The influence of perceived COVID-19 risk on the modal-split for long-distance travel in Europe

A Hierarchical Information Integration and Stated-Preference study approach

M.M. van Dalen

TU Delft

Royal HaskoningDHV
Enhancing Society Together

The influence of perceived COVID-19 risk on the modal-split for long-distance travel in Europe

A Hierarchical Information Integration and Stated-Preference study approach

by

M.M. van Dalen

to obtain the degree of Master of Science
in **Transport, Infrastructure & Logistics (TIL)**
at the Delft University of Technology,
to be defended publicly on Thursday June 30, 2022 at 10:45 AM.

Student number:	4444965	
Project duration:	October 14, 2021 – June 30, 2022	
Thesis committee:	Dr. O. Cats,	TU Delft, committee chair
	Ir. N. Geržinič,	TU Delft, 1^{st} supervisor
	Prof. dr. ir. C. G. Chorus,	TU Delft, 2^{nd} supervisor
	Ir. B. J. H. F. Donners,	Royal HaskoningDHV, company supervisor

TU Delft

Royal HaskoningDHV
Enhancing Society Together

Preface

This thesis is written as completion to the Master of Science Transport, Infrastructure & Logistics. This research is done as an internship for Royal HaskoningDHV. It studies the effects of perceived COVID-19 (infection) risk, on modal-split for travelling to European long-distance destinations.

I would like to express my gratitude for the support and guidance I have received during the process of writing this graduation thesis. First of all, I would like to thank Royal HaskoningDHV for giving me the opportunity to do this thesis as an internship at the company. Additionally, I would like to thank the Sustainable Mobility team members for helping me with questions during the different phases of the research.

Besides this, I would like to thank all of my committee supervisors. First of all, I want to thank both Nejc Geržinič and Barth Donners for the daily supervision. I would like to thank for being approachable, giving me sharp feedback and to spar with. I am very thankful for the amount of time both have taken for my research. Furthermore, I would like to thank Caspar Chorus as my second supervisor. Besides the official meetings, the two meeting together were very helpful. Moreover, I would like to thank my chair of the committee, Oded Cats. It was very helpful that I could always approach you and the 'catch-up' meetings were enjoyable as well. At last, I want to thank Jet van de Kar (NS international), for the collaboration with the NS panel. Because of the NS panel, this research has a high amount of respondents.

Maurizio van Dalen
Delft, June 2022

Contents

List of Tables

List of Figures

1
Introduction

1.1. COVID-19 and travel behaviour

The new coronavirus SARS-CoV-2 (from now: COVID-19) is an ongoing global pandemic that is considered one of the worst post-World War II pandemics that affects the world, surpassing both MERS and SARS outbreaks (Matiza, 2020). Based on rising infection rates in China and then to over the world, the WHO Emergency Committee declared a global health emergency on January 30, 2020 (Velavan and Meyer, 2020). Countries have taken several measures to reduce the number of infections; some of these measures were restrictions from entering other countries, the closing of restaurants/bars and measures during the trip (like a face mask or negative test). Because of these measures, COVID-19 has a huge impact on travel behaviour (De Vos, 2020). One of the elements of travel behaviour is mode choice. It is determined by a lot of factors that are often interrelated to each other to a smaller or larger extent (De Witte et al., 2013). Often it is the result of a compound choice process that unconsciously or consciously influences daily life and includes objective and subjective determinants (Buehler, 2011). Quite some research is done on the effects of mode choice, mainly focusing on daily (short-distance) travel behaviour (Buehler, 2011; Atasoy et al., 2013; De Witte et al., 2013). However, long-distance travel behaviour is often excluded from the analysis. At the same time, over 50% of the passenger-kilometers travelled come from long-distance travel (Aamaas et al., 2013). Several studies researched the main drive for passengers to choose a certain mode for long-distance travel in Europe. The main finding of these studies was that travel time seems to be the most important factor (van Goeverden, 2009). Furthermore, several variables seem to be important in the decision between modes: travel cost, comfort, reliability, access & egress time and number of interchanges (Román et al., 2010).

As COVID-19 dramatically changed travel behaviour, mode choice is potentially be influenced as well. Several papers researched the effects of COVID-19 infection on mode choice, but the focus was on short-distance travel. Abdullah et al. (2020) found that people tend to use less public transport (hereafter: PT) services and more private cars during the pandemic. There was a shift found to active modes as well. Moreover, sanitisation measures and social distancing characterise the perception of safety and, therefore, the willingness to use PT services (Aaditya and Rahul, 2021). According to de Haas et al. (2020) in the Netherlands, during the first wave, around 80% of the people reduced their activities, with an increase from 6% of the people working at home to 39%. For public transport, there was a drop in usage of 90%. Shamshiripour et al. (2020) noticed the following pattern in Chicago: people tend to shift more to individual and active modes of travel (e.g., walking or cycling) or not travelling at all. According to Wen et al. (2005), other viruses outbreaks like SARS and MERS in Asia showed that people reduced the amount of travel due to internal risk (perceived risk by travelling) and external risk (ban of travelling, orders to stay at home). Another possible reason for the reduction of travel by PT is that people believe that PT is an unhygienic place, with a high probability of infections with viruses (Troko et al., 2011).

At this stage, most of the COVID-19 measures are gone in the Netherlands. Partly this is a consequence of omicron-variant of the COVID-19 virus (Chen, 2022). This variant results in lower hospitalisation rates. The Dutch society accepted to 'live' with COVID-19 because COVID-19 will not fully be gone. Therefore, several scenarios can be possible, for example, that a new variant increases hospitalisation, and therefore again, measures are taken by the government. Or the situation remains the same with COVID-19 being a virus amongst other viruses like influenza.

By definition, a lot of different people are transported at the same time in PT systems; therefore, the virus can relatively easy be transmitted among travellers. To illustrate this, Krishnakumari and Cats (2020) found that, on average, a person interacts with 1200 other people on a single trip in the metro system network of Washington D.C. Now, it seems evident that measures were taken in order to reduce the number of infections and mitigate the public health crisis. As a consequence, PT usage was dramatically reduced. Because of the risk of infection of COVID-19 in PT systems, people might prefer private modes. However, all of the papers mentioned were focusing on daily travel behaviour. The risk effects of COVID-19 on mode choice for long-distance travel are very little researched. Therefore this research investigates if and how risk effects of COVID-19 affect mode choice for long-distance travel.

To study the effects of COVID-19 on mode choice, a new variable is introduced: perceived risk. Perceived risk is the risk people perceive during their trip regarding the probability of getting infected with COVID-19. It is often defined as the perceived likelihood of getting the disease times the perceived severity of the symptoms (Karlsson et al., 2021). For this study, however, perceived risk consists on mode-related and destination-related attributes. Perceived risk due to COVID-19 is complicated to measure, with several elements/attributes that could contribute to this perceived risk. These attributes possibly weigh different for each individual. Therefore, perceived risk is a complex variable, its score is determined by other variables (Molin, 2020). Perceived risk does also have an emotional dimension like worry and fear (Loewenstein et al., 2001). When looking at different aspects of perceived risk, people who perceive COVID-19 as a greater risk tend to engage better in efforts that are preventive, such as social distancing and hand-washing. Potentially, people change their travel behaviour (and thus mode choice) due to these factors as well as other factors of their trip. The outcome of this study could be beneficial in transport planning and policy-making during health crises. Providers of PT services (like airlines or rail operators) can use the information the optimise their services and operations.

1.2. Research questions

Altogether this leads to the following main research question:

To what extent does risk perception of infection with COVID-19 influence mode choice for long-distance trips in Europe?

The following sub-questions are stated in order to answer the main research question:

1. How do the different risk factors influence the perceived risk variable?
2. How do socio-demographic variables influence the perceived risk variable within the rating experiment?
3. How do socio-demographic and travel behaviour variables influence travel cost, travel time, travel comfort and perceived risk in the main choice experiment?
4. How are travel time, travel cost and travel comfort traded off against people's perception of COVID-19 risk?

1.3. Scope

Mode choice for long-distance travel in Europe can be very broad. To ensure the research is doable, the topic is scoped.

Modes that are included

It is important to consider which modes are included in this research. Only conventional high-speed train is taken into account for train options. Aeroplane is taken into account as well, as this is still the most popular mode for distances greater than 400 km (Álvarez, 2010; Reichert and Holz-Rau, 2015). The car is also a popular choice for long-distance travel (Reichert and Holz-Rau, 2015). Therefore this mode is included as well. Long-distance bus services have increased during the past years. However, bus only counts for approximately 2.5%

of market share (Blayac and Bougette, 2017). Due to this low percentage, bus is not taken into account. Moreover, for the sake of simplicity of this research, night trains are not included in the research.

Definition long-distance & distance classes

From literature, it becomes not clear what the definition of 'long-distance' is. Some papers state a distance of at least 160 kilometer (Georggi and Pendyala, 2000), but others do state a distance of at least 240 kilometer (Dargay and Clark, 2012). Furthermore, Miller (2019) recommends excluding daily or weekly travel from long-distance, as this is routine travel, even though it covers the distance for the long-distance definition. According to Givoni (2006), it becomes clear that the maximum distance for travelling by train is approximately 1000 km. The maximum distance for this research is around 1200 kilometres. At this distance, still, all three modes can be competitive alternatives. As an example, the Amsterdam - Barcelona route is around 1200 kilometres. Besides, a certain minimum distance has to be taken into account as well for certain distances. This is since for short distances aeroplane is not an option. The minimum distance is 400 kilometres. This is, for example, the Amsterdam - London and Amsterdam - Paris routes. On these routes, train and plane are heavily competing with each other (Román et al., 2010). Car is a viable and often chosen option as well. In order to approach certain distances, two distance classes is introduced; the 400 - 600 kilometre class and the 800 - 1200 kilometre class. It is chosen to not overlap the classes so that each distance class has different characteristics, and therefore the widest range of possible values is taken into account.

Travel purpose & time span

The research is held in the Netherlands, and therefore the decision-makers are mostly Dutch. The most common types of travellers are leisure, business and 'Visiting Friends or Relatives' (from now: VFR) travellers. Within these categories, there are still a lot of different travel purposes. An important factor that influences mode choice is luggage (Cascetta et al., 2011). Luggage is an important factor that significantly differs for different travel purposes. To test the 'perceived risk' variable, the effects of luggage is excluded.

1.4. Relevance of research

In this section, the societal, scientific and company relevance of this research is discussed.

Societal relevance

Most research does focus on daily travel behaviour, not on long-distance travel behaviour. Beginning more than 2.5 years ago, COVID-19 dramatically changed society. Due to several restrictions, travel behaviour significantly changed (de Haas et al., 2020). Little research has been done on COVID-19 effects on long-distance travel. Again, research that has been done is mainly focusing on daily travel behaviour, not on long-distance travel behaviour. Thus, this research can give insights into the effects of COVID-19 on long-distance travel behaviour. These insights can be used for society so that effective measures can be taken for the future.

Scientific relevance

Little research is done on the effects of COVID-19 on long-distance travel behaviour (see chapter 2). Papers that were found focused on daily travel behaviour. Besides, in general, research on mode choice is (almost) always based on daily travel behaviour. A lot of research is done on the 'shorter' daily trips (Buehler, 2011; Atasoy et al., 2013; De Witte et al., 2013). Several papers investigated COVID-19 effects on mode choice as well, these papers all focus on daily travel behaviour (de Haas et al., 2020; Aaditya and Rahul, 2021; Shaer and Haghshenas, 2021). To get insight into the knowledge of the effects of COVID-19 on the mode choice of long-distance travel, this research is done. Moreover, this research also gives insight on general mode choice for long-distance travel, alongside time valuation, individual characteristics, travel characteristics and interactions with background & socio-demographic variables.

Company relevance

This research is carried out as a graduate internship at Royal HaskoningDHV. The mission statement of Royal HaskoningDHV is 'Enhancing Society Together'. They want to achieve this through their innovations, partnerships and expertise. They combine their knowledge and skills with the strengths of their clients to co-create solutions that improve the lives of people all around the world (HaskoningDHV, n.d.).

Royal HaskoningDHV (from now: RHDHV) is one of the market leaders within the mobility field. An important aspect within their mobility field is sustainable mobility. They identified several steps to get to the point where train travel is a good alternative for short-haul flights. Little research has been done on the effects of COVID-19 and also the other main variables (such as travel time and travel cost) for long-distance travel. This research suits this field, as the focus of this thesis is to get insight into trade-offs on mode choice for long-distance travel within Europe. As there is little knowledge on mode choice for long-distance travel, this research gives insight into RHDHV and which trade-offs are made. This can be used for other rail-related projects.

Moreover, RHDHV is in the Train2EU initiative. This is a non-profit organisation which was formed together with 9292 (Dutch PT app), 'Jonge Veranderaars' (a platform for young professionals in public transport) and Natuur & Milieu. This initiative wants to improve and stimulate the usage of international trains. This thesis can help to achieve this goal as it gives insight into mode choice for long-distance travel.

1.5. Thesis outline

The thesis has the following structure. Chapter 2 discusses relevant literature on COVID-19 travel behaviour impact, the concept of perceived risk and mode choice for long-distance travel. Chapter 3 discusses the used method in a theoretical way. Then in chapter 4, the theoretical framework and survey design are discussed. In chapter 5, the sample is analysed. In chapter 6, the results of the perceived risk rating experiment and the main (mode) choice experiment are discussed. In chapter 7, these results are applied to three cases. In chapter 8, the conclusion to the research question is taken. At last, in chapter 9, the discussion and recommendations are discussed.

2
Literature

This chapter discusses relevant literature for this research. Section 2.1 is about COVID-19 travel behaviour impacts, section 2.2 about the impact of other virus outbreaks on travel behaviour, section 2.3 explains the concept of perceived risk, section 2.4 discusses perceived risk caused by COVID-19, section 2.5 elaborates on the relationship between risk and travel behaviour and at last, in section 2.6 literature on mode choice for long-distance travel is discussed.

