%). People in this group are more likely to be male (67%), have a university master (63%), are middle aged (48% between 40 and 64) and have a high income (33% more than 50.000).
4. 'Time sensitive people'. This group represents only 5.5% of the sample. Only attention is paid to travel time and arrival time. The arrival time context variable for the night train is most important (RI: 90%). For the morning plane option, they only consider the travel time and that has low importance (RI: 3%). Therefore, that is the favourite travel option (89%). In this group, there is a high probability one never travelled by international train (69%).
5. 'Evening plane favoured'. Probability of being in this class is 8.8%. People in this class pay a lot of attention to the cost of the morning plane (RI: 76%). The travel time of the night train (RI: 7%) and comfort (RI: 16%) play a smaller role. This results in a preference for the evening plane (96%). There is equal probability (50%) to travel for business or leisure, compared to other classes in which the leisure purpose varied between 61-88%. People are most likely to be male (78%), earn a high income (44%) and hold a university masters degree (56%).

6. ‘Plane users’. 5.4% of the sample belongs to this group. People mostly choose for the aeroplane options (79%). They are cost sensitive towards the morning plane (RI: 51%) and pay attention to night train comfort (RI: 36%). In that sense, this group is similar to class 3. However, the difference is in the covariates. People in this class are young (44% between 20-39), and most likely to earn a low income (35%),
7. ‘Frequent flyers’. The probability of belonging to this class is only 1.2%. People in this class do not opt for the night train (only 2%). Their choice behaviour is only described by the arrival time variables. 42% chooses the evening plane, 56% the morning plane. People in this class are middle aged (62% between 40 and 64), hold a university degree (49%) and are frequent flyers (38% more than 6 times in the past 2 years).

Based on choice behaviour it is possible to identify the classes that are most likely to use the night train. Those are the ‘comfort minded people’, ‘night train lovers’ and ‘morning plane disfavoured’. To a lesser extent ‘plane users’ should also be considered, because they are price sensitive and pay attention to comfort.

4. Application

The estimated mode choice model is applied to investigate the potential modal shares for several scenarios. The investigated scenarios are shown in Table 5. In this table the market share estimates of the night train are also shown for the different travel purposes and if one is a frequent train traveller or not.

Table 5: Different scenarios with the according attribute levels for the mode choice experiment and market shares for the night train (NT)

Scenario:	1. Reference	2. Upgraded trains	3. Fast, luxury trains, upgraded infrastructure	4. Current trains, optimised timetable	5. Luxury trains, optimised timetable	6. Low-cost competitor
Travel time train [hh:mm]	14:15	14:15	11:45	13:00	13:00	14:15
Travel cost train	80	120	160	80	120	40
’Perceived comfort’	3 stars	4 stars	5 stars	3 stars	5 stars	1 star
Travel time plane [hh:mm]	05:00	05:00	05:00	05:00	05:00	05:00
Travel cost plane	110	110	110	110	110	60
Market share NT						
Purpose business	46	48	53	48	56	27
Purpose leisure	63	63	67	66	71	43
Frequent train traveller	70	70	73	72	76	58
Non-frequent train traveller	59	59	63	62	68	39

Table 6: Potential numbers of passengers for a night train service.

Scenario	NT Combined Market share [%]	Passengers yearly [-]	Passenger daily [-]
1. Reference	59.23	44,423	122
4. Current trains, optimised timetable	61.89	46,418	127
5. Luxury trains, optimised timetable	67.67	50,753	139

When the night train is introduced as-is, meaning the line between Dusseldorf and Vienna gets extended, it reflects the reference scenario. Simulations show that upgrading the night train comfort level would have more influence on the night train market share than drastically reducing the travel time by investing in high-speed infrastructure. Positioning the night train as an alternative for low-cost flights with low comfort and low price levels do not benefit the market share, it would drop around 20%. Meanwhile optimising the timetable should be possible.

The timetable can be optimised by reducing the buffer time or the coupling/decoupling time required for splitting the night trains at intermediate stops. Furthermore, often night trains need to share the tracks with freight trains. In that case, better coordination can result in a reduction in travel time. Therefore, scenarios '4. Current trains, optimised timetable' and '5. Luxury trains, optimised timetable' are further considered.

To provide some estimates about how many passengers one could expect for a night train service, some assumptions had to be made. The number of aeroplane passengers between Amsterdam and Milan/Vienna is about 750.000 each. When assuming return trips, this results in 375.000 one-way passengers. To both these destinations around 11 to 14 flights depart daily, of which 2 or 3 in the early morning. This results in a percentage of about 20% if demand is uniformly distributed (and the night train only captures those early morning passengers). Both leisure and business purposes are considered. Data analysis showed that 22% of the respondents travelled for business. This is used to calculate a 'combined potential market share'. Table 6 shows the potential number of night train passengers one could expect for the scenarios.

The number of daily passengers is estimated to be roughly between 120 and 140. The maximum capacity of the current train from Dusseldorf is 246 passengers (vagonWEB, 2019). This would mean occupancy rates would be about 50%. This seems low, however, one should consider that the night train also serves other markets in Germany and the train composition can relatively easily be adapted.

5. Conclusion

5.1. Key findings

This paper aimed to explore what the Willingness to Use night trains, as an alternative for flying, for long-distance European travel is. The approach is based on the assumption that travelling by night train has several advantages and comes with a higher comfort level. To investigate this, two experiments were developed, a comfort rating experiment and a mode choice experiment. In the first experiment, the 'perceived comfort' level was the dependent variable, it was predicted by a set of night train characteristics (& background variables). During the second experiment, the 'perceived comfort' level was included as an independent variable, this to assess how it is traded-off to the more conventional mode choice attributes such as trip time. Additionally, a latent class choice model was estimated to gain insights on market segments.

It has been shown that the 'perceived comfort' rating is heavily influenced by the number of people in the shared accommodation. A non-stop night train with shared accommodation for 6 people with lockable compartments, showers, a restaurant car etc. would score lower than a basic night train that lacks all these 'luxury facilities', but does offer 2 people shared accommodation. With increasing age, people put less weight on the accommodation type (i.e. seat, couchette or sleeper).

Using focus groups it was determined that people do not take previously unknown attributes into account when considering the night train travel option. Travel time and cost, together with comfort for the night train were proved to be significant explanatory variables for mode choice. It has been shown that the marginal utility contribution decreases with increasing comfort levels.

People value one minute of travel time onboard the night train less than one minute onboard an aeroplane in the early morning. The arrival time does influence the mode choice. When people want to arrive in the early morning (08:30) there is a preference for the night train. Most important background variable is the travel purpose. If travelling for business one has a strong preference for taking the aeroplane the evening before and staying in a hotel.

Application of a latent class choice model revealed that 'comfort-minded people', 'night train lovers' and 'morning plane disfavoured' are the most likely segments to choose for the night train. To a lesser extent 'plane users' is found to be interested into using the night train. The 'comfort-minded people' find the comfort factor of the night train highly important and are mostly young. The 'night train lovers' are people who are older than 65, retired and have a high education level. In the 'morning plane disfavoured' segment, people are middle-aged, highly educated and earn a high income.

When the night train is introduced as-is, the combined market share for both leisure and business travel purposes could be around 60%. With improvements to the rolling stock and the timetable, it is predicted to increase to 67%. Positioning the night train as a low-cost alternative is not a good option, the market share suffers from the low comfort level. Using airline passenger numbers, it is expected daily passengers could be between 120 - 140. One can, therefore, conclude that people do indeed seem willing to use the night train as an alternative for flying for long-distance trips in Europe.

5.2. Limitations

This study considered a set context, i.e. a trip within Europe taking 12-14 hours, with a desired arrival time in the (early) morning. This means that conclusions regarding services that fall outside of this scope cannot be made.

A second limitation is the distribution of the survey through the NS panel. It resulted in a large number of responses. However, data analysis showed it is likely to be a 'convenience sample' as it consists of a substantial amount of regular train users. Variables accounting for being a frequent train traveller or highly educated were significant, confirming a bias in the sample.

As always with any Stated-Preference study, it remains to be seen if people opt for the night train in reality. It might have been hard for people to imagine the full travel experience for a non-existing mode. This, for example, showed in the low Willingness to Pay for a reduction in the number of stops. Their decisions could also be influenced by a single negative experience. Something that is currently ignored.

Social-desirability might have played a role when answering the questions. During the study period, there was a public debate about climate change and 'shame of flying'. At last, for some people the choice set would not offer a 'real choice', i.e. people who are in principle against flying or have a fear of flying would opt for the night train by default, while in reality, they might also use another travel mode.

5.3. Further research

Suggestions for further research include repeating the study with a more representative sample, to reduce the possible effects of bias.

Repeating the study in a country where the night train is already operating would allow combing Revealed-Preference data with Stated-Preference data to calibrate the choice model.

Further research is also needed into the mode choice for night trains outside the defined scope of this study, e.g. a preferred arrival in the afternoon. This would allow for more accurately determining the potential market share.

One could develop more advanced latent class choice models. This would allow to further explore the segments in the dataset and related covariates. As a follow-up, these models can be used to investigate how the specific segments would react to different scenarios.

References

Álvarez, A.G., 2010. Energy Consumption and Emissions of High-Speed Trains. Transportation Research Record: Journal of the Transportation Research Board 2159, 27–35. URL: http://journals.sagepub.com/doi/10.3141/2159-04, doi:10.3141/2159-04.

Bhat, C.R., 1998. Analysis of travel mode and departure time choice for urban shopping trips. Transportation Research Part B: Methodological 32, 361–371. URL: https://www-sciencedirect-com.tudelft.idm.oclc.org/science/article/pii/S0191261598000046, doi:10.1016/S0191-2615(98)00004-6.

Bierlaire, M., 2018. PandasBiogeme: a short introduction. Technical Report. Transport and Mobility Laboratory, ENAC, EPFL. URL: http://transp-or.epfl.ch/documents/technicalReports/Bier18.pdf.

Bos, I.D.M., Van der Heijden, R.E.C.M., Molin, E.J.E., Timmermans, H.J.P., 2004. The Choice of Park and Ride Facilities: An Analysis Using a Context-Dependent Hierarchical Choice Experiment. Environment and Planning A: Economy and Space 36, 1673–1686. URL: http://journals.sagepub.com/doi/10.1068/a36138, doi:10.1068/a36138.

Hensher, D.A., Rose, J.M., 2007. Development of commuter and non-commuter mode choice models for the assessment of new public transport infrastructure projects: A case study. Transportation Research Part A: Policy and Practice doi:10.1016/j.tra.2006.09.006.

de Jong, G., Daly, A., Pieters, M., Vellay, C., Bradley, M., Hofman, F., 2003. A model for time of day and mode choice using error components logit. Transportation Research Part E: Logistics and Transportation Review 39, 245–268. URL: https://www-sciencedirect-com.tudelft.idm.oclc.org/science/article/pii/S1366554502000376, doi:10.1016/S1366-5545(02)00037-6.

Louviere, J.J., 1984. Hierarchical Information Integration: a New Method For the Design and Analysis of Complex Multlattribute Judgment Problems. Advances in Consumer Research 11, 148–155. URL: http://acrwebsite.org/volumes/6233/volumes/v11/NA-11.

Molin, E., van Gelder, M., 2008. Freeway Access to Public Transport. Transportation Research Record: Journal of the Transportation Research Board 2076, 106–113. URL: http://journals.sagepub.com/doi/10.3141/2076-12, doi:10.3141/2076-12.

Morikawa, T., Ben-Akiva, M., Mcfadden, D., 2002. Advances in Econometrics Discrete choice models incorporating revealed preferences and psychometric data. Advances in Econometrics URL: https://doi.org/10.1016/S0731-9053(02)16003-8, doi:10.1016/S0731-9053(02)16003-8.

Paulssen, M., Temme, D., Vij, A., Walker, J.L., 2014. Values, attitudes and travel behavior: a hierarchical latent variable mixed logit model of travel mode choice. Transportation 41, 873–888. URL: http://link.springer.com/10.1007/s11116-013-9504-3, doi:10.1007/s11116-013-9504-3.

PWC, 2016. Staying power, European cities hotel forecast for 2016 and 2017. Technical Report. URL: https://www.pwc.com/gx/en/hospitality-leisure/pdf/european-cities-hotel-forecast-2016-2017.pdf.

PWC, 2018. Best placed to grow? European cities hotel forecast for 2018-2019. Technical Report. URL: https://www.pwc.fr/fr/assets/files/pdf/2018/04/pwc-european-cities-hotel-forecast-for-2018-and-2019.pdf.

Román, C., Espino, R., Martín, J.C., 2010. Analyzing competition between the high speed train and alternative modes. The case of the Madrid-Zaragoza-Barcelona corridor. Journal of Choice Modelling 3, 84–108. URL: https://www.sciencedirect.com/science/article/pii/S1755534513700307, doi:10.1016/S1755-5345(13)70030-7.

vagonWEB, 2019. NightJet 40421 train composition. URL: https://www.vagonweb.cz/razeni/vlak.php?zeme=OeBB&kategorie=NJ&cislo=40421&nazev=&rok=2019.

B

Ngene code

This appendix contains the Ngene code that was used to generate the choice sets for both the comfort rating and main mode choice experiments.

Comfort rating experiment

```
? Design for perceived comfort experiment
design
;alts = alt1 , alt2
;rows = 36
;orth = seq
;block = 6
;model:
U(alt1) = ASC
+ Accom.effects[0|0] * Accom[0,1,2]
+ Bpeople*People[2,4,6]
+ Block*Lock[0,1]
+ Cat.effects[0|0] * Cat[0,1,2]
+ Bbath*Bath[0,1]
+ Bstop*Stops[0,3,6]
$
```

Main mode choice experiment

```
? Design for main choice experiment
design
;alts = train, plane, planeref
;rows = 36
;block = 6
;orth = sim
;model:
U(train) = ASC_train
+ Btt_train * TT_train[11.75,13,14.25]
+ Btc_train * TC_train[40,80,120,160]
+ Bcom.effects[0|0] * Com[0,1,2]
+ Bctt * TT_train * Com /
U(plane) = ASC_plane
+ Btt_plane * TT_plane[4.5,5.0]
+ Btc_plane * TC_plane[60,110,160]
$
```

C

Focus group outcomes

This appendix gives an overview of the different answers given by the respondents that took part in one of the focus groups. In total there were 19 participants. In the following tables, one can find the recorded responses.

What are, for you, advantages and disadvantages of travelling by night train?

Table C.1: Table showing the answers to question about advantages and disadvantages of night trains.

Advantages	#Mentioned	Disadvantages	#Mentioned
Travel while sleeping	11	Long travel time	10
Efficient travel option	3	Limited comfort	10
Save hotel night	4	High cost	7
Arrive in city centre in morning	4	Point-to-point connections	2
Short check-in time	1	Pre- and post transport (compared to car)	1
Easier access	1	Privacy/Safety	3
Catering facilities	1	Buying tickets is complicated	2
More space during journey	2	Sanitary facilities	1
Sustainable travel	5	Delays uncertainty	1
No airsickness	1	Noise of trains	3
Departure times	1	Inflexible departure/arrival times	3
Keep luggage in compartment	1	Handle baggage yourself	1

Which attributes would you consider in your decision-making?

Table C.2: Table showing the answers to question about which attributes are considered.

Attribute:	#Mentioned
Travel time	14
Travel cost	17
Comfort	13
Departure/arrival times	7
Number of transfers	6
Luggage regulations	2
Environment	4
Privacy	1
Easiness of booking	1
Purpose	3
Alone or group travel	2
Transport at destination	2

Which factors do you consider to determine ’comfort’ and ’convenience’ of a night train?

Table C.3: Table showing the answers to question about what determines comfort and convenience of a night train.

Comfort:	#Mentioned	Convenience:	#Mentioned:
Accommodation type	15	Luggage (and safety of luggage)	8
Privacy	8	Wait times	2
Catering	7	Safety check	5
Wifi	1	Access distance	10
Vibrations	2	Booking/Comparison website	11
Noise	7	Staff	1
Lighting	6	Ability to get meals	1
Climate control	4	Flexibility (travel times/tickets)	1
Staff	3	Transfers	2
Number of stops	2		
Space to move	5		
Bathroom facilities	9		

D

Ordinal Regression

To check if the assumption that the 'perceived comfort' variable can be considered to be continuous is acceptable, an ordinal logistic regression is performed. The same procedure as during linear regression is applied for estimating the model: first only main effects were added, followed by all socio-demographics/background variables and interactions. If parameters turned out to be insignificant, they were removed by backward elimination. This resulted in the 'Final model' shown in Table D.1.

Table D.1: Parameter estimates, t and significance values from ordinal logistic regression

Model:	Main effects			Background & Interactions			Final model		
Parameter:	Value	Wald	Sig	Value	Wald	Sig	Value	Wald	Sig
Threshold [Stars = 1]	-3.941	1709.610	<0.001	-4.070	536.361	<0.001	-4.113	1029.296	<0.001
Threshold [Stars = 2]	-2.489	840.122	<0.001	-2.591	230.277	<0.001	-2.636	472.616	<0.001
Threshold [Stars = 3]	-1.057	178.264	<0.001	-1.130	45.692	<0.001	-1.178	102.966	<0.001
Threshold [Stars = 4]	0.740	78.325	<0.001	0.692	16.798	<0.001	0.642	29.549	<0.001
Accom1	0.629	270.821	<0.001	1.167	91.087	<0.001	1.160	99.967	<0.001
Accom2	0.259	46.459	<0.001	0.542	18.923	<0.001	0.505	18.306	<0.001
People	-0.448	689.517	<0.001	-0.426	375.850	<0.001	-0.440	629.244	<0.001
Lock	0.282	114.212	<0.001	0.291	120.271	<0.001	0.291	120.449	<0.001
Catering1	0.263	46.964	<0.001	0.264	47.018	<0.001	0.267	48.085	<0.001
Catering2	0.181	23.398	<0.001	0.189	25.260	<0.001	0.190	25.684	<0.001
Shower	0.225	69.743	<0.001	0.231	72.056	<0.001	0.227	70.871	<0.001
Stops	-0.120	122.848	<0.001	-0.115	78.543	<0.001	-0.122	127.508	<0.001
Gen				0.024	0.780	0.377			
HighIncome				-0.206	7.202	0.007	-0.202	46.722	<0.001
HighEdu				0.151	2.598	0.107			
Student				0.080	2.232	0.135			
Retired				0.005	0.014	0.906			
FreqTrain				-0.051	3.141	0.076			
AGE				-0.003	1.648	0.199	-0.004	5.865	0.015
ComfortAge1				-0.010	22.506	<0.001	-0.010	22.797	<0.001
ComfortAge2				-0.005	5.556	0.018	-0.005	4.918	0.027
IncomePeople				0.006	0.098	0.754			
EduPeople				-0.043	4.787	0.029	-0.022	9.066	0.003
GenderShower				0.026	0.931	0.335			
EduStops				-0.014	1.244	0.265			
IncomeAccom				-0.010	0.061	0.805			
IncomeAccom2				0.045	1.143	0.285			