2.1. COVID-19 travel behaviour impacts

The COVID-19 pandemic has had a huge impact on travel behaviour. Countries have taken several measures to prevent the spread of the virus, such as stimulating working from home, closing schools, closing bars/restaurants and prohibiting events (De Vos, 2020). People avoid travelling by public transport as people might believe that avoiding contact with other passengers might be difficult. Hu et al. (2020) quantified the transmission risk of COVID-19 on-board trains in China. The average attack rate was between 0 tot 10.3%. At the same time, people could believe that public transport is an unhygienic place full of viruses (Troko et al., 2011). People might travel more by car as this protects them from getting the virus. Therefore public transport operators should focus on improving a safer way of travelling when social distancing is needed. According to De Vos (2020), public transport operators should not reduce their services, even though they depend heavily on the revenue of fares. Governments could support the public transport operators. This study gives implications of social distancing on daily travel patterns; however, it is not based on quantitative research. The study also provided some suggestions for policymakers.

Shamshiripour et al. (2020) did a Stated-Preference survey on travel behaviour and implemented this in the Chicago metropolitan region. The data obtained demonstrate major shifts in several elements of people's travel habits. The main finding was that people shift towards more individual and active (e.g., walking or cycling) modes of transport. Another finding was that working from home seems to work and could also be implemented for near future policy goals towards sustainability. Moreover, people tend to shift from the aeroplane as a mode for leisure or business trips to private modes (especially car). This research did a comprehensive study, but its focus is on Chicago. Europe and the U.S. are quite different in terms of travel behaviour, especially on the long distances, with fewer mode choice possibilities in the U.S (Aditjandra et al., 2009)

De Haas et al. (2020) did a study in the Netherlands on how the 'intelligent lockdown' changed the travel behaviour of people. Around 80% of the respondents reduced activities, 27% expect to work more from home after COVID-19. The share of home-workers increased from 6% to 39%. The number of trips dropped by 55% and the distance travelled by 68%. In particular, public transport services are highly affected, with a 90% drop in usage. This study also confirms the increase in usage of active modes and also that people expect to fly less in the future. This article shows results from a big sample (N=2500), but this article does not focus on long-distance travel. It mainly focuses on daily travel behaviour.

The paper written by van Wee and Witlox (2021) discusses the possible long-term effects of COVID-19 on travel behaviour. This paper makes use of concepts and theories, sociology, psychology and geography. Lasting effects can be expected; therefore, peak effects for car and public transport can be expected. Lamb et al., (2020) noticed that the decision to fly among different users depends on perceived risk of COVID-19, affect, fear and agreeableness. This research looked into the willingness to fly given the COVID-19 with a regression model with a total sample of 632 participants. (Van Der Drift et al.) 2021, did also do a study in the Netherlands and noticed the same as de Haas et al., (2020): cycling showed to be an alternative option for travellers, and PT usage decreased dramatically. Van Der Drift et al., (2012) showed that by high impacts, such as 9/11 and the oil crisis, the long-lasting effects were relatively minor.

2.2. Impact of other virus outbreaks on travel behaviour

Wen et al. (2005) analysed the impacts of SARS on leisure travel in the China domestic market. The authors of this paper found that the outbreak of SARS had reduced the travel associated with different motivations. Both internal (perceived risk by travelling) and external risk (ban of travelling, orders to stay at home) reduced the amount of travel. However, this study did not explore what the effects were post-outbreak.

When focussing on the airline industry, a study from Fenichel et al. (2013) researched travel patterns of air travel by analysing 1.7 million flights records. This was done to study the behavioural response to the influenza pandemic back in 2009. It was estimated that 0.34% of missed flights were because of the pandemic. However, this pandemic was not so widespread as COVID-19. Therefore, it is hard to make conclusions about the effect of COVID-19. Another study by Liu et al. (2011) looked at the effects of SARS on international air travel. For this case, international air travel between China, Hong Kong & Taiwan and the U.S. was examined. This study found that the level of risk by aircraft was perceived differently in every country. However, in all countries, the number of trips made by aircraft reduced. Sobieralski (2020) did a study among several industries and economic variables of the COVID-19 pandemic, with capacity reductions of airlines. It was estimated that around 7% job loss could be expected, with a maximum of 13%. The COVID-19 (but also the SARS outbreak) show that airlines are highly vulnerable. Effects of COVID-19 will probably last for about four years. Focusing on public transport trips, a study done by Kim et al. (2017) analysed PT card data before and after the outbreak of MERS in South Korea. People that live in areas with higher income tend to stay more at home than people in other areas. The fear of exposure to the virus resulted in changes in the travel pattern. People with higher income are mostly less reliable on PT than people with lower income.

2.3. Perceived risk

However, not all people will comparably change habits due to COVID-19. Socio-demographic, attitudinal and psychological factors will possibly influence these habits. Risk is a widely studied topic. It is measured by multiplying probability with impact, which is shown in the following equation:

$$Risk = probability \times impact \tag{2.1}$$

This equation is an objective definition of risk. Risk can be defined as as the subjective evaluation of the risk of a dangerous situation and the severity of the situation (Moen et al., 2004; Moreira, 2008). Based on the evaluation of the situation, risk perception will consequently change an individual's behaviour (Weinstein, 1988). Risk also includes a personal (individual) element. The difference between the objective value of risk and the perceived (personal) one is known as the 'perception gap' (Faganel, 2010). Partly this can be attributed to the awareness of a person. Awareness is multidimensional and opaque. Several studies defined the term 'perceived risk'. Bauer (1960) was one of the first to measure perceived risk. After this, several studies tried to define the concept of perceived risk. Peter and Ryan (1976) mentioned that the difference among individuals in the perception of risk could be explained by the difference in the judgement of risk and the amount of negativity combined with this. This is similar to the equation above; however, it differs from the fact that the impact and probability are assessed by the individual instead of as aggregated terms beforehand. Yates and Stone (1992) conceptualised perceived risk as a mismatch that is multidimensional between the required and obtained outcome of a product or service. Also, importance was added to the calculation.

2.4. Perceived risk caused by COVID-19

Perceived risk does have an emotional dimension like worry and fear (Loewenstein et al., 2001). Perception of risk is thus a form of risk assessment. This is very convenient, as the axiom about expected utility theory in choice experiments uses this as well. Socio-demographic variables partly explain the difference in impact judgement and likelihood judgement. When focussing on COVID-19, the difference in impact judgement can be stated as the severity of COVID-19 on the individual (this means the impact of a COVID-19 infection). The impact of a COVID-19 infection is mostly caused by underlying health condition and age (Dong et al., 2020). Therefore it is expected that younger people have a lower impact on judgement than older people.

Moreover, several papers studied perceived risk of COVID-19 due to personal (individual) characteristics. These characteristics are based on both affective and cognitive risk assessment. Gerhold (2020) studied the risk perception of COVID-19 among German citizens. The data was gathered in March 2020. Risk perception was operationalised on a scale with affective (how serious is the person taking the pandemic) and cognitive (the perceived likelihood of contradiction with the virus) dimensions. Furthermore, the affective dimension also includes the level of worry. Gerhold (2020) measured the effects of several qualitative dimensions for perceived risk due to COVID-19. These dimensions were based on the psychometric paradigm of Slovic (1992). The elements of risk perception were covered in a survey with a Likert scale. Surprisingly, it was found that elderly people perceive the risk of getting infected to be lower than younger people, while at the same time, the consequences of infections are higher for elderly people (Dong et al., 2020). Another interesting observation was that people generally are more worried about other people (like friends and family) than about themselves. A lot of people were worried about the pandemic and the virus, but the fear of infection was relatively low. This might indicate that people accept the risk and, therefore, will not change their behaviour.

Dryhurst et al. (2020) studied risk perception of COVID-19 of persons from all over the world. The survey was held in March and April 2020, just when the pandemic started. The survey measured the risk perception by introducing a risk perception index to study the differences in these perceptions between countries. With psychological and demographic predictors, which were largely based on a study done by one of the authors (Dong et al., 2020). This research was based on risk perception regarding climate change. Of the socio-demographic variables, only gender turned out to be significant. Men did, in general, have a lower risk perception than women. Another study was done by Glöckner et al. (2020) in Germany, which also showed that people who perceived their likelihood of getting infected by COVID-19 as higher did take more measures to prevent infections, like staying at home and the usage of face-mask. The studies above showed that perceived risk of getting COVID-19 has a lot of impact on precautions that people take to prevent it. However, no connection has been made to travel behaviour. Potentially, people will change their mode choice due to factors of their trip (such as spacing and measures that are taken during the trip).

Brown et al. (2020) performed a study about health-related risk and experiences, which was held in the United Kingdom during the most strict measures. They found, in line with Dryhurst et al. (2020), that women perceive risk of COVID-19 higher than men. Moreover, a relation was found between education level and perception of risk. Lower educated people had a higher perceived risk for COVID-19 than higher educated people. Furthermore, Ahorsu et al. (2020) and Taylor et al. (2020) both developed a multidimensional scale for worry, fear and anxiety for COVID-19, with Taylor et al. (2020) having an additional scale for stress. Media usage and fear for COVID-19 turned out to be significant in a study done by Mertens et al. (2020). When looking at different aspects of 'perceived risk', people who perceive COVID-19 as a greater risk tend to engage better in efforts that are preventive, such as social distancing and hand-washing (Karlsson et al., 2021).

2.5. Risk and travel behaviour

Perceived risk of travelling is highly related to the intention of an individual to change the travel plan, for example travel to a certain destination or avoiding a destination (Reisinger and Mavondo, 2005; Pennington-Gray et al., 2011; Schroeder et al., 2013). Moreover, self-efficacy is something that is becoming relevant when an individual perceives risk as severe or likely and therefore will avoid the risk by avoiding/changing the destination or cancelling the trip (Rogers, 1975; Schroeder et al., 2013). So perceived risk will not only influence the decision for destination choice, but also to travel or not (Floyd et al., 2000; Reisinger and Mavondo, 2005; Rittichainuwat and Chakraborty, 2009). Furthermore, media is also an important factor for the relationship between risk perception and intention to travel (Neuburger and Egger, 2021). Travelers are likely to change travel plans when a certain destination is linked with negative events or a higher risk of incidents. This is done in order to avoid a perceived 'unsafe' destination and therefore travelers will seek for a safer alternative (Sönmez and Graefe, 1998). Travelers are likely to avoid destinations with a higher safety risk, such as natural disasters, terrorist attacks or a pandemic (Pizam and Fleischer, 2002; Rittichainuwat and Chakraborty, 2009).

2.6. Mode choice on long-distance travel

Van Goeverden (2009) did a study to explain factors for train use for European long-distance travel (with a focus on leisure travel). The study found several significant background variables, such as the size of the destination city, car ownership, home country and the number of participants in the journey. Significant quality variables were travel cost and travel time (for both train and car) and the number of interchanges. The frequency of the service did not turn out to be significant. The study showed that leisure travellers do not significantly differ from other long-distance travellers. This is because leisure travellers count for a high percentage of all travellers. The knowledge of background variables can be used for estimating future demand for train travel. As car ownership is growing, it could be expected that this has a negative influence on the market share for trains. The rising necessity for transfers, the increased obligation to make seat reservations, and the complexities of fare schemes are all negative aspects. If policymakers try to reduce this, market share could be increased. This study is comprehensive, showing what factors influence the usage of trains on long-distance European destinations. However, the study is already quite outdated (2009). Therefore, the results will not be representative for now as several variables have potentially changed. At the same time, they not studied what the potential effects are of a pandemic such as COVID-19.

Román et al. (2010) analysed the competition on the high-speed rail (HSR) Barcelona-Madrid route. The model specification was based on travel time, access and egress time, reliability, headway & comfort. By using these attributes, they explained the changes in demand for HSR. The study obtained different measures of willingness to pay (WtP) for improved service quality. Most of the time, the WTP is higher for mandatory trips than leisure trips. Moreover, the WTP is higher when the level of comfort is lower. There is also a high WtP for a reduction in delay. They also found that comfort attributes change the perception of time. On the shorter distances, demand is more sensitive to travel time than to access&egress time or price. Policies penalising the alternatives of the car are most effective. Demand for HSR is not sensitive to price and mostly headway. When aeroplanes compete with trains, demand is sensitive to airfares and access&egress time. Combining increasing travel cost of car with decreasing travel time for HSR obtained substantial gains for the train. Most of the market share increase for train is obtained from cars and bus travellers, not from aeroplane. The study is very comprehensive and discusses what factors influence the demand for HSR. At the same time, they stated several policy solutions. They concluded with the fact that due to high costs and low return on rate levels, HSR is not always profitable. At the same time, they note that other elements such as regional development and welfare also have a positive impact.

Dobruszkes et al. (2014) also found that air services are affected by HSR; more air services are offered if the travel time of HSR is longer than by aeroplane. The same picture can be taken for the number of seats offered, and the number of flights offered. Airlines do not apply frequency-based strategies to compete better with HSR. Moreover, they found the same effect as Román et al. (2010): HSR frequency has a very weak impact on air services. Another study done by Clewlow et al. (2014) does also confirm the effect of rail travel times on air traffic in Europe; reduction in travel times of train reduces short-haul traffic in Europe. At the same time, HSR does also reduce the amount of domestic air traffic in European countries. They found something interesting; when population density increases, short-haul air traffic declines when there is a good rail option. So rail may be more competitive in more densely populated areas. However, as low-cost carriers have been introduced to Europe, passenger kilometres travelled by aircraft have only increased.

In a study done by Behrens and Pels (2012), the inter-modal competition on the London-Paris route between HSR and air transport was studied. In contradiction with the studies above, they found that the frequency of HSR is one of the main factors of travel behaviour. The studies above concluded that frequency did not have a high impact on demand for HSR, but this study concludes that the frequency of HSR is an important variable for demand. However, they conclude that travel time and distance are the main factors of travel behaviour as well. Furthermore, they found that business and leisure travellers behave differently; leisure travellers are more heterogeneous for average fares in comparison with business passengers. Thus, inter-modal competition between train and aeroplane depends largely on trip purpose. It turns out that for the London-Paris route, the HSR is a viable option. They even suggested that HSR can be viable with larger differences in travel times due to larger distances of city pairs or lower average speed. Also, having multiple airline options is not a barrier to introducing HSR. At last, they concluded that HSR competes with both main and low-cost carriers. Airlines have difficulties in markets where HSR is dominant, like on the London-Paris route. This suggests that HSR can be a strong mode alternative. Another study done in Spain by Jiménez and Betancor (2012) concluded that the entry of HSR in Spain reduced air operations by around 17%. The total demand for HSR increased between 8% and 35%, depending on the route. The highest increase was found on the Barcelona-Madrid route. However, they did not identify which share was switching from the other modes and which share was newly generated (e.g. induced demand).