When considering the outcomes of the ordinal regression and comparing them to the 'final model' from linear regression one can notice that the background variable 'TrainTrav' was insignificant ($p = 0.065$). Furthermore, all other parameters are significant and have the expected sign. One remarkable observation is that the parameter values differ quite a lot between the linear and ordinal regression models, in almost all cases the parameters have a larger value. While this is a difference, the interpretation is less straightforward. This is because of the usage of ordinal *logistic* regression, e.g. with a certain combination of variables the model calculates a score and together with the thresholds provides a probability for a certain rating. The final ordinal regression model does significantly predict the dependent variable over an intercept-only model, $\chi^2(13) = 1573.032, p < .001$. As said, for the reasons provided in Chapter 7, it is chosen to use the linear regression.

E

PandasBiogeme code & results

This appendix provides the used PandasBiogeme coding, furthermore, the estimation results of the different discrete choice models are presented.

E.1. Multinomial Logit model

E.1.1. PandasBiogeme syntax

```
import pandas as pd
import biogeme.database as db
import biogeme.biogeme as bio
import biogeme.models as models
pandas = pd.read_table("Dataset_Timev2.dat")
database = db.Database("Dataset_Timev2",pandas)

from headers import *
#Main estimates
ASC_Train = Beta('ASC_Train',0,-1000,1000,0)
ASC_Plane = Beta('ASC_Plane',0,-1000,1000,0)
ASC_Planeref = Beta('ASC_Planeref',0,-1000,1000,1)
Beta_TTTrain = Beta('Beta_TTTrain',0,-1000,1000,0)
Beta_TTPlane = Beta('Beta_TTPlane',0,-1000,1000,0)
Beta_TCTrain = Beta('Beta_TCTrain',0,-1000,1000,0)
Beta_TCPlane = Beta('Beta_TCPlane',0,-1000,1000,0)
Beta_Com = Beta('Beta_Com',0,-1000,1000,0)
Beta_Com_Q = Beta('Beta_Com_Q',0,-1000,1000,0)
Beta_Time = Beta('Beta_Time',0,-1000,1000,0)
Beta_TimePlane = Beta('Beta_TimePlane',0,-1000,1000,0)

#SocioDemographics
Beta_Gender_Train = Beta('Beta_Gender_Train',0,-1000,1000,0)
Beta_HighEdu_Train = Beta('Beta_HighEdu_Train',0,-1000,1000,0)
Beta_Age_Train = Beta('Beta_Age_Train',0,-1000,1000,0)
Beta_HighIncome_Train = Beta('Beta_HighIncome_Train',0,-1000,1000,0)
Beta_FreqTrain_Train = Beta('Beta_FreqTrain_Train',0,-1000,1000,0)
Beta_Student_Train = Beta('Beta_Student_Train',0,-1000,1000,0)
Beta_Retired_Train = Beta('Beta_Retired_Train',0,-1000,1000,0)
Beta_Purpose_Train = Beta('Beta_Purpose_Train',0,-1000,1000,0)

#Interaction
Beta_TimeComfort = Beta('Beta_TimeComfort',0,-1000,1000,0)
```

```
Beta_CostPurpose = Beta('Beta_CostPurpose',0,-1000,1000,0)
Beta_ComAge = Beta('Beta_ComAge',0,-1000,1000,0)

#Same for plane
Beta_Gender_Plane = Beta('Beta_Gender_Plane',0,-1000,1000,0)
Beta_HighEdu_Plane = Beta('Beta_HighEdu_Plane',0,-1000,1000,0)
Beta_Age_Plane = Beta('Beta_Age_Plane',0,-1000,1000,0)
Beta_HighIncome_Plane = Beta('Beta_HighIncome_Plane',0,-1000,1000,0)
Beta_FreqTrain_Plane = Beta('Beta_FreqTrain_Plane',0,-1000,1000,0)
Beta_Student_Plane = Beta('Beta_Student_Plane',0,-1000,1000,0)
Beta_Retired_Plane = Beta('Beta_Retired_Plane',0,-1000,1000,0)
Beta_Purpose_Plane = Beta('Beta_Purpose_Plane',0,-1000,1000,0)

#Rescale data with factor 10 or 100
TT_Train_Scaled = DefineVariable('TT_Train_Scaled', TT_Train / 10,database)
TC_Train_Scaled = DefineVariable('TC_Train_Scaled', TC_Train / 100,database)
TT_Plane_Scaled = DefineVariable('TT_Plane_Scaled', TT_Plane / 5,database)
TC_Plane_Scaled = DefineVariable('TC_Plane_Scaled', TC_Plane / 100,database)

#Define utility functions
V0 = (ASC_Train
      + Beta_TTTrain * TT_Train_Scaled
      + Beta_TCTrain * TC_Train_Scaled
      + Beta_Com *  StarsModel
      + Beta_Com_Q * (StarsModel * StarsModel)
      + Beta_Time * Time
      + Beta_Gender_Train * Gen
      + Beta_HighEdu_Train * HighEdu
      + Beta_Age_Train * Age
      + Beta_HighIncome_Train * HighIncome
      + Beta_FreqTrain_Train * FreqTrain
      + Beta_Student_Train * Student
      + Beta_Retired_Train * Retired
      + Beta_Purpose_Train * Purpose
      + Beta_TimeComfort * TT_Train * StarsModel
      + Beta_CostPurpose * TC_Train_Scaled * Purpose
      + Beta_ComAge * StarsModel * Age)
V1 = (ASC_Plane
      + Beta_TTPlane * TT_Plane_Scaled
      + Beta_TCPlane * TC_Plane_Scaled
      + Beta_Gender_Plane * Gen
      + Beta_HighEdu_Plane * HighEdu
      + Beta_Age_Plane * Age
      + Beta_HighIncome_Plane * HighIncome
      + Beta_FreqTrain_Plane * FreqTrain
      + Beta_Student_Plane * Student
      + Beta_Retired_Plane * Retired
      + Beta_Purpose_Plane * Purpose
      + Beta_TimePlane * Time)
V2 = ASC_Planeref

# Associate utility functions with the numbering of alternatives
V = {0: V0,
     1: V1,
     2: V2}
```

```
# Associate the availability conditions with the alternatives
AV0 = 1
AV1 = 1
AV2 = 1

av = {0: AV0,
      1: AV1,
      2: AV2}

# The choice model is a standard logit, with availability conditions
logprob = bioLogLogit(V,av,ChoiceTime)
biogeme  = bio.BIOGEME(database,logprob)
biogeme.modelName = "Thesis_MNL_SocioDemo_Scaled_AllEffects1"
results = biogeme.estimate()
print("Results=",results)
```

E.1.2. Estimation results

Table E.1 provides the estimation results for the different parameters. This is provided for the model with all included effects.

Table E.1: Parameter estimations for initial MNL model.

Model:	MNL All		
Parameter:	Value	t	Sig
ASC_Train	-0.177	-0.273	0.785
ASC_Plane	2.03	3.33	<0.001
Com	1.02	5.09	<0.001
Com_Q	-0.0351	-2.99	0.003
Beta_TCPlane	-1.04	-15.5	<0.001
Beta_TCTrain	-0.733	-12.5	<0.001
Beta_TTPlane	-0.853	-1.61	0.108
Beta_TTTrain	-0.511	-1.11	0.266
Beta_Time	-0.0621	-2.29	0.022
Beta_TimePlane	0.0933	2.89	0.004
Beta_Gender_Plane	0.228	6.7	<0.001
Beta_Gender_Train	0.131	4.5	<0.001
Beta_HighEdu_Plane	-0.471	-11.4	<0.001
Beta_HighEdu_Train	-0.19	-5.01	<0.001
Beta_HighIncome_Plane	-0.11	-2.85	0.004
Beta_HighIncome_Train	-0.213	-6.88	<0.001
Beta_Purpose_Plane	-0.431	-10.8	<0.001
Beta_Purpose_Train	-0.693	-10.4	<0.001
Beta_Retired_Plane	0.101	1.9	0.058
Beta_Retired_Train	0.0688	1.57	0.116
Beta_Student_Plane	0.328	5.12	<0.001
Beta_Student_Train	0.133	2.22	0.026
Beta_FreqTrain_Plane	-0.515	-3.13	0.002
Beta_FreqTrain_Train	0.165	1.57	0.117
Beta_Age_Plane	-0.00528	-1.69	0.090
Beta_Age_Train	0.0254	7.1	<0.001
Beta_TimeComfort	-0.0156	-1.11	0.266
Beta_CostPurpose	0.19	3.25	0.001
Beta_ComAge	-0.00472	-5.81	<0.001

E.2. Nested Logit model

E.2.1. PandasBiogeme syntax

```
import pandas as pd
import biogeme.database as db
import biogeme.biogeme as bio
import biogeme.models as models

pandas = pd.read_table("Dataset_Timev2.dat")
database = db.Database("Dataset_Timev2",pandas)

from headers import *

#Main estimates
ASC_Train = Beta('ASC_Train',0,-1000,1000,0)
ASC_Plane = Beta('ASC_Plane',0,-1000,1000,0)
ASC_Planeref = Beta('ASC_Planeref',0,-1000,1000,1)
Beta_TTTrain = Beta('Beta_TTTrain',0,-1000,1000,0)
Beta_TTPlane = Beta('Beta_TTPlane',0,-1000,1000,0)
Beta_TCTrain = Beta('Beta_TCTrain',0,-1000,1000,0)
Beta_TCPlane = Beta('Beta_TCPlane',0,-1000,1000,0)
Beta_Com = Beta('Beta_Com',0,-1000,1000,0)
Beta_Com_Q = Beta('Beta_Com_Q',0,-1000,1000,0)
Beta_Time = Beta('Beta_Time',0,-1000,1000,0)
Beta_TimePlane = Beta('Beta_TimePlane',0,-1000,1000,0)

MUmorning = Beta('MUmorning',1,1.0,10,0) #ADAPT HERE

#SocioDemographics
Beta_Gender_Train = Beta('Beta_Gender_Train',0,-1000,1000,0)
Beta_HighEdu_Train = Beta('Beta_HighEdu_Train',0,-1000,1000,0)
Beta_Age_Train = Beta('Beta_Age_Train',0,-1000,1000,0)
Beta_HighIncome_Train = Beta('Beta_HighIncome_Train',0,-1000,1000,0)
Beta_Student_Train = Beta('Beta_Student_Train',0,-1000,1000,0)
Beta_Purpose_Train = Beta('Beta_Purpose_Train',0,-1000,1000,0)

#Interaction
Beta_CostPurpose = Beta('Beta_CostPurpose',0,-1000,1000,0)

#Same for plane
Beta_Gender_Plane = Beta('Beta_Gender_Plane',0,-1000,1000,0)
Beta_HighEdu_Plane = Beta('Beta_HighEdu_Plane',0,-1000,1000,0)
Beta_HighIncome_Plane = Beta('Beta_HighIncome_Plane',0,-1000,1000,0)
Beta_FreqTrain_Plane = Beta('Beta_FreqTrain_Plane',0,-1000,1000,0)
Beta_Student_Plane = Beta('Beta_Student_Plane',0,-1000,1000,0)
Beta_Purpose_Plane = Beta('Beta_Purpose_Plane',0,-1000,1000,0)

#Rescale data with factor 10 or 100
TT_Train_Scaled = DefineVariable('TT_Train_Scaled', TT_Train / 10,database)
TC_Train_Scaled = DefineVariable('TC_Train_Scaled', TC_Train / 100,database)
TT_Plane_Scaled = DefineVariable('TT_Plane_Scaled', TT_Plane / 5,database)
TC_Plane_Scaled = DefineVariable('TC_Plane_Scaled', TC_Plane / 100,database)

#Define utility functions
V0 = (ASC_Train
      + Beta_TTTrain * TT_Train_Scaled
```

```
        + Beta_TCTrain * TC_Train_Scaled
        + Beta_Com *  StarsModel
        + Beta_Com_Q * (StarsModel * StarsModel)
        + Beta_Time * Time
        + Beta_Gender_Train * Gen
        + Beta_HighEdu_Train * HighEdu
        + Beta_Age_Train * Age
        + Beta_HighIncome_Train * HighIncome
        + Beta_Student_Train * Student
        + Beta_Purpose_Train * Purpose
        + Beta_CostPurpose * TC_Train_Scaled * Purpose)
V1 = (ASC_Plane
        + Beta_TTPlane * TT_Plane_Scaled
        + Beta_TCPlane * TC_Plane_Scaled
        + Beta_Gender_Plane * Gen
        + Beta_HighEdu_Plane * HighEdu
        + Beta_HighIncome_Plane * HighIncome
        + Beta_FreqTrain_Plane * FreqTrain
        + Beta_Student_Plane * Student
        + Beta_Purpose_Plane * Purpose
        + Beta_TimePlane * Time)
V2 = ASC_Planeref

# Associate utility functions with the numbering of alternatives
V = {0: V0,
     1: V1,
     2: V2}

# Associate the availability conditions with the alternatives
AV0 = 1
AV1 = 1
AV2 = 1

av = {0: AV0,
      1: AV1,
      2: AV2}

#Definition of nests: -> ADAPT HERE
# 1: nests parameter
# 2: list of alternatives
morning = MUmorning , [0,1]
evening = 1.0 , [2]
nests = morning, evening

# The choice model is a nested logit, with availability conditions
logprob = models.lognested(V,av,nests,ChoiceTime)
biogeme  = bio.BIOGEME(database,logprob)
biogeme.modelName = "Thesis_Nested_SocioDemo_Scaled1"
results = biogeme.estimate()
print("Results=",results)
```

E.2.2. Estimation results

Table E.2 provides the estimation results for the Nested Logit model. Only the nest with the two alternatives that arrive in the morning was significant. Therefore, it is not needed to apply a Cross-Nested Logit (CNL) model. Only the parameters that are significant in the final MNL model are estimated for NL. In addtion, the travel time for the morning plane parameter is included.

Table E.2: Parameter estimations for NL model.

Model:	NL		
Parameter:	Value	t	Sig
ASC_Train	0.919	3.19	0.001
ASC_Plane	1.65	3.27	0.001
Com	0.491	7.15	<0.001
Com_Q	-0.0283	-2.65	0.008
Beta_TCPlane	-0.9	-11.7	<0.001
Beta_TCTrain	-0.667	-11.5	<0.001
Beta_TTPlane	-0.829	-1.81	0.071
Beta_TTTrain	-0.888	-4.38	<0.001
Beta_ArrivalTime	-0.0547	-2.04	0.041
Beta_ArrivalTimePlane	0.0782	2.52	0.012
Beta_Gender_Plane	0.219	6.76	<0.001
Beta_Gender_Train	0.133	4.68	<0.001
Beta_HighEdu_Plane	-0.442	-10.8	<0.001
Beta_HighEdu_Train	-0.205	-5.45	<0.001
Beta_HighIncome_Plane	-0.134	-3.81	<0.001
Beta_HighIncome_Train	-0.22	-7.36	<0.001
Beta_Purpose_Plane	-0.447	-12.3	<0.001
Beta_Purpose_Train	-0.672	-10.8	<0.001
Beta_Student_Plane	0.331	6.35	<0.001
Beta_Student_Train	0.159	2.98	0.003
Beta_FreqTrain_Plane	-0.57	-4.51	<0.001
Beta_Age_Train	0.0134	8.52	<0.001
Beta_CostPurpose	0.167	3.08	0.002
Mumorning	1.18	15.1	<0.001