Givoni and Dobruszkes (2013) did an extensive ex-post review from several studies of mode substitution and induced demand when HSR is introduced. This study again confirmed that the main factor that explains HSR demand is travel time (and, therefore, average speed). On the HSR Rome-Naples, route travel time was the main factor in choosing HSR. Other factors that were found (in order from most important to least important): comfort, novelty, frequency and fare. For travellers that stick to flying, their main reason was comfort, and also onward connections to other destinations were important. For the car, travel time and comfort were most important. Furthermore travel time to stations as well as the number of transfers required is also important. In Korea, a lot of people do not use the HSR because stations are not easily reachable, and also, the price is an important factor not to use HSR. In Taiwan, stations are often away from city centres. Respondents said that this was one of the main factors in not choosing HSR. Estimates of the value of time show that savings of one minute access&egress time are the same as a two minute saving for in-vehicle time. Often it is seen that travel time is the starting point of research with only a few other factors discussed. This can lead to the effect that travel time masks other effects that are not examined. Usually, information on fares is not available and thus not included in the research. Low-cost airlines are also mentioned. Low-cost carriers might also increase the substitution of rail to air. In Germany, it was found that the entry of low-cost carriers into the German domestic market led to a decrease in demand for rail. The longer the route, the stronger the effect was of decrease in rail. This could be expected as Dobruszkes et al. (2014) confirmed this picture as well. However, when looking at the Paris-London route, HSR travellers are less sensitive to fare, frequencies and total travel time than airline passengers. The difference with Germany could be partly explained by the fact that in Germany for a lot of OD-pairs travel time is more than three hours, while the Paris-London route is less than three hours. Therefore, for airlines on the Paris-London route, it is hard to compete this travel time.

Cascetta et al. (2011) showed that the cross-price elasticity of demand did have a low potential for modal shift from car to rail concerning travel cost and travel time. Car does still offer the most flexibility concerning the schedule of a traveller and also route choice. Luggage is another factor in choosing car over HSR, especially with heavy luggage. However, travel time variability (due to congestion) is a factor why people choose HSR over car, as well as the fact that travel time can be used effectively (e.g., working on the train). At last 'attitudes' of travellers might play an important role. People with a 'green attitude' might be more willing to use HSR; this attitude was also found in the study of Molin et al. (2016).

Another study done by Moeckel et al. (2015) introduced a new nested multinomial model for mode choice sensitive to distance, travel cost, service frequency, transit station availability, number of transfers & parking cost. This was done for car driving alone and sharing rides for 2-4 people. On the transit side, the following choices were available: rail, bus & air. The study was based on travel in North Carolina (U.S.), so mode choice alternatives are mostly different than for Europe. Europe has a more frequent and reliable train network. The study showed that for short distances, car is dominant. The greater the distance, the more dominant the aeroplane becomes. For distances to approximately 550 kilometres car has a share of over 80%. From then on, aeroplane becomes almost dominant. Train only has a 2% share and bus only a 5% share for to a maximum of 450 kilometres. As can be seen in this study, car is very dominant as could be expected from the U.S. Therefore, this cannot represent travel behaviour mode choice in Europe.

3
Methodology

This chapter discusses the methodology that is used to provide an answer to the main research question and sub-questions. For this research, a rating experiment and a Stated-Preference experiment is conducted. A survey is held with a representative sample within the Netherlands. By using this methodology, mode choice between train, plane and car in Europe is examined. Section 3.1 discusses the data collection by using a Stated-Preference survey. Section 3.2 discusses the (modified) Hierarchical Information Integration (HII) approach that is done by using a perceived risk rating experiment. Section 3.3 discusses the main (mode) choice experiment. Section 3.4 discusses the theory used to predict the modal split. A summary can be found in section 3.5.

3.1. Data collection by using a Stated-Preference survey

To collect data, a survey has to be conducted to research the effect of COVID-19 on mode choice between train, plane and car in Europe. There are two general ways of collecting data: Stated-Preference (SP) data or Revealed-Preference (RP) data. A study done by Wardman (1988) compered SP and RP. Both methods have drawbacks and advantages. For RP, only existing alternatives can be taken into account; no new alternatives can be tested. In the case of this project, RP could be partly used to compare the market shares before and during COVID-19. As there is the wish to investigate the risk effects of contradicting COVID-19, additional data is needed, which cannot be obtained with existing RP data. Therefore an SP research is conducted. SP has an additional benefit in that it has total control within the experiment. This includes, alternatives, attributes of the alternatives and values for alternatives (Molin, 2020). At the same time, as the design of the survey/choice experiment can be controlled, correlations can stay low between different attributes. This ensures more valid results. However, there are also drawbacks to Stated-Preference. It is sometimes hard to create sufficient variation in choice situation to ensure that utility functions can be estimated, making the parameters unreliable (Molin, 2020). Parameters are reliable if they have small standard errors. This can be done by choosing a proper experimental design; therefore this research uses both an orthogonal or D-efficient design (Molin, 2020). If the parameters are reliable, they resemble the true parameters and therefore resemble the 'real world'. Another disadvantage of SP is hypothetical bias, which is defined as the disparity between stated and observed behavior. (Brownstone and Small, 2005). This study showed that there were significant differences between willingness to pay values that were derived from SP and RP studies.

Ben-Akiva et al. (2019) and Molin (2020) presented a lot of elements that require attention when a SP experiment is conducted. This must be done in order to get valid results on the assessment of the change in modal split due to the new variable perceived risk. Four elements that are important when conducting an SP choice experiment are mentioned and discussed in how they are used in this research. *Recruitment, sampling and background* are important regarding the sample. This research uses a representative sample. In order to get 95% significance for a population of >100.000, 400 respondents in the sample are needed (SurveyMonkey, 2021). A minimum goal of 400 respondents is therefore used for the survey. Responses are collected by sharing the survey on social media with family and friends and by recruiting people at (train) stations & Amsterdam Schiphol Airport. In collaboration with NS, the survey is with their panel; this ensures a lot of extra respondents. In order to test whether the sample is representative, socio-demographic variables are compared to data from the Statistics Netherlands (CBS). *Familiarity and attribute formatting* are important as well; respondents must be familiarised with the subject being examined. Furthermore, *experimental design* is important; this means that the survey must be precisely constructed regarding the number of alternatives, attributes and attribute levels. As explained above, this is done by using either an orthogonal or D-efficient design. *Calibration & testing* is recommended for the results. This is done by comparing the SP experiment with RP. This is limited as the effects of COVID-19 can almost not be compared to other outbreaks/pandemics.

3.2. Rating experiment of 'perceived COVID-19 risk'

Perceived risk due to COVID-19 is complicated to measure, with several elements/attributes that could contribute to this perceived risk. These attributes possibly weigh different for each individual. Therefore, perceived risk is a complex variable, its score is determined by other variables (Molin, 2020). For this research both *internal (psychological), route, mode specific* and *socio-demographic* attributes probably influence the perceived risk (Molin, 2020). Therefore, the first step is to find these attributes in literature. After these attributes are determined, the 'rating' experiment is done to measure how the attributes influence the score on the complex variable perceived risk. Perceived risk (in rating points) is the dependent variable, and the attributes the independent variables. With this experiment, the perceived risk of travellers can be predicted (Molin, 2020) when using a certain mode for a certain OD-pair. A regression model is used to analyse the rating experiment. This research uses the Hierarchical Information Integration (HII) theory, which was developed by Louviere (1984). This theory is used when decision-makers are confronted with many attributes. Decision-makers categorise these attributes into 'decision constructs'. Decision-makers (respondents) trade-off attributes that belong to such a 'decision-construct' in the first (sub) experiment, the 'rating' experiment. Then in the 'bridging' experiment, decision-makers make a trade-off between the construct evaluations that are done in the 'rating' experiment (Molin, 2020). This research used an altered version of the original HII experiment, with just one 'decision construct,' to determine the perceived risk. Estimating the 'rating' model allows for predicting the perceived risk. Then in the main choice experiment, the perceived risk attribute is shown among the other attributes that are defined. This is not a true 'bridging' experiment as in a conventional 'bridging' experiment more sub experiments for the decision constructs are used; in this research; however, only one. Molin and van Gelder, (2008) showed that this approach is successful.

3.3. Main (mode) choice experiment

After the perceived risk is retrieved from the 'rating' experiment, the main choice experiment is done. In this case, the perceived risk variable is an attribute among the other main attributes regarding mode choice between train, plane and car. The main attributes are reviewed from the literature and are then used in order to construct the survey. Choice sets do consist of three parts: attributes, attribute levels and alternatives. An orthogonal or D-efficient design for constructing choice sets is used. The orthogonal design minimises the standard errors, and therefore makes the parameters more reliable (Molin, 2020). However, dominance of choices within the choice set is possible. This can be avoided by using a D-efficient design. Furthermore, A D-efficient design has the advantage of getting more reliable parameters with fewer respondents. The downside of a D-efficient design is that prior values are necessary. Therefore, a pilot research could be needed. There is the risk of wrong priors (and therefore biased parameters) as well. Both designs are evaluated, and the best-suited design is chosen. Then the choice sets (based on either D-efficient or orthogonal design) are presented to the respondents. In the survey, the respondents have to make trade-offs between attributes when choosing between train, plane and car. When a sufficient number of respondents have responded to the survey, the trade-offs between attributes are analysed. This is in line with the goal of the research, to identify how perceived risk due to COVID-19 among the main attributes do influence the mode choice between train, plane and car in Europe.

3.4. Predicting modal-split using the Discrete choice modelling theory

The data is gathered in the first two steps. With this choice data, trade-offs and preferences of the respondents are deduced. Based on these trade-offs, future choices are predicted using the Discrete Choice Modelling Theory. This theory is a theory to get insight in trade-offs that respondents make, introduced by McFadden et al. (1973). Especially for trade-offs regarding travel behaviour, this theory is widely used. The theory does presume that a respondents' choice is captured by a specific 'utility,' and that a respondent picks the option with the greatest value for 'utility.' Most of the time, utilities are negative, such as travel time and travel cost. This is since an increase in those attributes decreases the utility for a decision-maker. This concept is mathematically formulated in the following equation:

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

Equation 3.1 indicates that the decision maker prefers alternative *i* over alternative *j* if the utility of *i* is larger, but *i* and *j* cannot be equal. As there is never perfect information, an error term is added in the next equation:

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

All options in the choice set are evaluated by the decision-maker. There is also a weight assigned to the particular decision maker. If all parameters are linear, the deterministic utility function is given by the following equation:

$$V_i = \sum_m \beta_m \times x_{im} + \epsilon_i \tag{3.3}$$

Several kinds of discrete choice models may be used to estimate the market shares of various modes. The most often utilised model is the Multinomial Logit Model (MNL). The following equation represents this model:

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

This equation does show the probability that decision maker chooses alternative *i* from the total choice set *C*.

However, the MNL model has some significant shortcomings. At first, it is assumed that the error term is Type 1 extreme value distributed (Chorus, 2020). Secondly, the Irrelevant Alternative (IIA) property assumption holds. This means that the relative popularity of alternatives does not depend on other ones. For this research, this becomes a problem, as the rail options and aeroplanes share unobserved (transit) factors. The third issue with the MNL model is that it ignores the heterogeneity in attribute weights of the respondents (Chorus, 2020). The fourth issue with MNL is numerous choices made by a single decision-maker. The MNL assumes independence from every decision, while in reality, those decisions that are made by the respondents are not independent of each other, the so-called 'panel' effect (Chorus, 2020).

3.5. Summary

Main points of chapter

- Data is collected by using a survey
- In order to measure 'perceived risk by COVID-19', an adapted form of the HII approach is done. This is done by using a rating experiment.
- The rating experiment is estimated with a regression model.
- The main choice experiment is conducted to get insight in trade off for mode choice.
- The data is analysed using the discrete choice modeling theory. The modal-split is estimated using an MNL model.

The approach that is used is explained in detail in this chapter. The data for the Stated-Preference is obtained by a survey. To guarantee that the data is usable from the perspective of an economist and to reduce the potential bias, several elements that require attention have been highlighted. To integrate the perceived risk rating experiment and mode choice experiment, this research employs a modified version of the Hierarchical Information Integration (HII) theory. The dependent variable in the comfort rating experiment is perceived risk. A regression analysis is used to estimate the rating experiment. Perceived risk is an independent variable in the main mode choice experiment. The data is analysed using Discrete Choice Modelling. An MNL-model is used.

4

Theoretical framework & survey design

This chapter is about the theoretical framework that is used in the model. Also, the steps for survey design are discussed. Travel behaviour consists of several elements, with mode choice as one of these factors. Therefore in the first paragraph, this is discussed. Section 2 discusses different perspectives on travel behaviour. In section 3, the perceived risk attributes are elaborated. In section 4, the same is done for the main choice experiment. Section 5 discusses the included socio-demographic and travel behaviour attributes. In section 6, the theoretical framework of this research is shown. Section 7 elaborates on the steps that are used to generate the choice sets. Section 8 discusses how the survey is constructed and implemented in the software. The last section gives a summary.

4.1. Travel behaviour modelling

The influence risk effects of COVID-19 on mode choice for long-distance travel in Europe is part of a broader context. Mode choice is one of the elements of travel behaviour. de Dios Ortúzar and Willumsen (2011) created the 4-step model, which serves as the basis for modelling travel behaviour; this is shown in figure 4.1.