E.3. Panel Mixed Logit model

E.3.1. PandasBiogeme syntax

```
import pandas as pd
import biogeme.database as db
import biogeme.biogeme as bio
import biogeme.models as models

pandas = pd.read_table("Dataset_Timev2.dat")
database = db.Database("Dataset_Timev2",pandas)
database.panel("ResponseID")

from headers import *

#Main estimates
ASC_Train = Beta('ASC_Train',0,-1000,1000,0)
ASC_Plane = Beta('ASC_Plane',0,-1000,1000,0)
ASC_Planeref = Beta('ASC_Planeref',0,-1000,1000,1)
Beta_TTTrain = Beta('Beta_TTTrain',0,-1000,1000,0)
Beta_TTPlane = Beta('Beta_TTPlane',0,-1000,1000,0)
Beta_TCTrain = Beta('Beta_TCTrain',0,-1000,1000,0)
Beta_TCPlane = Beta('Beta_TCPlane',0,-1000,1000,0)
Beta_Com = Beta('Beta_Com',0,-1000,1000,0)
Beta_Com_Q = Beta('Beta_Com_Q',0,-1000,1000,0)
Beta_Time = Beta('Beta_Time',0,-1000,1000,0)
Beta_TimePlane = Beta('Beta_TimePlane',0,-1000,1000,0)

#SocioDemographics
Beta_HighEdu_Train = Beta('Beta_HighEdu_Train',0,-1000,1000,0)
Beta_Age_Train = Beta('Beta_Age_Train',0,-1000,1000,0)
Beta_HighIncome_Train = Beta('Beta_HighIncome_Train',0,-1000,1000,0)
Beta_Student_Train = Beta('Beta_Student_Train',0,-1000,1000,0)
Beta_Purpose_Train = Beta('Beta_Purpose_Train',0,-1000,1000,0)

#Interaction
Beta_CostPurpose = Beta('Beta_CostPurpose',0,-1000,1000,0)

#Same for plane
Beta_Gender_Plane = Beta('Beta_Gender_Plane',0,-1000,1000,0)
Beta_HighEdu_Plane = Beta('Beta_HighEdu_Plane',0,-1000,1000,0)
Beta_FreqTrain_Plane = Beta('Beta_FreqTrain_Plane',0,-1000,1000,0)
Beta_Purpose_Plane = Beta('Beta_Purpose_Plane',0,-1000,1000,0)

#Rescale data with factor 10 or 100
TT_Train_Scaled = DefineVariable('TT_Train_Scaled', TT_Train / 10,database)
TC_Train_Scaled = DefineVariable('TC_Train_Scaled', TC_Train / 100,database)
TT_Plane_Scaled = DefineVariable('TT_Plane_Scaled', TT_Plane / 5,database)
TC_Plane_Scaled = DefineVariable('TC_Plane_Scaled', TC_Plane / 100,database)

#Random parameter to model the nest
Morning_mean = Beta('Morning_mean',0,-1000,1000,1)
Morning_std = Beta('Morning_std', 0,-1000,1000,0)

# Define a random parameter, normally distirbuted, designed to be used
# for Monte-Carlo simulation
Sig_Morning = Morning_mean + Morning_std * bioDraws('Sig_Morning','NORMAL')
```

```
#Define utility functions
V0 = (ASC_Train
      + Beta_TTTrain * TT_Train_Scaled
      + Beta_TCTrain * TC_Train_Scaled
      + Beta_Com *  StarsModel
      + Beta_Com_Q * (StarsModel * StarsModel)
      + Beta_Time * Time
      + Beta_HighEdu_Train * HighEdu
      + Beta_Age_Train * Age
      + Beta_HighIncome_Train * HighIncome
      + Beta_Student_Train * Student
      + Beta_Purpose_Train * Purpose
      + Beta_CostPurpose * TC_Train_Scaled * Purpose
      + Sig_Morning)
V1 = (ASC_Plane
      + Beta_TTPlane * TT_Plane_Scaled
      + Beta_TCPlane * TC_Plane_Scaled
      + Beta_Gender_Plane * Gen
      + Beta_HighEdu_Plane * HighEdu
      + Beta_FreqTrain_Plane * FreqTrain
      + Beta_Purpose_Plane * Purpose
      + Beta_TimePlane * Time
      + Sig_Morning)
V2 = ASC_Planeref

# Associate utility functions with the numbering of alternatives
V = {0: V0,
     1: V1,
     2: V2}

# Associate the availability conditions with the alternatives
AV0 = 1
AV1 = 1
AV2 = 1


av = {0: AV0,
      1: AV1,
      2: AV2}

#ChoiceModel is a panel model
obsprob = models.logit(V,av,ChoiceTime)
condprobIndiv = PanelLikelihoodTrajectory(obsprob)
logprob = log(MonteCarlo(condprobIndiv))

biogeme  = bio.BIOGEME(database, logprob, numberOfDraws=200)
biogeme.modelName = "Thesis_panel_socio_scaled200"
results = biogeme.estimate(bootstrap=0)
print("Results=",results)
```

E.3.2. Estimation results

Table E.3 shows the model estimation results for the Panel Mixed Logit model with 100 and 200 draws. When one calculates the standard error, it is noticed that the changes in parameter estimations between both models are within 2 times the standard error. Therefore, it can be concluded that the model has converged. The estimation resulting from the model with 200 draws are used in this research.

Table E.3: Parameter estimations for Panel ML model, for 100 and 200 draws.

Model:	Panel ML 100 draws			Panel ML 200 draws		
Parameter:	Value	t	Sig	Value	t	Sig
ASC_Train	1.9	4.77	<0.001	2.06	5	<0.001
ASC_Plane	2.95	4.62	<0.001	3.1	4.8	<0.001
Com	0.648	8.09	<0.001	0.648	8.09	<0.001
Com_Q	-0.0321	-2.42	0.015	-0.0322	-2.43	0.015
Beta_TCPlane	-1.13	-15.9	<0.001	-1.12	-15.9	<0.001
Beta_TCTrain	-0.94	-13.7	<0.001	-0.939	-13.7	<0.001
Beta_TTPlane	-1.28	-2.27	0.023	-1.27	-2.26	0.024
Beta_TTTrain	-1.22	-4.85	<0.001	-1.21	-4.84	<0.001
Beta_ArrivalTime	-0.0777	-2.1	0.036	-0.0776	-2.1	0.036
Beta_ArrivalTimePlane	0.0868	2.1	0.035	0.0869	2.11	0.035
Beta_Gender_Plane	0.12	4.07	<0.001	0.121	4.08	<0.001
Beta_HighEdu_Plane	-0.795	-5.29	<0.001	-0.874	-5.07	<0.001
Beta_HighEdu_Train	-0.499	-3.33	<0.001	-0.578	-3.37	<0.001
Beta_HighIncome_Train	-0.114	-3.26	0.001	-0.117	-3.36	<0.001
Beta_Purpose_Plane	-1.12	-7.74	<0.001	-1.11	-6.92	<0.001
Beta_Purpose_Train	-1.35	-8.59	<0.001	-1.35	-7.81	<0.001
Beta_Student_Train	-0.209	-3.96	<0.001	-0.209	-3.98	<0.001
Beta_FreqTrain_Plane	-0.726	-4.9	<0.001	-0.729	-4.91	<0.001
Beta_Age_Train	0.017	8.15	<0.001	0.017	8.15	<0.001
Beta_CostPurpose	0.185	2.71	0.007	0.184	2.7	0.007
Sigma_morning_std	3.37	21.9	<0.001	3.32	21.5	<0.001

E.4. Simulation model

E.4.1. PandasBiogeme syntax

```
import pandas as pd
import biogeme.database as db
import biogeme.biogeme as bio
import biogeme.models as models
import biogeme.results as res
import sys
pandas = pd.read_table("Dataset_Time.dat")
database = db.Database("Dataset_Time",pandas)
#database.panel("ResponseID") -> Simulation is not compatible with panel effect

from headers import *

#Main estimates
ASC_Train = Beta('ASC_Train',0,-1000,1000,0)
ASC_Plane = Beta('ASC_Plane',0,-1000,1000,0)
ASC_Planeref = Beta('ASC_Planeref',0,-1000,1000,1)
Beta_TTTrain = Beta('Beta_TTTrain',0,-1000,1000,0)
Beta_TTPlane = Beta('Beta_TTPlane',0,-1000,1000,0)
Beta_TCTrain = Beta('Beta_TCTrain',0,-1000,1000,0)
Beta_TCPlane = Beta('Beta_TCPlane',0,-1000,1000,0)
Beta_Com = Beta('Beta_Com',0,-1000,1000,0)
Beta_Com_Q = Beta('Beta_Com_Q',0,-1000,1000,0)
Beta_Time = Beta('Beta_Time',0,-1000,1000,0)
Beta_TimePlane = Beta('Beta_TimePlane',0,-1000,1000,0)

#SocioDemographics
Beta_HighEdu_Train = Beta('Beta_HighEdu_Train',0,-1000,1000,0)
Beta_Age_Train = Beta('Beta_Age_Train',0,-1000,1000,0)
Beta_HighIncome_Train = Beta('Beta_HighIncome_Train',0,-1000,1000,0)
Beta_Student_Train = Beta('Beta_Student_Train',0,-1000,1000,0)
Beta_Purpose_Train = Beta('Beta_Purpose_Train',0,-1000,1000,0)

#Interaction
Beta_CostPurpose = Beta('Beta_CostPurpose',0,-1000,1000,0)

#Same for plane
Beta_Gender_Plane = Beta('Beta_Gender_Plane',0,-1000,1000,0)
Beta_HighEdu_Plane = Beta('Beta_HighEdu_Plane',0,-1000,1000,0)
Beta_FreqTrain_Plane = Beta('Beta_FreqTrain_Plane',0,-1000,1000,0)
Beta_Purpose_Plane = Beta('Beta_Purpose_Plane',0,-1000,1000,0)

#Rescale data with factor 5, 10 or 100
TT_Train_Scaled = DefineVariable('TT_Train_Scaled', TT_Train / 10,database)
TC_Train_Scaled = DefineVariable('TC_Train_Scaled', TC_Train / 100,database)
TT_Plane_Scaled = DefineVariable('TT_Plane_Scaled', TT_Plane / 5,database)
TC_Plane_Scaled = DefineVariable('TC_Plane_Scaled', TC_Plane / 100,database)

#Random parameter to model the nest
Morning_mean = Beta('Morning_mean',0,-1000,1000,1)
Morning_std = Beta('Morning_std', 0,-1000,1000,0)

# Define a random parameter, normally distirbuted, designed to be used
# for Monte-Carlo simulation
```

```
Sig_Morning = Morning_mean + Morning_std * bioDraws('Sig_Morning','NORMAL')

#Set the scenario parameters
TT_Train_Scaled_sim =  1.425   #1.175 - 1.425
TC_Train_Scaled_sim =  0.4   #.40-1.60
StarsModel_sim = 1 #1-5
TT_Plane_Scaled_sim = 1.0   #1 or 1.1
TC_Plane_Scaled_sim = 0.6    #.60-1.60
Purpose_sim = 1 #1 or -1

#Define utility functions
V0 = (ASC_Train
      + Beta_TTTrain * TT_Train_Scaled_sim
      + Beta_TCTrain * TC_Train_Scaled_sim
      + Beta_Com *  StarsModel_sim
      + Beta_Com_Q * (StarsModel_sim * StarsModel_sim)
      + Beta_Time * Time
      + Beta_HighEdu_Train * HighEdu
      + Beta_Age_Train * Age
      + Beta_HighIncome_Train * HighIncome
      + Beta_Student_Train * Student
      + Beta_Purpose_Train * Purpose_sim
      + Beta_CostPurpose * TC_Train_Scaled_sim * Purpose_sim
      + Sig_Morning)
V1 = (ASC_Plane
      + Beta_TTPlane * TT_Plane_Scaled_sim
      + Beta_TCPlane * TC_Plane_Scaled_sim
      + Beta_Gender_Plane * Gen
      + Beta_HighEdu_Plane * HighEdu
      + Beta_FreqTrain_Plane * FreqTrain
      + Beta_Purpose_Plane * Purpose_sim
      + Beta_TimePlane * Time
      + Sig_Morning)
V2 = ASC_Planeref

# Associate utility functions with the numbering of alternatives
V = {0: V0,
     1: V1,
     2: V2}

# Associate the availability conditions with the alternatives
AV0 = 1
AV1 = 1
AV2 = 1


av = {0: AV0,
      1: AV1,
      2: AV2}

prob_nt = models.logit(V,av,0)
prob_mp = models.logit(V,av,1)
prob_ep = models.logit(V,av,2)

logprob_nt = MonteCarlo(prob_nt)
logprob_mp = MonteCarlo(prob_mp)
```

```
logprob_ep = MonteCarlo(prob_ep)

simulate = {'Prob.␣NT': logprob_nt,
            'Prob.␣MP': logprob_mp,
            'Prob.␣EP': logprob_ep}

biogeme = bio.BIOGEME(database, simulate, numberOfDraws=100)
biogeme.modelName = "Thesis_ML_Simulation"

betas = biogeme.freeBetaNames
results = res.bioResults(pickleFile='Thesis_panel_socio_scaled200.pickle')

betaValues = results.getBetaValues()

simulatedValues = biogeme.simulate(betaValues)

b = results.getBetasForSensitivityAnalysis(betas, size=100)
left, right = biogeme.confidenceIntervals(b,0.9)

simulatedValues['ProbNT2real'] = 1.0 * simulatedValues['Prob.␣NT']
left['ProbNT2real'] = left['Prob.␣NT']
right['ProbNT2real'] = right['Prob.␣NT']

probability_NT = simulatedValues['Prob.␣NT'].mean()
probability_NT_left = left['Prob.␣NT'].mean()
probability_NT_right = right['Prob.␣NT'].mean()

print(probability_NT)
print(probability_NT_left)
print(probability_NT_right)

simulatedValues['ProbMP2real'] = 1.0 * simulatedValues['Prob.␣MP']
left['ProbMP2real'] = left['Prob.␣MP']
right['ProbMP2real'] = right['Prob.␣MP']

probability_MP = simulatedValues['Prob.␣MP'].mean()
probability_MP_left = left['Prob.␣MP'].mean()
probability_MP_right = right['Prob.␣MP'].mean()

print(probability_MP)
print(probability_MP_left)
print(probability_MP_right)

simulatedValues['ProbEP2real'] = 1.0 * simulatedValues['Prob.␣EP']
left['ProbEP2real'] = left['Prob.␣EP']
right['ProbEP2real'] = right['Prob.␣EP']

probability_EP = simulatedValues['Prob.␣EP'].mean()
probability_EP_left = left['Prob.␣EP'].mean()
probability_EP_right = right['Prob.␣EP'].mean()

print(probability_EP)
print(probability_EP_left)
print(probability_EP_right)
```

E.4.2. Simulation results

Simulations were performed to do a sensitivity analysis and to predict the potential market shares of the night train for different scenarios. Tables E.4 to E.7, show the numerical results from the scenario calculations, including the 90% confidence intervals. This is shown for the different purposes as well as if one is a frequent train traveller or not.

Table E.4: Probabilities (Prob.) of choosing the night train (NT), morning plane (MP) and evening plane (EP) alternatives, for the different scenarios for the business trip purpose. The 90% Confidence Intervals (CI) are also shown.

Scenario	Prob. NT [%]	90% CI NT [%]		Prob. MP [%]	90% CI MP [%]		Prob. EP [%]	90% CI EP [%]	
1. Reference	45.54	39.52	51.41	15.86	11.64	20.56	38.6	33.57	44.13
2. Upgraded trains	47.59	41.6	53.23	14.76	11.13	18.9	37.65	32.92	42.87
3. Fast, luxury trains, upgraded infrastructure	53.42	46.41	59.96	11.74	8.57	16.81	34.84	29.79	39.98
4. Current trains, optimised timetable	48.05	42.47	54.37	14.47	10.61	18.26	37.48	32.7	42.62
5. Luxury trains, optimised timetable	55.77	49.64	61.95	10.61	7.49	14.4	33.62	28.35	38.74
6. Low-cost competitor	26.94	21.3	33.09	32.14	25.45	38.98	40.92	34.21	45.81

Table E.5: Probabilities (Prob.) of choosing the night train (NT), morning plane (MP) and evening plane (EP) alternatives, for the different scenarios for the leisure trip purpose. The 90% Confidence Intervals (CI) are also shown.

Scenario	Prob. NT [%]	90% CI NT [%]		Prob. MP [%]	90% CI MP [%]		Prob. EP [%]	90% CI EP [%]	
1. Reference	63.34	57.87	68.12	18.51	14.91	23.09	18.15	15.41	21.08
2. Upgraded trains	62.83	57.64	67.21	18.84	15.41	22.99	18.32	15.87	21.21
3. Fast, luxury trains, upgraded infrastructure	66.52	61.47	71.15	16.3	12.97	20.16	17.18	14.54	19.95
4. Current trains, optimised timetable	65.94	60.76	70.67	16.68	13.37	20.91	17.37	14.75	20.18
5. Luxury trains, optimised timetable	71.28	66.52	75.04	13.13	10.42	16.69	15.59	13.34	18.25
6. Low-cost competitor	43.06	36.98	49.17	37.37	31.9	43.12	19.57	16.87	22.33

Table E.6: Probabilities (Prob.) of choosing the night train (NT), morning plane (MP) and evening plane (EP) alternatives, for the different scenarios for a frequent train traveller. The 90% Confidence Intervals (CI) are also shown.

Scenario	Prob. NT [%]	90% CI NT [%]		Prob. MP [%]	90% CI MP [%]		Prob. EP [%]	90% CI EP [%]	
1. Reference	70.25	62.03	75.03	5.28	1.8	13.41	24.47	21.06	27.95
2. Upgraded trains	70.45	62.4	75.38	5.29	2.06	12.87	24.25	20.85	27.79
3. Fast, luxury trains, upgraded infrastructure	73.25	65.12	77.82	4.37	1.99	11.97	22.38	19.01	25.77
4. Current trains, optimised timetable	72.01	63.70	76.37	4.66	1.83	12.58	23.33	20.12	26.84
5. Luxury trains, optimised timetable	75.86	67.57	79.81	3.45	1.12	11.42	20.69	17.35	23.96
6. Low-cost competitor	58.03	45.64	67.42	13.63	3.98	28.28	28.34	24.25	32.33

Table E.7: Probabilities (Prob.) of choosing the night train (NT), morning plane (MP) and evening plane (EP) alternatives, for the different scenarios for a non-frequent train traveller. The 90% Confidence Intervals (CI) are also shown.

Scenario	Prob. NT [%]	90% CI NT [%]		Prob. MP [%]	90% CI MP [%]		Prob. EP [%]	90% CI EP [%]	
1. Reference	59.04	53.24	64.31	18.15	14.36	22.72	22.81	19.26	26.35
2. Upgraded trains	59.13	53.65	64.19	18.18	14.51	22.77	22.69	19.42	25.95
3. Fast, luxury trains, upgraded infrastructure	63.27	57.55	68.01	15.5	12.31	19.65	21.24	18.09	24.76
4. Current trains, optimised timetable	61.66	56.52	66.49	16.4	12.87	20.32	21.94	19.00	25.25
5. Luxury trains, optimised timetable	67.51	62.73	71.84	12.73	9.84	16.17	19.75	16.83	22.99
6. Low-cost competitor	39.01	33.25	44.55	36.59	30.95	42.33	24.41	21.5	27.89

F

Latent class choice model

This appendix provides the output files of the latent class choice model from the LatentGold Choice software. Using these tables, the different market segments were identified.

	LL	BIC(LL)	Npar	Class.Err.	$R^2(0)$
1-Class Choice	-7344.9678	14741.897	8	0.0000	0.1844
2-Class Choice	-6119.3972	12466.128	35	0.0324	0.3767
3-Class Choice	-5147.6472	10698.000	62	0.0272	0.5146
4-Class Choice	-4858.2265	10294.531	89	0.0289	0.5657
5-Class Choice	-4601.4423	9956.3353	116	0.0399	0.6129
6-Class Choice	-4492.6851	9914.1931	143	0.0396	0.6333
7-Class Choice	-4389.1156	9882.4264	170	0.0353	0.6515
8-Class Choice	-4314.1983	9907.9638	197	0.0582	0.6543
7-Class Choice	-4506.1116	9941.0462	143	0.0352	0.6251

Figure F.1: Overview of model fit, the lowest 7-class model is the 'final' model.