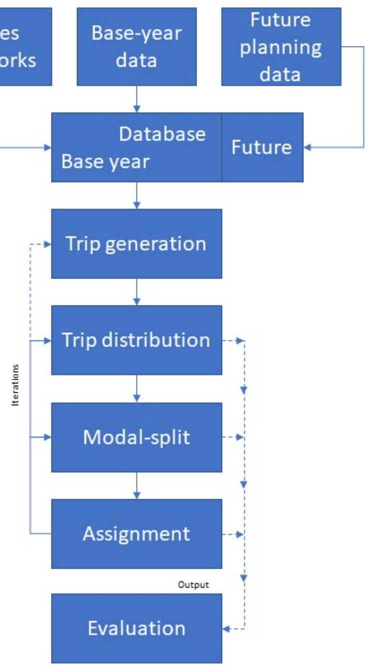


Figure 4.1: 4-step model (de Dios Ortúzar and Willumsen, 2011)

There are also more steps that decision-makers (individuals) might consider, such as period of day, the choice to travel (or not) and destination choice. This research will investigate if risk factors of infection with COVID-19 will influence the choice of mode for European long-distance travel. Therefore, the focus will be only on the modal-split step.

4.2. Perspectives of travel behaviour

There are several perspectives on travel behaviour, with the two most known being the 'econometric' and the 'marketing' perspective (Anable, 2005; Arentze and Molin, 2013). The most often used perspective is the econometric one. From this econometric perspective, it is assumed that travel behaviour is the result of a decision-making process. It makes use of the Discrete Choice Modelling (DCM) theory, which describes choice behaviour using mathematical formulas. The advantage of the econometric is thus that it is a very powerful and intuitive tool for predicting demand. This method uses a mathematical model; therefore, varying attribute levels influence choices made, and thus the direction of the causality is clear. The theoretical underpinning of the econometric perspective is plausible and consistent with welfare theory (Kroesen, 2020). The maximisation of utility, which is the same as the reduction of dis-utility, is the theory that underpins this paradigm. According to this theory, the respondent making a choice chooses the choice that provides the respondent with the highest amount of utility. The model used is called the Random Utility Model (RUM). This is not the only decision rule; another model is the Random Regret Minimisation (RRM) model. In this case, it is assumed that the respondent minimises regret. In a study done by Chorus et al. (2013), it was shown that the differences in performance could be very small. As the RUM model is more widely used, and in addition to more knowledge and software, this model will be chosen.

4.3. Perceived risk attributes

In this section, the perceived risk determinants are discussed. The determinants/attributes are based mainly on the papers of Dryhurst et al. (2020), Mertens et al. (2020), Tirachini and Cats, (2020) and Leppin and Aro (2009) (who did an extensive research on perception of risk due to SARS). All attributes have four levels. In order to research perceived risk of COVID-19, the rating experiment is done. One of the main interests of this thesis is to analyse if and how perceived risk of COVID-19 infection has an influence on mode choice for long-distance travel in Europe. The perceived risk experiment will only be done for travelling by plane and train; no distinction between these modes is made.

4.3.1. Mode-related attributes

The first four factors are mode-related attributes. These factors presumably have an influence on perceived risk during travel. Both plane and train have their characteristics, such as type of air conditioning, face mask policy and cleaning policy.

Load factor

Load factor is one of the most crucial elements of perceived risk (Tirachini and Cats, 2020). This is since the virus can be easily transmitted between people when it is crowded on board a vehicle. Transmission in the air within a closed environment is studied by several papers (Morawska and Cao, 2020; Shen et al., 2020; Tang et al., 2020), therefore an closed environment is more risky than an open environment (Tirachini and Cats, 2020). The Dutch government introduced a 1.5-meter distance rule between people to reduce the probability of infecting each other (RIVM, 2020). Service providers all over the world reduced the number of passengers allowed on board vehicles; sometimes also, the frequency was reduced. Services providers in the UK reduced ridership by 30% in trains (Guardian, 2020). In order to reduce the transmission risk, 1, 1.5 or 2 meters is needed (Jarvis et al., 2020). Several papers researched the necessary reduction in capacity to meet the social distancing rules. Krishnakumari and Cats, (2020) stated that for the Washington D.C. metro system, 80% of the capacity reduction is needed in order to maintain 1.5-meter social distancing and only 10% when implementing a 2-meter distance rule. At the same time, (physical) distancing conflicts with the concept of public transport as volume is needed to cover the cost (Musselwhite et al., 2020). Tirachini et al., (2013) concluded that crowding could negatively affect passengers onboard a vehicle in several ways when the risk of infection is not included. This includes health and safety concerns, like stress. It was found that time valuation is increased with higher load factors on-board vehicles (De Palma et al., 2017). COVID-19 makes crowding onboard vehicles even more important than already was the case. Shelat et al., (2020) showed that when passengers are offered less crowded vehicles, willingness to board (train) vehicles is increased. For this attribute, it is expected that there will be a positive relationship between crowding and perceived risk. This is due to two reasons. If passengers are more close to each other, there will be an increased risk of infecting each other. The second reason is that crowding increases the probability of getting exposed to an infected (other) passenger. Thus, an increase of both reasons will increase both perceived risk and risk of infection with COVID-19. Occupation of seats is one of the most used methods to operationalise crowding onboard a vehicle (Li and Hensher, 2011). Therefore this percentage of occupied seats will be used to operationalise onboard crowding.

Attribute levels: Onboard crowding is often specified as the percentage of seats occupied onboard the vehicle/plane. The levels are: 25% of the seats occupied, 50% of the seats occupied, 75% of the seats occupied and 100% of the seats occupied.

Face mask policy

The usage of face masks is one of the measures that has been in use (and is still in use) to reduce COVID-19 infections in areas where 1.5 meters is not always possible (Greenhalgh et al., 2020). There are several arguments against the usage of face masks. Some of these arguments are limited evidence of efficiency, false security as people do not comply with basic rules and misuse of the masks due to lack of information (Tirachini and Cats, 2020). First, the World Health Organization (WHO) did recommend only using face masks for people with COVID-19 symptoms (WHO, 2020b). After this, the WHO advised wearing non-medical masks in public transport and public places and medical masks for the more vulnerable groups (WHO, 2020a). Partly the advice for non-medical masks was because, at the beginning of the pandemic, there was a shortage of medical masks (Greenhalgh et al., 2020). Konda et al. (2020) found that when using different fabrics (silk, cotton) and different layers combined could approach a similar level of protection to that of medical masks. Even though there is doubt about the efficiency of face masks, more recent research does advise using face masks. According to (Chu et al. (2020) and Eikenberry et al. (2020), face masks can reduce the particles of COVID-19 in exhaled breath significantly, in particular for people that have mild symptoms and/or are asymptomatic. The effectiveness of the fabric masks is more than 80%-90% for certain combinations of fabric, so cotton, chiffon, silk and flannel (Konda et al., 2020). In the New York and Washington states, 80% usage by people of 50% effective masks, could prevent 17-45% of deaths by COVID-19 in these states (Eikenberry et al., 2020). Even very low protective masks (with a 20% protection rate) can be useful. The greatest gain of face masks will be in combination with other measures like social distancing (which is described above) (Eikenberry et al., 2020). Wearing a face mask serves a dual preventive purpose: one protects one another from getting a viral infection, and one protects itself from others (Abboah-Offei et al., 2021). The effectiveness of face masks will presumably influence perceived risk. Perceived risk is subjective rather than the objective effectiveness of face masks. Within the rating experiment, it is measured if and how face masks influence perceived risk. It is expected that the mandatory usage of face masks will have negative relationships with perceived risk, as face masks will decrease the probability of infecting each other.

Attribute levels: The attribute levels are a representation of the types of masks available. The four levels are: no mask mandatory, any face mask mandatory, at least a surgical (type II) mask mandatory or at least an FFP2 mask mandatory. Every increase in the level of the type of mask gives better protection.

Cleaning policy

The COVID-19 virus can stay infectious from hours to days on several different surfaces, which includes stainless steel and plastic (Van Doremalen et al., 2020; Chin et al., 2020). A study done by Kampf et al. (2020) showed that the COVID-19 virus could even stay as long as nine days on surfaces like glass, metal and plastic. It could therefore be that COVID-19 will be transmitted via surfaces. This is the reason why public transport operators and airlines have increased their cleanings policies due to COVID-19 (KLM, n.d.). However, it is not 100% clear if cleaning policies are very effective in reducing the transmission of COVID-19. There are several experts stating that infection with COVID-19 via surfaces is very rare (Thompson, 2020). Moreover, the Centers for Disease Control and Prevention (CDC) stated that transmission via surfaces is one of the least common ways of COVID-19 transmission (Newsroom, 2020). Nevertheless, companies all over the world adopted increased levels of (extra) cleaning and sanitisation within vehicles (Krishnakumari and Cats, 2020). It is expected that extra levels of sanitisation and cleaning in vehicles (and planes) will decrease the perceived risk. If people know that vehicles are more often cleaned, they presumably believe that the probability of infection with COVID-19 is smaller.

Attribute levels: The four levels are: same cleaning policy as before COVID-19, increased cleaning policy (focus on touching points), weekly disinfection of the whole vehicle and daily disinfection of the whole vehicle. The levels are based on several policies that airlines and rail companies implemented and still implement.

Ventilation

Whether the usage of air conditioning and ventilation systems does contribute to the spreading of COVID-19 remains partly unclear. Likely, it could happen that recirculated air will be used which is not filtered (Tirachini and Cats, 2020). Some evidence can be found about the fact that air conditioning can contribute to infection with COVID-19 in indoor environments (Lu et al., 2020). The CDC recommends using air condition systems that do not use recirculated air, so only on the non-recirculating mode (CDC, 2021). Another preventive measure that often is taken and recommended, is (frequent) ventilation of closed environments (Lu et al., 2020); CDC, 2021). In vehicles, this becomes relevant for both passengers and drivers. Especially for trips (with people often sitting for hours in the same cabin) this is important (Tirachini and Cats, 2020). The United Kingdom recommended the following flow rate for ventilation for buildings of 8-10 liters per second per person of fresh air, excluding re-circulation (Gartland et al., 2021; Bhagat et al., 2020). Several filters are used and can be used to filter the air onboard vehicles. Often the so-called HEPA (high-efficiency particulate air) filters are used. These HEPA filters are very efficient in filtering air, they are as effective as 99.9% (Chuaybamroong et al., 2010; Saini and Saini, 2020). (Almost) all planes have HEPA filters within the air conditioning systems. Every 2-3 minutes, the air on-board a plane is renewed (Hunt and Space, 1994). The distributed air from the air conditioning system is distributed via panels above the passengers. The air that was already in the cabin gets sucked down by vans on the cabin floor (Hunt and Space, 1994). The ventilation system onboard TGVs routinely exchanges inside air with outside air. The ventilation system draws air that is inside the train, filters it, and then it will be mixed with air from outside (SNCF, n.d.). The proportion of filtered and fresh air depends on the type of train but in general, 2/3 filtered air and 1/3 fresh air (SNCF, n.d.). Instead of an airflow downwards, the airflow is upwards (vertical). Both trains and aircraft have ventilation systems with filters; the systems work (on average) the same. These systems will have a negative relationship with perceived risk. This means 'having' such a system will decrease perceived risk. For this attribute, three levels can be found.

Attribute levels: Airplanes do (almost) always have HEPA-filter on-board (Korbee, 2020). For trains, it does not become fully clear if trains have HEPA filters onboard or not. SNCF stated that the air is refreshed every few minutes but does not state that the train has a HEPA filter. Italo (a private Italian train company) do state that their trains have HEPA filters onboard (Italo, n.d.). So it depends on the train. The levels are: no ventilation and air conditioning, only ventilation, air conditioning without HEPA filters and air conditioning with HEPA filters.

4.3.2. Destination-related attributes

The last four perceived risk attributes are destination-related. Either travelling by train or plane will not change these attributes as these are related to the destination.

Vaccination/recovery and/or testing requirements

Due to COVID-19, almost all countries changed or introduced entry rules or sometimes closed the border completely. Because of the ban or extra formalities, people travel less while travelling is harder (Chinazzi et al., 2020). At the same time, the extra measures can have a negative contribution on perceived risk. The reason for this will presumably be that the need for testing/recovery/vaccination decreases the chance that someone is infected, and therefore the risk of infection will be decreased as well (RIVM, 2021). Due to this decreased infection risk, presumably perceived risk by people decreases as well. For the different modes, the requirements can be different, with, on average, more strict measurements for flying. Digital health passports, for example, might assist in standardising screening criteria at airports and border crossings, allowing for a safer return to travel (Khatib et al., 2020). The same would count for trains; however, as there are, on average fewer security/checks, it is easier to travel without recovery/vaccination/test. So it is expected that the need for testing/vaccination/recovery has a negative relationship with perceived risk. This means that the increased need does decrease perceived risk.

Attribute levels: Several policies are implemented within Europe, like 3G or 2G. The following levels are chosen as they reflect different policies within Europe: no mandatory requirements, either testing, vaccination or recovery proof required (3G-rule), only vaccination or recovery proof required (2G-rule), and vaccination or recovery + testing required or booster required (2G+-rule).

Infection rate

The infection rate is the number of confirmed cases in a country; in this case, this will be the Netherlands. The RIVM (National Institute for Public Health and the Environment in the Netherlands) keeps track of the number of confirmed cases. According to the RIVM (2021): "We can measure the rate at which the virus is spreading before it reaches intensive care by keeping track of the number of confirmed cases. This allows us to be prepared." The number of confirmed cases is thus an indication of the actual number of infections within a country. This increase in infections will also increase the probability of being infected. This risk factor is an exogenous factor, in contradiction to the aforementioned risk factors. In a study done by Wang (2014), it was found that during the SARS outbreak in 2003 in Taiwan that the number of infections was an important factor for predicting the usage of public transport. The infection rate can be treated in two ways. As a first way, it can be included in the rating experiment as a risk factor. In a second manner, it can be used as a context variable in the main choice experiment. For this research, it will be chosen to include it in the rating experiment. This is due to the fact that, in this case, a direct relationship can be found between infection rate and the perception of risk. This factor will be operationalised by number of positive tests per day (RIVM, 2021). As stated earlier, this is still an estimate, but it gives a good indication of the percentage of infections on average.

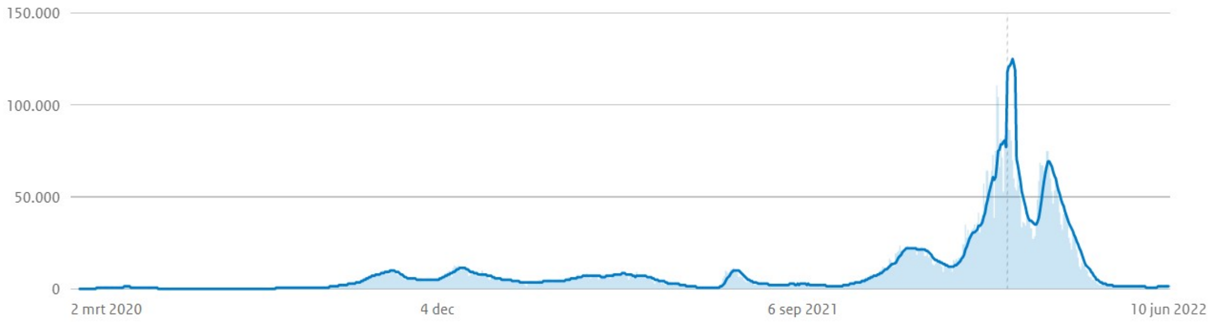


Figure 4.2: Number of positive tests per day (RIVM, 2022)

Attribute levels: The infection rate levels do reflect different time moments during the pandemic. The levels are: 100 positive tests per day (reflects the situation of the summer of 2020 or in June 2021), 10.000 positive tests per day (reflects the situation of November 2020 and July 2021), 25.000 positive tests per day (reflects the situation of November 2021) and 100.000 positive tests per day (fictitious situation).

Vaccination rate

At the beginning of 2021, the BioNTech/Pfizer vaccine was the first vaccine used in the Netherlands (Brabants-Dagblad, 2021). From then on, as more vaccines got available, more and more people got vaccinated. At the moment of writing, there is a vaccination rate of 68.4% (June 2022) (RIVM, 2022). Research showed that the increase in vaccination rate does decrease the probability of getting infected (Chen, 2021). At the same time, someone can still get infected with COVID-19, but the probability of severe symptoms is very small. Therefore, it is expected that the vaccination rate has a negative relationship with perceived risk. This means that a higher vaccination rate presumably decreases perceived risk. However, vaccines do not protect 100%. Therefore there is always a risk of getting infected even when someone is vaccinated. But this research will focus on risk perception, so this (subjective) risk perception will presumably decrease with higher vaccination rates.

Attribute levels: The levels reflect the vaccination rates in different European countries. The levels are from December 2021: 15% vaccination rate (level as in Bulgaria), 30% vaccination rate (level as in Romania), 70% vaccination rate (level as in the Netherlands and E.U. average), 90% vaccination rate (level as in Portugal).

Travel advice

Europe and other countries all over the world introduced travel advice regarding COVID-19. Most of the time, this travel advice was already there, stating which risk factors could be expected when travelling to a certain country (Rijksoverheid, 2021). The travel advice website (Union, 2021) rates countries on severity by giving colors. Green means low risk, yellow means potential risk, orange means high risk and dark red means very high risk.

Attribute levels: These are the travel advice from the Dutch government (Rijksoverheid, 2021). The levels are: green, yellow, orange and red.

Table 4.1: Rating attributes and attribute levels

Risk factors	# levels	Explanation attribute levels	Category of risk	Attribute	Type
On-board crowding/load factor	4	0: 25 % of seats occupied 1: 50 % of seats occupied 2: 75% of seats occupied 3: 100 % of seats occupied	Mode/trip	crow	Ratio
Face mask policy	4	0: No face mask mandatory 1: Face mask mandatory (any) 2: Surgical mask mandatory 3: FFP2 mask mandatory	Mode/trip	mask	Ordinal
Cleaning policy	4	0: Same cleaning policy as before COVID-19 1: Increased cleaning policy (touch points) 2: Weekly disinfection whole vehicle 3: Daily disinfection whole vehicle	Mode/trip	clean	Ordinal
Ventilation/air conditioning	4	0: No ventilation and air conditioning 1: Only ventilation 2: Air conditioning without HEPA filters 3: Air conditioning with HEPA filters	Mode/trip	airco	Ordinal
Vaccination/recovery/testing requirements	4	0: No mandatory requirements 1: Either testing, vaccination or recovery required (3g-rule) 2: Only vaccination or recovery (2g-rule) 3: Always testing, and vaccination or recovery (2G+-rule)	Destination	req	Ordinal
Infection rate	4	0: 100 positive tests per day (summer 2020/June 2021) 1: 10.000 positive tests per day (October/November 2020/July 2021) 2: 25.000 positive tests per day (November 2021) 3: 100.000 positive tests per day (fictitious extreme number)	Destination	infect	Ordinal
Vaccination rate	4	0: 15 % fully vaccinated people (Bulgaria) 1: 30 % fully vaccinated people (Romania) 2: 70 % fully vaccinated people (Netherlands and EU average) 3: 90 % fully vaccinated people (Portugal)	Destination	vacc	Ratio
Travel advice	4	0: Green advice 1: Yellow advice 2: Orange advice 3: Red advice	Destination	advice	Ordinal

4.3.3. Context of rating experiment

For respondents, the context of the survey must be clear. If not, respondents will make their own assumptions when information is missing. This will result in data that has a lower quality. Of course, there is not 'one approach' for constructing a survey with its context, but it is important to add as much information as possible. This is always in consideration with the length of the survey. Because of this, decisions regarding the context of the survey are discussed. The main goal of the research is to research whether perceived risk of COVID-19 does influence mode choice for long-distance travel within Europe. First, the context of the rating experiment is discussed. After this, the main choice experiment context is discussed.

In section 4.3 it is discussed which attributes are included in the rating experiment. All of the attributes have four levels and are based on the context in Europe regarding COVID-19. This research uses two distance classes to research the main choice attributes. Respondents are told as an assumption that in the rating experiment, the travel they are taking always has the same duration. This was done to account for the time component in the rating experiment. Krishnakumari and Cats, (2020) stated that a longer duration could possibly increase the risk of infection with COVID-19, and therefore perceived risk can increase as well. By combining the rating experiment with both distance classes of the main choice experiment, it can be analysed that for longer trips, people have a higher perceived risk. Furthermore, respondents had to assume that they were either travelling by plane or by train. Thus, there was no distinction made between these modes. This was done to keep the survey easy and not too long. In order to analyse the relationship between perceived risk and each mode, interaction effects between perceived risk and mode-specific attributes will be estimated.

4.4. Main choice attributes

In this section, the main choice experiment variables/attributes will be discussed. The determinants of the main choice experiment are travel time, travel cost, travel comfort and perceived risk. All attributes have three levels. The attributes of travel time and travel cost are varied for all modes. Travel comfort and perceived risk are only varied for plane and train. Perceived risk and travel comfort are not varied for car. This is because it is assumed that respondents are not sharing their car with strangers. As a consequence, perceived risk in the car is always very low. For travel comfort, it is assumed that people' own' the same car within the experiment. Therefore, the comfort of the car does not change; thus, the levels of comfort are not varied. The attribute levels can be found in table 4.2, 4.3, 4.4 & 4.3.

4.4.1. Included mode choice attributes

Travel time

Travel time is (almost) always taken into account in studies regarding mode choice (Morikawa et al., 2002; van Goeverden, 2009; Román et al., 2010). Some studies refer to the total travel time (thus including access, egress, transfer and waiting time). Other studies only refer to the in-vehicle time. In this study, the total travel time will be used. This includes in-vehicle time and transfer time (if this applies). **Access and egress** are incorporated in the total travel time in this study. This is done in order to keep alternatives simple and understandable and because of the fact that this is not relevant for this study. The aim of this study is to investigate the effects of COVID-19 risk on mode choice. Therefore total travel time will be simplified. There will be made use of distance classes for this case. For each of these distance classes, there will be different attribute levels. In total, there will be two distance categories:

- 400-600km: Destinations such as Paris, London, Zürich, Berlin en Copenhagen.
- 800-1200km: Destinations such as Bordeaux, Milan, Barcelona, Warsaw en Stockholm.

Attribute levels

Travel time is calculated using different sources in order to construct realistic travel times for the different modes. The main choice experiment has two distance classes, 400-600 kilometres and 800-1200 kilometres. For each distance class, the following steps were taken in order to construct the travel time levels for the main choice experiment. First, the travel time for the shortest OD-pair within the distance class is calculated. After this, the travel time for the longest OD-pair is calculated. Then the mid-point of these two travel times is taken. The construction of values for travel time is rather an interactive approach.

- **Car**: For the calculation of the travel times of car, the websites ViaMichelin.nl, Rome2Rio.com and Google Maps are used. ViaMichelin gives a very detailed indication for both travel time and travel cost (ViaMichelin, n.d.). After this, Rome2Rio and Google Maps are used to compare travel times found at ViaMichelin. (Rome2Rio, n.d.; Google, n.d.). This study does not focus on travel time variability. Therefore for the shortest OD-pair within the distance class, the free-flow travel time is used. For the longest OD-pair, the free-flow travel time plus some extra travel time to account for busy roads or traffic jams. For the 400-600 kilometres distance class, an hour is added to the longest travel time. For the 800-1200 kilometres, one and a half hours is added to the longest travel time. Now, a wide range of travel time is available so that as much possible travel time values are within this range. The levels for the 400-600 kilometres distance class: are 4.5 hours, 6.5 hours and 8.5 hours. The levels for the 800-1200 kilometres are 10 hours, 13 hours and 16 hours.
- **Train**: For train, mainly the website Rome2Rio is used. About the same approach is taken as in the calculation of the travel times for car. For each distance class, the travel time for the shortest OD-pair is calculated for the first level using the website Rome2Rio. The same is done for the longest OD-pair within the distance class for the third level. Then the mid-point was taken as the second level. The travel time found on this website does include transfer time but does not include travel time from the station to the city centre. Therefore, the travel time from the station to the city centre was calculated and added as well. For this calculation of the travel time, Google Maps was used. To find the central

point as the city centre, Google Maps was used as well. Google Maps was also used to calculate the total travel time door to door for train as a check from Rome2Rio. Sometimes, if no information was available from Google Maps, the website from N.S. international was used as well (NS, n.d.). For the OD-pairs, almost always the same travel time was found. This confirmed realistic travel times levels for the main choice experiment. The levels for the 400-600 kilometres distance class: 3 hours (Amsterdam-Paris), 4.5 hours and 6 hours. The levels for the 800-1200 kilometres are 6 hours, 9 hours and 12 hours.

- **Plane**: For the calculation of the travel times for plane, Rome2Rio was used as well. Rome2Rio calculated the total travel time, which includes access and egress times to the airport, waiting time at the airport and flight time for the flight route. Rome2Rio calculates one and a half-hour of waiting time for the short distance class and two hours of waiting time for the longer distance class. Again for the first level, the shortest possible travel time is calculated for the shortest OD-pair. Then for the third level, the total travel time for the longest OD-pair is calculated. Then for the second level, the mid-point is taken. Google Maps and Skyscanner are used to check the travel times calculated by Rome2Rio. Again almost all the travel times from Rome2Rio were similar to the ones calculated with Google Maps and Skyscanner (Skyscanner, n.d.). The levels for the 400-600 kilometre distance class: are 3 hours, 4 hours and 5 hours. The levels for the 800-1200 kilometre distance class are 4 hours, 5 hours and 6 hours.

Table 4.2: Travel time levels

Travel time							
400-600 km				800-1200km			
# levels	Mode	Value attribute levels		# levels	Mode	Value attribute levels	
3	Train	180 min	3h	3	Train	360 min	6h
		270 min	4.5h			540 min	9h
		360 min	6h			720 min	12h
	Airplane	180 min	3h		Airplane	240 min	4h
		240 min	4h			300 min	5h
		300 min	5h			360 min	6h
	Car	270 min	4,5h		Car	600 min	10h
		390 min	6,5h			780 min	13h
		510 min	8.5h			960 min	16h

Travel cost

Travel cost is one of the most important variables in travel behaviour research and (almost) always included within stated choice experiments. Travel cost does refer to the cost of making a trip. This can be either the ticket price or the total price for driving the car (fuel + any additional cost). Travel cost is often used in choice experiments to retrieve the willingness to pay for improvements in one of the other attributes of interest. In the case of this research, it would be researched how COVID-19 risk influences mode choice for long-distance travel within Europe. In this case, it is studied how price is traded against COVID-19 (infection) risk. From the literature review, it can be concluded that travel cost was one of the most important variables for mode choice on long-distance travel (van Goeverden, 2009; Román et al., 2010; Dobruszkes et al., 2014). For the train option, the ticket price is often based on the distance travelled. However, for long-distance, this is not always the case. There is a possibility of specially reduced prices. The prices on https://www.nsinternational.com/ are used as inspiration for the price levels. Ticket prices for airlines are based on demand and supply and, to a lesser extent, on distance. For these ticket prices, realistic values are to be found based on certain OD-pairs (classes). For cars, the total cost is varied, including fuel cost and all other costs.

Attribute levels

The values of the levels of travel cost are calculated using the following websites: Skyscanner, KLM, NS international and ViaMichelin. A similar approach is used for the travel time. For the shortest OD-pair within the distance class, the cheapest possible fare is used for the first level. For the third level, the highest possible fare for the longest OD-pair is used. The mid-point is used for the second level.

- **Car**: For car ViaMichelin is used. It gives a detailed overview of the costs of travelling by car. The cost includes toll costs and fuel costs. The calculation was done in December 2021; this is important because fuel costs fluctuate. For the car, the 'hatchback' type is used and the average price of EURO 95 fuel at the moment of search. To test whether the costs at ViaMichelin are a good approximation, the cost were also calculated by using Google Maps (for calculating the distance), average fuel costs and the average fuel consumption of a car. The approximations done by ViaMichelin were very close to the calculation via Google Maps. The levels for the 400-600 kilometres distance class: are €80, €115 and €150. The levels for the 800-1200 kilometres are €100, €150 and €200.