	Class1	Class2	Class3	Class4	Class5	Class6	Class7	Overall	Observed	StdResid	UniResid
Class Size	0.3229	0.2906	0.1773	0.0549	0.0880	0.0543	0.0120				
Set Average (n=7944)											
Choice 1	0.5180	0.9557	0.5338	0.0644	0.0370	0.2117	0.0236	0.5586	0.5600	0.1766	0.1309
2	0.3945	0.0291	0.0277	0.8885	0.0022	0.3688	0.5582	0.2162	0.2171	0.1733	
3	0.0875	0.0152	0.4385	0.0470	0.9608	0.4195	0.4182	0.2252	0.2228	-0.4480	

Figure F.2: Overview of mode choices for each class

	Class1	Class2	Class3	Class4	Class5	Class6	Class7
Maximum							
Time_train	0.2182	0.7685	0.4780	0.6233	1.1604	0.4720	0.0000
Cost_train	2.7948	1.0786	0.0000	0.0000	0.0000	0.0000	0.0000
Comfort_train	6.1316	1.8981	3.6773	0.0000	2.8381	2.0980	0.0000
Time_plane	0.2806	0.2806	0.2806	0.2806	0.2806	0.2806	0.2806
Cost_plane	1.3420	2.4657	7.5005	0.0000	13.2246	2.9617	0.0000
Arrival_plane	0.6272	0.0000	0.0000	0.0000	0.0000	0.0000	9.3068
Arrival_train	0.0000	0.0000	0.0000	8.5925	0.0000	0.0000	8.5925
ComfortQ_train	3.3662	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Relative							
Time_train	0.0148	0.1184	0.0400	0.0656	0.0663	0.0812	0.0000
Cost_train	0.1893	0.1662	0.0000	0.0000	0.0000	0.0000	0.0000
Comfort_train	0.4154	0.2924	0.3081	0.0000	0.1621	0.3610	0.0000
Time_plane	0.0190	0.0432	0.0235	0.0295	0.0160	0.0483	0.0154
Cost_plane	0.0909	0.3798	0.6284	0.0000	0.7555	0.5096	0.0000
Arrival_plane	0.0425	0.0000	0.0000	0.0000	0.0000	0.0000	0.5119
Arrival_train	0.0000	0.0000	0.0000	0.9048	0.0000	0.0000	0.4726
ComfortQ_train	0.2281	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Figure F.3: Table showing the maximum and relative (lower half) importance of the different attributes.

Model for Choices																				
	Class1		Class2		Class3		Class4		Class5		Class6		Class7		Overall					
R^2	0.3170		0.0222		0.3216		0.0460		0.0404		0.0212		0.6535		0.5760					
$R^2(0)$	0.4305		0.8767		0.4708		0.7351		0.8929		0.0622		0.7316		0.6251					
Attributes	Class1	z-value	Class2	z-value	Class3	z-value	Class4	z-value	Class5	z-value	Class6	z-value	Class7	z-value	Wald	p-value	Wald(=)	p-value	Mean	Std.Dev.
Time_train																				
	0.8729	4.8513	3.0739	9.3666	-1.9122	-13.8631	-2.4931	-4.0528	-4.6414	-6.7028	-1.8881	-5.6182	0.0000	.	464.7783	3.2e-97	464.7783	3.2e-97	0.1883	2.4567
Cost_train																				
	-2.3290	-17.0224	-0.8988	-3.3167	0.0000	.	0.0000	.	0.0000	.	0.0000	.	0.0000	.	299.6497	8.5e-66	299.6497	8.5e-66	-1.0133	0.9796
Comfort_train																				
	1.5329	9.6772	0.4745	5.2814	0.9193	16.9222	0.0000	.	0.7095	3.5790	0.5245	4.9067	0.0000	.	491.8388	4.6e-104	491.8388	4.6e-104	0.8868	0.4968
Time_plane																				
	2.8058	26.0404	2.8058	26.0404	2.8058	26.0404	2.8058	26.0404	2.8058	26.0404	2.8058	26.0404	2.8058	26.0404	678.1009	1.7e-149	0.0000	.	2.8058	0.0000
Cost_plane																				
	-1.3420	-12.9714	-2.4657	-7.9045	-7.5005	-18.9098	0.0000	.	-13.2246	-8.3020	-2.9617	-14.0940	0.0000	.	499.0004	1.3e-105	499.0004	1.3e-105	-3.8046	3.7036
Arrival_plane																				
	0.3136	6.1498	0.0000	.	0.0000	.	0.0000	.	0.0000	.	0.0000	.	4.6534	11.2538	163.0347	4.0e-36	37.8203	7.8e-10	0.1570	0.5164
Arrival_train																				
	0.0000	.	0.0000	.	0.0000	.	4.2962	5.5322	0.0000	.	0.0000	.	4.2962	5.5322	30.6056	3.2e-8	30.6056	3.2e-8	0.2872	1.0730
ComfortQ_train																				
	-0.1403	-5.4734	0.0000	.	0.0000	.	0.0000	.	0.0000	.	0.0000	.	0.0000	.	29.9577	4.4e-8	29.9577	4.4e-8	-0.0453	0.0656
Model for Classes																				
Intercept	Class1	z-value	Class2	z-value	Class3	z-value	Class4	z-value	Class5	z-value	Class6	z-value	Class7	z-value	Wald	p-value				
	4.0724	2.4547	0.6497	0.3800	-1.4816	-0.3122	0.2478	0.0645	-0.1466	-0.0472	-1.0024	-0.2194	-2.3393	-0.3966	19.5336	0.0034				
Covariates	Class1	z-value	Class2	z-value	Class3	z-value	Class4	z-value	Class5	z-value	Class6	z-value	Class7	z-value	Wald	p-value				
Travel_exp_train																				
	0.1504	1.5417	0.3723	3.7943	0.3160	3.0983	-0.4332	-2.0853	0.0902	0.7542	-0.1053	-0.6135	-0.3903	-1.0246	22.0600	0.0012				
Travel_exp_plane																				
	-0.0063	-0.0826	-0.5232	-5.9638	0.0309	0.3651	-0.0720	-0.5229	0.0725	0.7088	0.1179	0.8683	0.3802	1.5994	40.1709	4.2e-7				
Gender																				
0	-0.1706	-1.5728	-0.0091	-0.0794	0.1250	0.9799	-0.1278	-0.6809	0.3018	1.7862	-0.1217	-0.6016	0.0025	0.0070	8.0267	0.24				
1	0.1706	1.5728	0.0091	0.0794	-0.1250	-0.9799	0.1278	0.6809	-0.3018	-1.7862	0.1217	0.6016	-0.0025	-0.0070						
Education																				
1	1.3536	0.3386	1.1943	0.2977	1.9722	0.3793	2.9858	0.5714	-3.8129	-0.2800	-1.9966	-0.1425	-1.6964	-0.1207	35.5980	0.91				
2	0.0865	0.0320	0.0013	0.0005	1.2985	0.3055	-4.2582	-0.4013	1.9728	0.5280	2.1036	0.4211	-1.2045	-0.1099						
3	0.0878	0.0441	-0.1876	-0.0934	-0.0230	-0.0059	1.3849	0.3535	0.0689	0.0208	-3.9247	-0.4110	2.5937	0.5373						
4	2.1583	0.2541	0.7749	0.0909	-3.4652	-0.1315	-2.5695	-0.0975	4.0442	0.4553	-0.8579	-0.0324	-0.0847	-0.0032						
5	-0.1414	-0.0626	0.3250	0.1435	-4.4204	-0.5236	2.2785	0.5632	1.1858	0.3460	3.1572	0.6642	-2.3847	-0.2677						
6	-1.1716	-0.8543	-1.0331	-0.7509	0.4646	0.1314	0.0658	0.0181	-0.0858	-0.0296	1.1495	0.2620	0.6106	0.1341						
7	-1.3466	-0.9666	-0.7970	-0.5666	0.1149	0.0323	-1.1896	-0.3209	0.2385	0.0815	1.1757	0.2673	1.8040	0.3926						
8	-1.4202	-1.0354	-0.9123	-0.6639	0.8856	0.2507	-0.8332	-0.2294	0.2156	0.0745	0.9604	0.2189	1.1043	0.2430						
9	0.3936	0.0817	0.6345	0.1320	3.1728	0.5468	2.1356	0.3624	-3.8271	-0.2312	-1.7671	-0.1047	-0.7423	-0.0439						
Work																				
0	0.6960	0.6012	-0.4550	-0.3801	1.6083	0.4982	0.1491	0.1026	-0.5171	-0.4001	1.3768	1.0766	-2.8580	-0.4814	25.5530	0.38				
1	-0.7556	-0.9657	-0.2044	-0.2587	1.2561	0.4053	-1.0051	-1.1272	-0.3556	-0.4260	-0.7676	-0.8804	1.8322	0.5434						
2	-0.3430	-0.4409	0.1828	0.2322	0.7643	0.2466	-0.1799	-0.2099	-0.4785	-0.5797	-1.2422	-1.4215	1.2965	0.3832						
3	-0.6442	-0.7918	-0.4643	-0.5646	0.6844	0.2201	-0.2818	-0.3028	0.0753	0.0856	-0.6531	-0.6661	1.2838	0.3711						
4	1.0469	0.3609	0.9409	0.3227	-4.3130	-0.3497	1.3177	0.4413	1.2759	0.4294	1.2861	0.4302	-1.5545	-0.1253						
TrainTravel																				
	-0.0983	-1.1506	0.2319	2.3392	0.0657	0.6284	-0.0552	-0.4035	0.0499	0.3893	0.1355	0.7813	-0.3295	-1.1974	10.5396	0.10				
Purpose																				
-1	0.3328	2.5135	0.4250	2.8646	0.1308	0.9159	-0.1600	-0.7240	-0.3333	-2.0535	-0.3302	-1.5105	-0.0652	-0.1586	21.9765	0.0012				
1	-0.3328	-2.5135	-0.4250	-2.8646	-0.1308	-0.9159	0.1600	0.7240	0.3333	2.0535	0.3302	1.5105	0.0652	0.1586						
AgeGroup																				
	-0.2557	-1.1126	0.8833	3.4885	-0.0570	-0.2173	0.3989	1.0043	-0.2317	-0.7326	-0.5530	-1.2913	-0.1848	-0.2553	19.6853	0.0031				
IncomeGroup																				
	-0.0958	-1.0892	-0.2655	-2.9375	0.0841	0.8500	-0.0146	-0.0997	0.1712	1.4572	0.2444	1.3624	-0.1237	-0.4490	15.9701	0.014				