- **Train**: For the train Rome2Rio is used. It gives a price indication but does not include access and egress to/from the station. Often stations are in the city centre. Therefore it is assumed that the train fares reflect the full price. Often the costs of access and egress from the station can be neglected. In addition, the prices were checked by checking fares on NS international, SNCF and Deutsche Bahn (SNCF, n.d.; Bahn, n.d.). Ticket prices start already at €35 from Amsterdam to Paris. To account for the full range of possible values, the first level for the 400-600 kilometre distance class is set to €30. The highest value for travel cost in the 400-600 kilometre distance class is set €300. This reflects the price of a last-minute trip to London in the 1st (premier) class on board the Eurostar. Higher prices are possible, but the €300 is assumed to be a realistic price within this distance class. For the second level, the mid-point is taken, so the levels are €30, €165 and €300. For the 800-1200 kilometres, the cheapest possible fare for the Nightjet is around €50, so this is chosen as the lowest level. For the highest level, €350 is chosen as the maximum, but also higher fares are possible. However, €350 is assumed as a realistic price for the highest level. The mid-point is €200.
- **Plane**: For plane, the website Skyscanner.nl is used (Skyscanner, n.d.). This website finds prices for flights to destinations. Also, the website of KLM is used to find fares that are valid. For plane, within Europe, distance does not always reflect the height of the costs. For example, a return ticket from Amsterdam to London is often more expensive than a return ticket from Amsterdam to Barcelona (Skyscanner, n.d.). For KLM, tickets are offered for €100 return; therefore (as all travels are one-way), the lowest level for the plane is €50. This is for both the 400-600km distance class and the 800-1200km distance class. The highest level for the 400-600km distance class is set to €300. This fare is found for business class to London on KLM, one week before departure (Skyscanner, n.d.). Higher values are possible, and also lower values are possible, but this fare was the average found. Also, it is the most offered fare on the website. For the second level, again, the mid-point is chosen, which is €175. For the 800-1200km level, the highest value is set to €400. This is because, on average, the prices for the longer distance class are a bit higher. For the second level, the mid-point is chosen, which is €225.

Table 4.3: Travel cost levels

Travel cost					
400-600 km			800-1200km		
# levels	Mode	Value attribute levels	# levels	Mode	Value attribute levels
3	Train	30 euro	3	Train	50 euro
		165 euro			200 euro
		300 euro			350 euro
	Airplane	50 euro		Airplane	50 euro
		175 euro			225 euro
		300 euro			400 euro
	Car	80 euro		Car	100 euro
		115 euro			150 euro
		150 euro			200 euro

Travel comfort

Travel comfort is also an important factor regarding mode choice for long-distance travel. Román et al. (2010) included comfort as an attribute in their research. The willingness to pay increased if the level of comfort was lower. Furthermore, they found that increased levels of comfort in the plane did decrease the perception of time. Train companies and airlines do offer different levels of comfort by distinguishing travel classes. For plane, often, there is a choice between economy class and business class. For the train, there is (almost) always 2^{nd} class and 1^{st} class.

Attribute levels

Travel comfort has two levels for both plane and train. As earlier explained, only legacy carriers are included in the main choice experiment. Therefore, two levels are included, economy class and business class. For train, there are also two levels, 2^{nd} class and 1^{st} class. For car, it is assumed that the level of comfort does not change.

Table 4.4: Travel comfort levels

Travel comfort		
# levels	Mode	Value attribute levels
2	Train	2nd class 1st class
2	Airplane	Economy Business
1	Car	Same level

Perceived risk

Perceived risk is the last attribute that will be included in the main choice experiment. This attribute is directly connected with the rating experiment. In the rating experiment, respondents rated their risk of getting infected with COVID-19 on the train or plane journey based on eight factors. Respondents rated their journey on a Likert scale with 1-very low, 2-low, 3-medium, 4-high and 5-very high. In the main choice experiment, perceived risk is an attribute amongst the other main attributes. As all levels in the main choice experiment have three levels, perceived risk has three levels as well. Therefore, the levels are 1-low, 3-medium and 5-high.

Attribute levels

This attribute is only varied for the plane and train modes. This is because it is assumed that people do not share their car with strangers. As a consequence, perceived risk in the car is always assumed to be very low. For train and plane, three levels will be varied. These levels are directly derived from the rating experiment. However, the 2^{nd} and 4^{th} level are not included. In this case, perceived risk is a given, and respondents do not rate their risk as they did in the rating experiment.

Table 4.5: Perceived risk levels

Perceived risk		
# levels	Mode	Value attribute levels
3	Train	1-Very low 2-Medium 3-Very High
3	Airplane	1-Very low 2-Medium 3-Very High
1	Car	1-Very low

4.4.2. Mode choice attributes *not* included

More attributes are important regarding mode choice for long-distance travel in Europe. However, they are not included in the research. Partly, this is due to the size of the research and the sake of simplicity. Moreover, this is because of the fact that often these attributes do not turn out to be significant. The attributes will be shortly discussed.

Access and egress time: This factor is included within the total travel time. As this research does not focus on one specific destination but rather distance classes, access and egress time would be very hard to incorporate. Often it is included in studies about public transport, but these studies are on the smaller scale (Morikawa et al., 2002; Román et al., 2010).

Transfers: In the first design phase of the survey, transfers were included in the main choice part. For the sake of simplicity and due to the focus of this research on perceived COVID-19, transfers are not included. Moreover, in order to keep the research (and survey) clear and easy to understand, transfers are not included.

Reliability: Travel time reliability is an important factor for travel behaviour. Román et al. (2010 included this factor into their research. Delays for both trains and planes are not uncommon. However, it is hard to incorporate this into the main choice experiment. For example, a percentage (extra) travel time could be varied in the main choice. However, as the main choice will consist of two distance classes, there will be too many attributes for the respondents. Therefore, it is chosen to not include reliability in this research.

Attributes regarding car: Several studies included attributes regarding car in their research. For example Hensher and Rose (2007) included daily parking cost and toll costs. For this research, it is assumed that the cost of the car does already include the toll costs. This will be further elaborated on in the section on survey design. Parking costs are not taken into account, as the focus of this research is only on the trip from origin to destination. Parking costs arise after reaching the destination, so, therefore, are not included in this research. Thus, for the cost attribute for car, toll cost is already incorporated into the total cost.

Frequency: Within the public transport area, frequency is one of the most important factors. Again Román et al. (2010) used this in their research on the Madrid-Barcelona corridor. As this research is more a general approach to long-distance travel and not based on a case study like the Madrid-Barcelona corridor, it is hard to incorporate frequency for all potential OD-pairs. Moreover, the focus of this research is on travel time, cost, comfort and COVID-19. Thus, frequency is not included in this research.

Waiting time: Hensher and Rose (2007) included waiting time into their research. In this research it is included in the 'total' travel time.

Departure / arrival time: Bhat (1998) and Hensher and Rose (2007) both found that including departure and arrival time did increase the modal fit. As this research will have a general approach rather than a certain OD-pair, including a departure and arrival time would not make a lot of sense. Therefore, this will also not be included in the research.

4.4.3. Context of main choice experiment

This research takes a general approach to travel within Europe, starting from the Netherlands. As stated earlier, the main choice experiment will be divided into two distance classes, 400-600 kilometres and 800-1200 kilometres. This is done to keep the survey clear and understandable for respondents. In addition, a few example destinations will be mentioned so that respondents can imagine what kind of trip they will be making.

It is assumed that respondents are travelling from city centre to city centre. For travel time, this means that it included access and egress time. For travel cost, it is assumed that for each mode, travel cost reflects total cost, so including the costs for access and egress. For car, the costs reflect the total costs, including gas costs, toll costs and costs regarding wear and depreciation. Perceived risk is a given, and respondents do not have to rate it by themselves in the main choice experiment.

Moreover, for the mode plane, it is assumed that respondents are travelling by legacy carriers like KLM, Lufthansa or Air France, so no low-cost carriers like Ryanair or EasyJet. The reason for this is that low-cost carriers often offer such low prices that are not realistic. Often people plan the destination based on price rather than choosing a destination and then looking for tickets. In order to 'give' every mode a fair probability of being chosen, such low prices are not included; hence low-cost carriers are excluded. For trains, it is assumed that people are using common-rail services starting from the Netherlands like Thalys/TGV, ICE, Eurostar and the NS international intercity Berlin. At last, respondents have to assume that the trip was paid for as specified earlier in the survey by the respondent him or herself, together with the travel purpose that was specified.

4.5. Socio-demographic variables and travel behaviour related variables

Other than the rating and main choice part, the survey does contain questions about socio-demographics and travel behaviour. The questions are added to the survey in order to account for interaction effects of socio-demographic variables and give insight into how the travel behaviour of the respondents influences mode choice. The questions on travel behaviour are asked in the beginning to introduce the respondents to the subject and get familiar. At the ending of the survey, socio-demographic questions are answered, as these questions are easy for the respondents and will not exhaust too much. This order is done to account for as many completes as possible. Socio-demographic characteristics are crucial to include in the study since they provide insight into the composition of the respondent sample. In addition, given these factors, it is feasible to get insight into how these socio-demographic characteristics affect the main choice attributes and, therefore, mode choice. Additionally, acquired data may be utilised to identify distinct market segments and user groups. When estimating choice models, Ben-Akiva and Bierlaire (1999) stated that socio-demographic factors must be included to explain probable individual heterogeneity. It can also enhance modal fit. Socio-demographic characteristics may have both direct and indirect influences on utility via interaction effects with other variables. The next paragraph will address these factors.

4.5.1. Socio-demographic variables

Age: Is one of the most common used socio-demographic variables. Several papers and research did look include this into their research: Buehler and Nobis (2010), Hensher and Rose (2007), Paulssen et al. (2014), Román et al. (2010). Often different age groups have different preferences for certain modes.

Gender: This socio-demographic variable is also a very common variable to include in stated-choice experiments. Almost all studies include this socio-demographic variable (Buehler and Nobis 2010; Hensher and Rose 2007; Paulssen et al. 2014; Román et al. 2010; Johansson et al. 2006). With this variable, it can be analysed if women and men have different preferences regarding the variables in the main choice experiment, this could be for example in the preference for a certain mode.

Income: Also an important socio-demographic variable to take into account. It is not always clear how this income is asked in the survey. Some papers ask about dispensable income, while at the same time other papers ask for gross income (Buehler and Nobis 2010; Hensher and Rose 2007; Paulssen et al. 2014; Román et al. 2010; Johansson et al. 2006). As it is expected that higher income will influence mode choice, this variable is included in the model. It is expected that higher income will increase the willingness to pay for the attributes, like time and comfort for example.

Car availability: Often considered as well. Both Buehler and Nobis (2010) and Limtanakool et al. (2006) included car availability into their research. For this research, it is not included.

Work status: Hensher and Rose (2007) included this variable in their research in order to check if the sample was representative, but they did not include this in the model specification. The same will be done for this research, and in addition to this, it is included in the model interactions.

Education level: Socio-demographic variable that is often included in models as well. Johansson et al. (2006) stated that he previously did not find any literature on including education level for long-distance travel; however, in his research, it turned out to be significant. Education level is also expected to have an influence on mode choice and, therefore, will be included in the model.

Household size and composition: This socio-demographic variable is sometimes included in studies and sometimes not. For example Buehler and Nobis (2010), Hensher and Rose (2007), Limtanakool et al. (2006) and Johansson et al. (2006). For this research it will be included to test if the sample is representative, but it will not be included in the model.

4.5.2. Travel behaviour/trip characteristics

The survey will lastly consist of a few questions on the travel behaviour of the respondents. This is done to test if these factors will influence the main choice variables and, as a consequence, the mode choice.

Travel frequency: This variable is sometimes included in research. Several research included this such as Román et al. (2010), Van Loon and Rouwendal (2013) & Nieto-García et al. (2020). For this research, this attribute will be included to test whether travel frequency of respondents influences perceived risk.

Preferred travel mode: Hensher and Rose (2007) included the preferred travel mode in their study. For this mode choice research, it is interesting to see if the preferred travel mode influences the mode choice. It is expected that people will stick to their main preferred mode when making a choice. In this research, it will be questioned for both the 400-600km and 800-1200km distance classes.

Payment: This attribute is included in order to test whether the value of time values changes when payment is made by the respondents or by someone else, or education/work. Kouwenhoven et al. (2014) found significant different in the VoT values for different purpose of work.

Trip purpose: Both Buehler and Nobis (2010) and Román et al. (2010) included trip purpose in their research and for both studies this turned out to be an important factor. Often willingness to pay for business trips is higher (as the respondent does not pay by him or herself) than for leisure trips. Therefore this will also be taken into account in this research. However, this is not done in the same manner as in these studies. In this research, people are asked if they pay for themselves or if someone else is paying for their trip. In this case, it can still be analysed if the willingness to pay is higher if respondents do not have to pay for themselves.

Travel company: A study done by Mertens et al. (2020) showed that risk for family and loved ones was one of the most important predictors of COVID-19 fear. Therefore it is also included in this research to test whether travel company contributes to perceived risk.

Worry COVID-19: This attribute is about the fact that people worry more or less about the omicron-variant in comparison to the delta-variant. It is included to test whether respondents are more or less worried about the omicron-variant in comparison to the delta-variant.

4.6. Theoretical framework

The theoretical framework incorporates the many factors described in the previous sections. Figure 4.3 provides a graphical presentation of the framework used to construct the Discrete Choice Model. On the left side, the perceived risk rating attributes are represented. The first four attributes are mode-related, and the last four attributes are destination-related. In the upper part, the included socio-demographic variables are represented. On the right side, the travel behaviour attributes are represented. The unobserved variable, which is the utility of choice, is represented by the oval box.

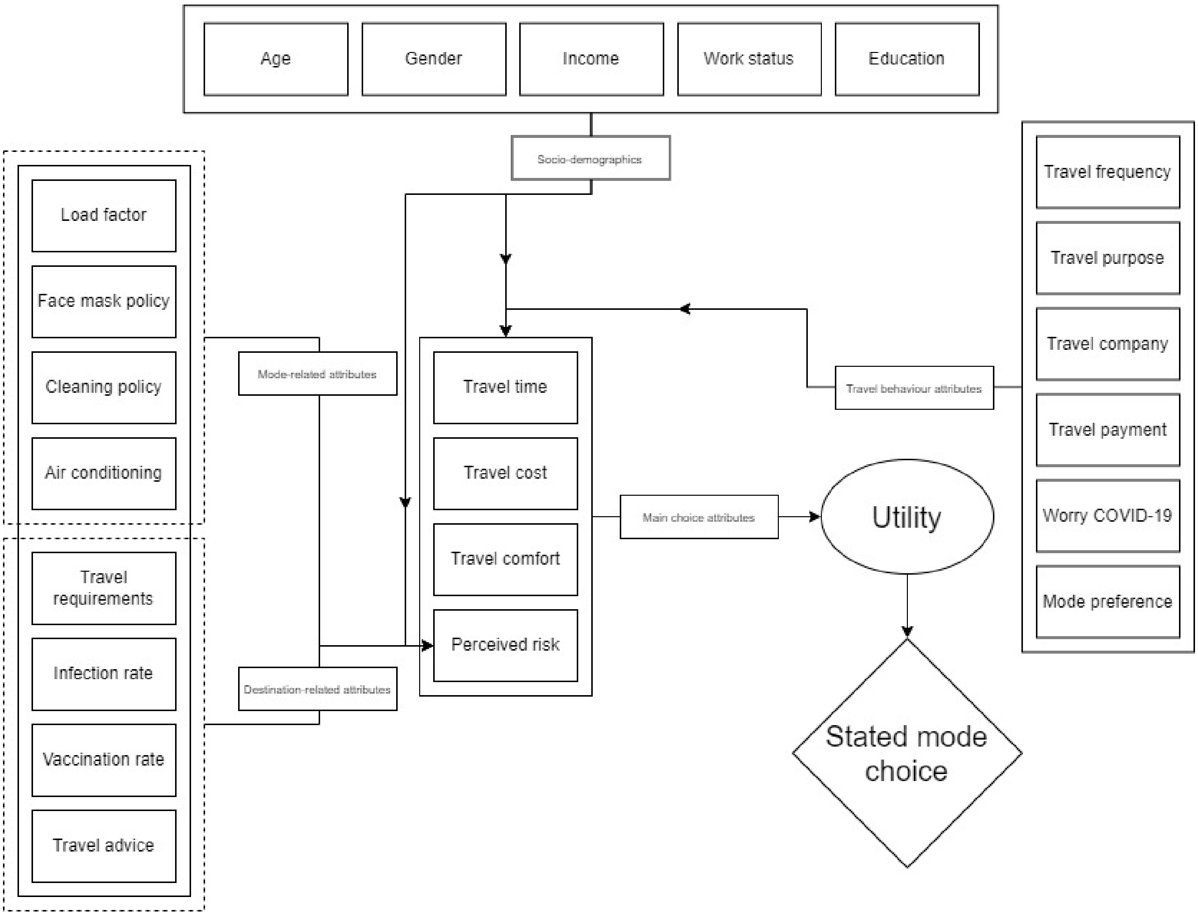


Figure 4.3: Theoretical framework

4.7. Generation of the choice sets

When generating choice sets in order to construct the survey. To create choice sets, the Ngene software is used (the 'so-called' experimental design). Appendix B contains the Ngene code. The Ngene manual contains detailed instructions for defining the syntax, which is used to make the following judgments.

Perceived risk rating experiment

There is no prior information available for the rating experiment. This is required in order to create efficient designs. As a result, a more 'conventional' design is used. A full factorial design is not recommended since it produces an excessive amount of choice sets. To get a sufficient number of responses, the survey should be as brief as feasible. As a result, a fractional factorial design is adopted in this study. An orthogonal design is chosen because it ensures attribute level balance; this means that all attribute levels appear an equal number of times in the choice sets. This results in minor correlations. It is decided to do an unlabeled experiment. The rating experiment will only be conducted for train and plane modes. To simplify, there is no differentiation inside the rating experiment. This suggests that people are either taking the train or flying. The rating experiment makes no difference between these modes. As a consequence, the choice sets may be constructed in sequential order. Ngene can generate an orthogonal design with 20 rows. Respondents will get exhausted if they are asked too many questions. As a consequence, blocking is used. Four blocks are used within the design, and each respondent is given five questions to assess their level of perceived COVID-19 risk. Because some of the attributes are categorical, such as face mask policy and travel advice, dummy coding is used. Effects coding could also be used; the findings would remain the same. The difference is in how the parameters are interpreted. The Ngene design can be found in appendix B.

Main choice experiment

For the main choice experiment, it is chosen to go for an efficient design. An efficient design results in fewer choice sets for the survey than an orthogonal design. In this case, there is prior information available; however, not for all four variables. Therefore a Bayesian D-efficient design is chosen, so the prior values can differ around the mean. The value of time values in the Netherlands are investigated by Netherlands Institute for Transport Policy Analysis (Kim). The value of time (VoT) for car drivers is set at around 8 euros per hour; for train travellers, it is set to around 10 euros per hour (KiM, 2020). Using a β_{cost} of 0.01 and a β_{time} of 0.1 leads to a VoT of 10 euro per hour. For β_{risk} the value of 5 euros per level of decrease in perceived risk is chosen. This is chosen as earlier research of van de Wiel (2021) found this value. For comfort, 50 euros per increase in class is chosen; this is rather an 'educated' estimate. No exact prior values can be found for comfort. As a consequence, this prior value can be anywhere between 25 euros and 75 euros. Ngene finds an efficient design with only ten rows. It is chosen to go for a design with 12 rows as this number can be divided by three. The reason for this is the fact that 12 choice sets give more information. As the main choice is divided into two distance classes, this gives a total of eight main choice questions for the respondents. Both designs can be found in appendix B.

4.8. Survey construction

The construction of the survey is done by using the MWM2 (Crowdtech) survey software.

4.8.1. Survey implementation in software

The designs found by Ngene will be used to construct the actual survey. The values of the variables are varied in the design; however, the representation of such a table is not attractive and easy to understand for respondents. The software that will be used is from MWM2 (Crowdtech). This software contains all the elements to construct the survey. The data can be downloaded as a .CSV file or as a .sav (SPSS) file.

Rating experiment

As earlier discussed, there will be 5 rating questions per respondent. Respondents need to rate their perceived risk of COVID-19 from 1-very low risk to 5-very high risk. An example from the real survey can be found in the figure below.

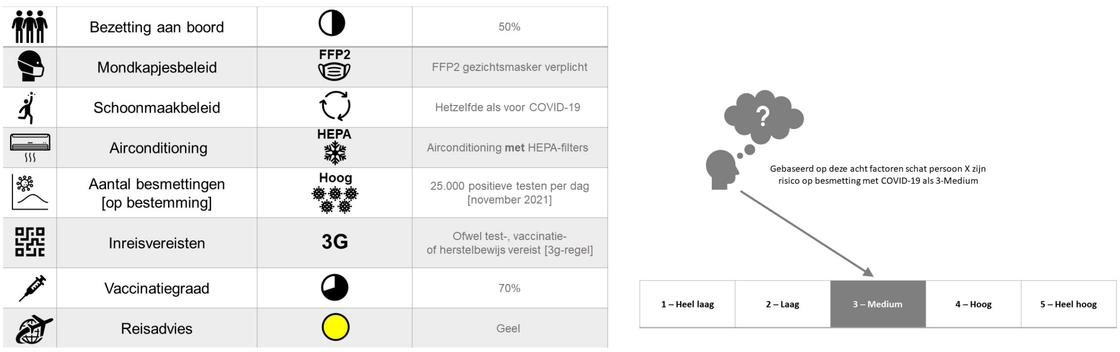


Figure 4.4: First question of rating experiment (block 1)

Main choice experiment

The resulting experimental design found by Ngene is transformed into choice situations for the main mode choice experiment. Respondents will have three alternatives, train, car or plane. The figure below shows an example of the main choice experiment. As there are two distance classes, respondents answer in total eight questions.

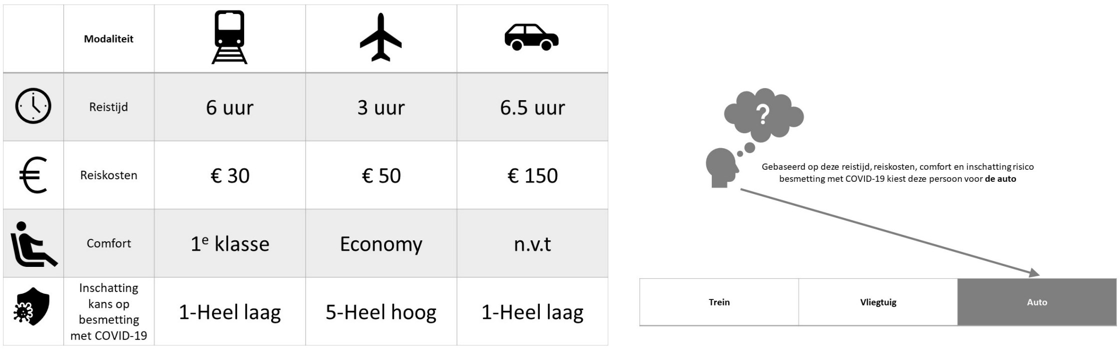


Figure 4.5: Example of main choice experiment question

4.8.2. Testing of survey

The first versions of the survey were tested among family and friends who were not familiar with transportation. This was done because people that are not experts also need to understand what the survey is about. The target group is a representative sample of the population of the Netherlands, so everyone must understand the survey. Furthermore, the committee members of this thesis gave extensive feedback. The most important feedback was: to make use of icons within the rating experiment so it is easy to interpret. Furthermore, the length of the survey was reduced as much as possible. At the same time, information was given if needed. At last, some textual changes were made. It is chosen to only present the survey in Dutch as the target group is the Dutch population.

4.9. Summary

Main points of chapter

- The survey contains two experiments, the rating experiment and the main choice experiment.
- The rating experiment consists of eight attributes, with each four levels.
- The main choice consists of four attributes, with each three levels.
- Each respondent answers five rating questions and eight main choice questions.
- The choice set consists of train, plane and car.
- For the rating part, only train and plane are taken into account, as it is assumed that respondents do not travel with strangers in their car.

This chapter gives details about the process of designing a survey. At first, the context of the survey was discussed. People are travelling to European destinations. A more general approach is chosen rather than a case study for an OD-pair. As a consequence, there is no case study on one OD-pair. To categorise destinations, two distance classes are introduced, the 400-600 kilometre class and the 800-1200 kilometre class.

In addition, the choice experiments were specified. The survey includes two separate experiments: a perceived risk rating and the main choice experiment. Three options were included in the main choice set: train, plane and car. For both the perceived risk rating and the main choice experiment, attribute values were given. For the rating experiment, an orthogonal design was selected; for the main choice experiment, an efficient design was chosen. Moreover, the choices to generate the experimental designs were discussed; Ngene will be used for this. A blocked design reduces the number of questions that a respondent must need to answer. For the rating experiment, four blocks were selected; for the main choice, three blocks were selected. The perceived risk rating experiment consists of five questions, while the main choice experiment consists of eight questions.

5 Data analysis

This chapter elaborates on the characteristics of the sample. In the first section, elements about data collection are discussed. Then, in the second section, the characteristics of the sample are explained. In the third section, the sample is discussed regarding travel behaviour. In section four, a summary is given.

5.1. Data collection

The research question is about how risk perception on COVID-19 contamination influences mode choice in Europe. The research does focus on the Dutch population. Because of this reason, the sample needs to be a representation of the Dutch population. A problem with the collection of own responses is that often certain population groups are overrepresented and other groups underrepresented. The minimum age for the survey is 16 years. Therefore, it needs to be accounted that people under 16 years are not included in the sample (but, of course, this group is part of the population). Another important step for this research is to collect enough responses. Johnson and Orme (2010) suggested a rule of thumb for calculating the minimum necessary number of responses. The formula is presented below:

$$\frac{n \times t \times a}{c} \geq 500 \tag{5.1}$$

Within this equation n means the number of respondents that are required, t is the number of choice tasks, a is the number of alternatives, and c is the highest number of levels used in the choice set. In this case, the minimum number of respondents is necessary to know.

Rating experiment
Substituting the known values for the rating experiment leads to the following equation:

$$n \geq \frac{500 \times c}{a \times t} = \frac{500 \times 4}{1 \times 20} = 100 \tag{5.2}$$

This equation shows that there are at least 100 respondents needed when they are faced with 20 choice tasks. This is too much for respondents. The choice tasks are divided into four blocks; thus, there are $100 \times 4 = 400$ respondents needed for the rating experiment.

Main choice experiment
Substituting the known values for the main choice experiment leads to the following equation:

$$n \geq \frac{500 \times c}{a \times t} = \frac{500 \times 3}{3 \times 12} = 41.67 \tag{5.3}$$

As there are three blocks, the minimum number of required respondents is: $42 \times 3 = 126$. For both main choice parts (400-600km & 800-1200km), an efficient design is used. The Ngene software calculates the so-called 'sp-estimate' for all betas. The sp-estimate is the minimum number required respondents needed. The sp-estimate for the *400-600 kilometer* distance class is 41.51 for β_{pr}. So the minimum number of respondents is 126. This number is (almost) the same as the calculated one from the equation. For the *800-1200 kilometer* distance class is 45.13 also for β_{pr}. So the minimum required number of respondents is 138.

Data was gathered in many methods. The survey was open for a response from the 8^{th} of February to the 8^{th} of March. Cooperation with NS has taken place. This was because of the fact that they were interested in the topic of this thesis. Therefore the survey was also distributed to their panel. In total, the survey was sent to 5000 people. In total, 938 respondents took part in the survey and fully completed the survey. This is a response rate of 18.7%. In addition, the survey link was shared with friends, family, and coworkers, as well as on social media sites such as LinkedIn, Instagram, and the Royal HaskoningDHV C-Infra department. This resulted in a total of 209 completed responses. All in all, the total number of completed responses of this survey was 1147, which is way above the minimum needed respondents. This number surpassed expectations.

5.2. Sample characteristics

Several background and socio-demographic questions are included in the survey. This is done because of two reasons: the first reason is to test whether the sample is representative to the Dutch population, and the second reason is to interact these variables in both the rating and main choice experiment. To check whether the sample is representative, the statistics from CBS are used to compare with the sample characteristics. For age, the second percentage found from CBS is the percentage without the 0-19 age group. This is done, as this gives a more 'fair' comparison, as the minimum age of this survey is 16 years old. It can be seen from table 5.1 that for some variables, the sample is quite representative, while for the other variables, it is not.