	Class1	Class2	Class3	Class4	Class5	Class6	Class7
Class Size	0.3229	0.2906	0.1773	0.0549	0.0880	0.0543	0.0120
Attributes							
Time_train							
1.175	0.2977	0.2162	0.4154	0.4408	0.5339	0.4143	0.3333
1.3	0.3320	0.3175	0.3271	0.3228	0.2989	0.3272	0.3333
1.425	0.3703	0.4663	0.2575	0.2364	0.1673	0.2584	0.3333
Mean	1.3091	1.3313	1.2803	1.2744	1.2542	1.2805	1.3000
Cost_train							
0.4	0.6210	0.3960	0.2500	0.2500	0.2500	0.2500	0.2500
0.8	0.2446	0.2764	0.2500	0.2500	0.2500	0.2500	0.2500
1.2	0.0964	0.1929	0.2500	0.2500	0.2500	0.2500	0.2500
1.6	0.0380	0.1347	0.2500	0.2500	0.2500	0.2500	0.2500
Mean	0.6205	0.8265	1.0000	1.0000	1.0000	1.0000	1.0000
Comfort_train							
1	0.0021	0.0975	0.0214	0.3333	0.0450	0.0833	0.3333
3	0.0444	0.2519	0.1343	0.3333	0.1860	0.2378	0.3333
5	0.9535	0.6506	0.8444	0.3333	0.7689	0.6789	0.3333
Mean	4.9028	4.1062	4.6460	3.0000	4.4479	4.1912	3.0000
Time_plane							
1	0.4303	0.4303	0.4303	0.4303	0.4303	0.4303	0.4303
1.1	0.5697	0.5697	0.5697	0.5697	0.5697	0.5697	0.5697
Mean	1.0570	1.0570	1.0570	1.0570	1.0570	1.0570	1.0570
Cost_plane							
0.6	0.5642	0.7265	0.9765	0.3333	0.9987	0.7818	0.3333
1.1	0.2884	0.2118	0.0230	0.3333	0.0013	0.1778	0.3333
1.6	0.1474	0.0617	0.0005	0.3333	0.0000	0.0404	0.3333
Mean	0.8916	0.7676	0.6120	1.1000	0.6007	0.7293	1.1000
Arrival_plane							
-1	0.3481	0.5000	0.5000	0.5000	0.5000	0.5000	0.0001
1	0.6519	0.5000	0.5000	0.5000	0.5000	0.5000	0.9999
Mean	0.3037	0.0000	0.0000	0.0000	0.0000	0.0000	0.9998
Arrival_train							
-1	0.5000	0.5000	0.5000	0.0002	0.5000	0.5000	0.0002
1	0.5000	0.5000	0.5000	0.9998	0.5000	0.5000	0.9998
Mean	0.0000	0.0000	0.0000	0.9996	0.0000	0.0000	0.9996
ComfortQ_train							
1	0.7352	0.3333	0.3333	0.3333	0.3333	0.3333	0.3333
9	0.2394	0.3333	0.3333	0.3333	0.3333	0.3333	0.3333
25	0.0254	0.3333	0.3333	0.3333	0.3333	0.3333	0.3333
Mean	3.5243	11.6667	11.6667	11.6667	11.6667	11.6667	11.6667
Covariates							
Travel_exp_train							
1 - 1	0.4333	0.3056	0.2101	0.6899	0.3128	0.4126	0.6154
2 - 2	0.2410	0.2406	0.2482	0.1917	0.1528	0.2264	0.2576
3 - 3	0.1837	0.2639	0.2720	0.0551	0.3263	0.2500	0.0003
4 - 6	0.1420	0.1899	0.2697	0.0633	0.2081	0.1110	0.1268
Mean	1.1413	1.4933	1.7277	0.5201	1.5330	1.1267	0.6385
Travel_exp_plane							
1 - 1	0.1255	0.3407	0.0898	0.1094	0.0687	0.0852	0.0000
2 - 2	0.1496	0.2456	0.1870	0.2483	0.2168	0.0876	0.1257
3 - 3	0.3557	0.2845	0.2723	0.3094	0.2650	0.3486	0.3856
4 - 4	0.2014	0.0820	0.1946	0.2217	0.1539	0.2640	0.1052
5 - 6	0.1678	0.0473	0.2563	0.1111	0.2956	0.2146	0.3835
Mean	2.1645	1.2631	2.4522	2.0051	2.5668	2.5462	2.8747
Gender							
0	0.5104	0.5866	0.6737	0.5517	0.7817	0.5980	0.4889
1	0.4896	0.4134	0.3263	0.4483	0.2183	0.4020	0.5111
Education							
1	0.0324	0.0211	0.0085	0.0828	0.0000	0.0000	0.0000
2	0.0624	0.0292	0.0254	0.0000	0.0515	0.0568	0.0000
3	0.0989	0.0448	0.0104	0.0828	0.0170	0.0000	0.1281
4	0.0093	0.0052	0.0000	0.0000	0.0172	0.0000	0.0000
5	0.0671	0.0733	0.0000	0.2215	0.0512	0.1261	0.0000
6	0.2778	0.3188	0.2386	0.3835	0.2329	0.2427	0.2597
7	0.1597	0.0827	0.0599	0.0281	0.0687	0.2200	0.1259
8	0.2821	0.4053	0.6316	0.1728	0.5615	0.3545	0.4862
9	0.0103	0.0195	0.0256	0.0286	0.0000	0.0000	0.0000
Work							
0	0.2746	0.0320	0.0623	0.0364	0.0344	0.3196	0.0000
1	0.1344	0.1679	0.2567	0.1096	0.1636	0.1763	0.3819
2	0.3825	0.3322	0.4723	0.4120	0.5466	0.3671	0.3637
3	0.1803	0.4418	0.2086	0.4144	0.2383	0.1116	0.2543
4	0.0283	0.0262	0.0000	0.0275	0.0172	0.0254	0.0000
TrainTravel							
1 - 4	0.2389	0.2016	0.1417	0.4667	0.2257	0.1506	0.3823
5 - 5	0.2586	0.3418	0.2815	0.2575	0.2398	0.2383	0.3834
6 - 6	0.2558	0.2919	0.2950	0.0551	0.2493	0.3091	0.1085
7 - 7	0.2467	0.1647	0.2818	0.2208	0.2853	0.3019	0.1257
Mean	4.3799	4.3828	4.6573	3.8663	4.5770	4.7046	3.8497
Purpose							
-1	0.8448	0.8752	0.6529	0.7525	0.5002	0.6060	0.6407
1	0.1552	0.1248	0.3471	0.2475	0.4998	0.3940	0.3593
AgeGroup							
0	0.0326	0.0000	0.0001	0.0000	0.0000	0.0562	0.0000
1	0.4245	0.1311	0.2679	0.1192	0.2310	0.4383	0.1282
2	0.3650	0.3772	0.4814	0.4663	0.4960	0.3674	0.6175
3	0.1733	0.4863	0.2339	0.4144	0.2730	0.1381	0.2544
4	0.0047	0.0054	0.0166	0.0000	0.0000	0.0000	0.0000
Mean	1.6929	2.3661	1.9991	2.2952	2.0419	1.5873	2.1262
IncomeGroup							
1 - 2	0.3620	0.1769	0.0731	0.1197	0.0678	0.3522	0.1259
3 - 3	0.1644	0.2323	0.1538	0.2483	0.1153	0.1209	0.2563
4 - 4	0.2135	0.3001	0.2650	0.2763	0.2050	0.1256	0.2563
5 - 5	0.1012	0.1084	0.1783	0.1627	0.1756	0.1235	0.1259
6 - 8	0.1589	0.1823	0.3297	0.1931	0.4364	0.2778	0.2356
Mean	2.3119	2.8512	3.5672	3.0249	3.8661	2.6494	3.1941