Age

Let's start with the variable age. Both the 20-40 and 40-65 values are nearby the values of CBS. For the 0-19 category, the sample rate is way lower, but this is a logical consequence as respondents had to be 16 years or older. The 65-80 and 80+ groups are overrepresented, this is a consequence of using the NS panel. Within this panel, there are more respondents in the higher age groups. Total frequency deviates from the total number of participants as some respondents did not wish to fill in age.

Gender

Gender is more straightforward. The percentages of both groups are very close to the CBS ones. However, in this sample, there are a little bit more males than females. People that did not want to specify or identify themselves as 'other' are not included in comparison with CBS.

Income

The variable income is not very representative in comparison to the Dutch population. It was expected that higher-income classes would be more presented than the lower-income classes. This is because a lot of respondents are from the author's own environment, with often more 'higher' educated people that often do better-earning jobs. However, as also the NS panel was used, it would be expected to be less. But also within the NS panel, there are more higher-income classes than the lower one. In the range from €20.000 to €50.000, the percentages are quite similar. As a consequence of the overrepresentation of the higher income classes, it can be expected that the willingness-to-pay values and Value of Time calculations are possibly too high.

Education

For education, only data is available from the NS panel group. This question was only asked in the NS panel. Again here, education cannot be seen as very representative of the Dutch population. The 'lower' levels of education are always lower in the sample than in the population, and the opposite can be seen for the 'higher' level of education. This was to be expected for the same reason explained with income. Only havo, vwo and hbo-,wo-bachelor are similar in percentages.

To summarise, the representativeness of this sample of Dutch people travelling to European destinations is questionable. This is partly due to the fact that most of the respondents do use the train on a regular basis. For this research, this is not considered a big problem. When considering the findings, there have to be kept in mind that the sample was primarily made up of frequent rail passengers.

Table 5.1: Characteristics of the sample

Socio-demographic variable	Category	Frequency	Percentage sample	Percentage CBS
Age	0-19	27	2.4%	21
	20-40	261	23.1%	25% / 34%[1]
	40-65	418	36.9%	34% / 43%[1]
	65 to 80	371	32.8%	15% / 19%[1]
	80+	55	4.9%	5% / 6%[1]
	Total	1132		
Gender	Female	547	49.0%	50.3%
	Male	569	51.0%	49.7%
	Total	1116		
Income	€10.000	77	8.5%	13.6%
	€10.000-€20.000	85	9.4%	23.3%
	€20.000-€30.000	117	12.9%	18%
	€30.000-€40.000	156	17.2%	14.7%
	€40.000-€50.000	141	15.5%	10.9%
	€50.000-€100.000	247	27.2%	16.5%
	€100.000-€200.000	76	8.4%	2.6%
	€200.000 or more	10	1.1%	0.4%
	Total	909		
Education	Basisonderwijs	7	0.8%	8.3%
	Vmbo-b/k, mbo1	22	2.4%	10.7%
	Vmbo-g/t, vwo-onderbouw	65	7.0%	8.4%
	Mbo2, mbo3 en mbo4	119	12.9%	26.6%
	Havo, vwo	93	10.0%	9.5%
	Hbo-,wo-bachelor	258	27.9%	21.9%
	Hbo-,wo-master, doctor	323	34.9%	13%
	Do not know	39	4.2%	1.7%
	Total	926		

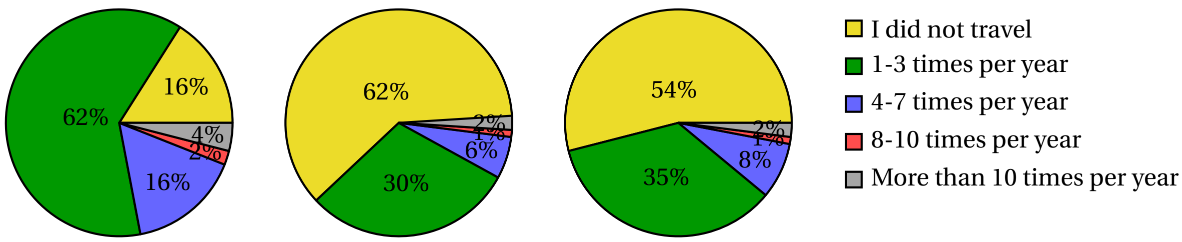
5.3. Travel behaviour of the sample

In the first part of the survey, questions on travel behaviour were included. Respondents were asked how often they travelled to European destinations in 2021. Also, travel purpose was asked, together with travel company and preferences for mode (in both distance classes). This was done to characterise the sample in terms of travel behaviour. These questions will be used to interact with main variables as well. The percentages are shown in pie charts. The first pie chart is always about the own response group, the second one always about the NS panel group and the last is combined.

Travel frequency

In terms of travel frequency, both groups do differ substantially. In the own response group, most of the people did travel at least once in 2021. In the NS panel, more than half of the respondents did not travel at all. This will possibly give another perceived imagination for these respondents as they did not know how travelling during COVID-19 was. Other groups were all quite small, so most of the respondents either travelled only once or not at all.

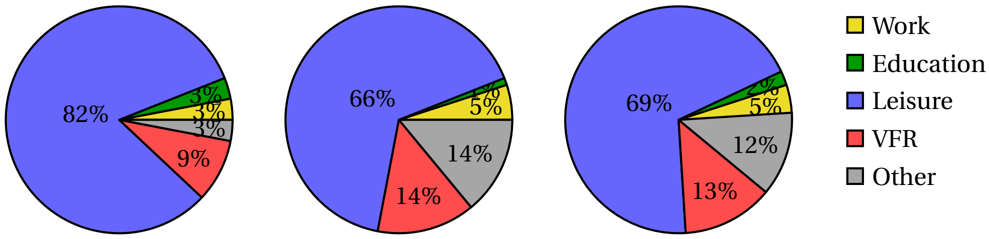
Figure 5.1: Pie chart for own response group, the NS panel and combined for travel frequency



Travel purpose

For both response groups and the combined group, most people travel for leisure. VFR travel is also important in all groups; work is not very often chosen. This can be partly due to the COVID-19 pandemic and the way we also can easily meet digitally. At last, there are a few respondents travelling for school, study or educational institution. There are no worthy of appointment differences between the groups.

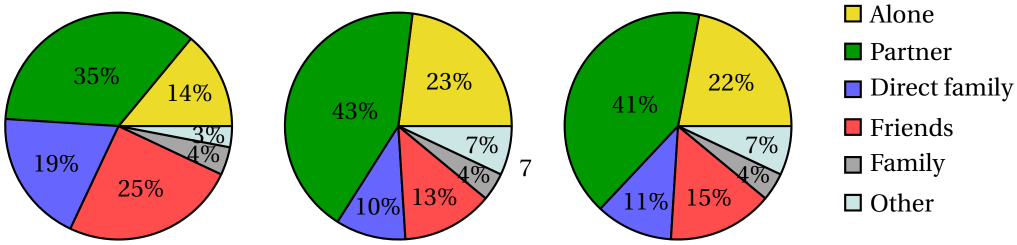
Figure 5.2: Pie chart for own response group, the NS panel and combined for travel frequency



Travel company

In terms of travel company, it can be seen that for the own response group and the NS panel, there are some differences. More people are travelling alone within the NS panel group, but for the own response group travelling with friends is more chosen than with the NS panel. The rest of the outcomes are about the same.

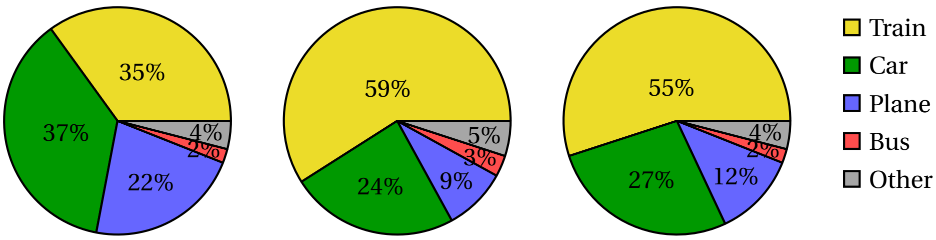
Figure 5.3: Pie chart for own response group, the NS panel and combined for travel company



Preferred mode *400-600 kilometer distance class*

Here may be some bias seen. For the own response group, the preferred mode in most cases is the car, with the train close to it. With the NS panel, more than half of the respondents do favour the train. Especially plane seems to underrepresented in the NS panel group.

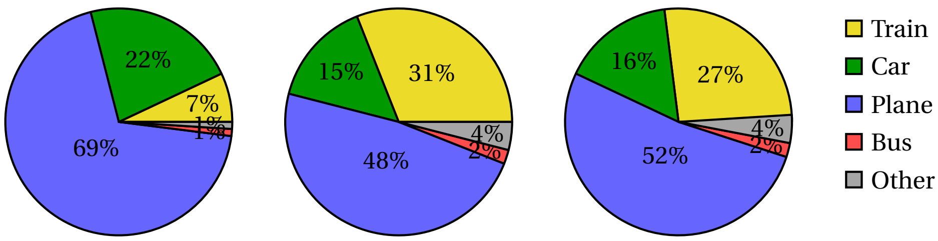
Figure 5.4: Pie chart for own response group, the NS panel and combined for mode preference 400-600km



Preferred mode *800-1200 kilometer distance class*

Again potential bias can be seen, as within the NS panel group, even for this longer distance, a huge part chooses train as their preferred mode. Within the own response group, most respondents chose the plane. This also counts for the NS panel group, but still almost 25% of the respondents less than the own response group.

Figure 5.5: Pie chart for own response group, the NS panel and combined for mode preference 800-1200km



As the last step, the mode choices of the respondents are analysed. In total, there were 24 choice sets, 12 for the 400-600 kilometre class and 12 for the 800-1200 kilometre distance class. For every different choice task, the percentages of each mode being chosen are shown in the figure below. On average the train is chosen 39% of the time (especially in the shorter distance class), car 29% and plane 33%. In every choice set, every mode is chosen at least once. This is a consequence of the efficient design; therefore, dominating alternatives are not in the design. F choice task 8, it can be seen that plane is very little chosen. This is a logical consequence of the values for plane in this choice set. For this case train and car do have the better attribute values, but plane has a 'better' value for comfort, i.e., business class. Therefore, there is no dominance. For choice set 18, plane is often chosen, in choice set 22 train has a high market share. However, in every choice set, every mode is chosen at least once.

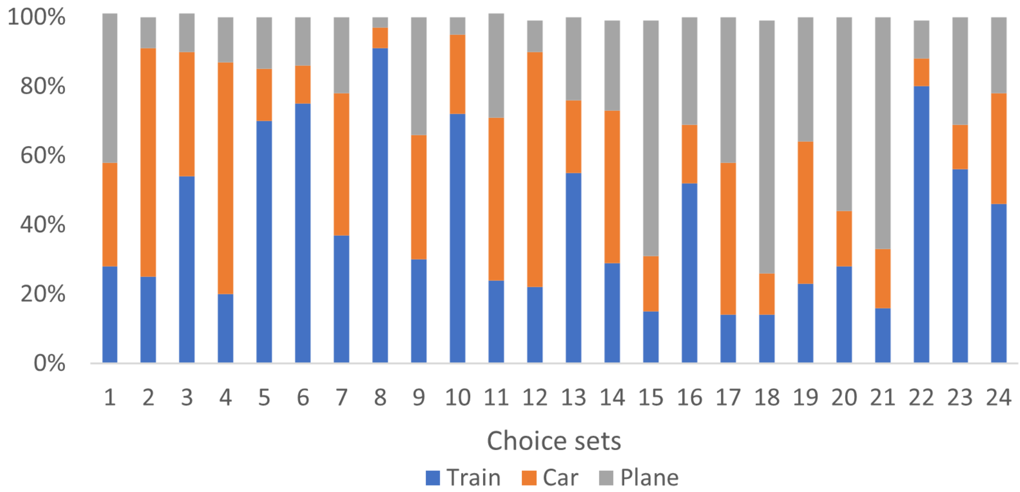


Figure 5.6: Percentages of mode being chosen in each choice task

5.4. Summary

Main points of chapter

- In total 1147 completed responses were collected.
- The sample composes mainly on highly educated people, that earn more than average.
- Age (and gender) are very well represented in the sample in comparison to the Dutch population
- Train is the most chosen mode, with car and plane about the same.

This chapter did provide an overview of the data collection. For this research, in total, 1147 responses were completed. Responses were collected via the NS customer panel, social media and the family, friends and knowns of the author. The sample is partly representative to the Dutch population, as some groups are over-represented and others underrepresented; the consequences should be taken into account when calculating the VoT and willingness-to-pay values.

6

Model estimation results

This chapter presents the outcomes of the models used for both the rating and main choice experiment. In the first section, the (linear) regression model is discussed that is used for the rating experiment. The second section discusses the MNL model that is used for the main choice experiment, and elaborates on the combination of the rating and main choice experiment. In the third section, a summary of this chapter is given.

6.1. Perceived risk rating experiment: Regression

In order to study the effects of perceived risk on COVID-19 on mode choice for long-distance travel, the perceived risk variable is introduced. Perceived risk is both the dependent and the independent variable; in the rating experiment, it is the dependent variable, while in the main choice experiment it is the independent variable. Perceived risk consists of eight factors; these factors are load factor, face mask policy, cleaning policy, air conditioning, travel requirements, infection rate, vaccination rate and travel advice.

6.1.1. Model estimation

To study how the perceived risk attributes contribute to perceived risk, a regression analysis is done. Respondents rated their perceived risk of COVID-19 based on a Likert scale. A Likert scale is a method for interrogating data that is difficult to quantify and providing it with an ordinal level of measurement (Joshi et al., 2015). Therefore it is widely used in questionnaires and surveys. It was chosen to go for a five-point scale, as this is easy for respondents to rate. The rating number (the target variable) is considered a continuous variable, and is also estimated as a continuous variable. The data is analysed using IBM SPSS statistics 26.0.

Six of the eight factors are ordinal scale type; the other two are ratio scales. The ordinal levels need to be coded into dummy or effects coded variables. This is due to the fact that every different level of an ordinal variable has a different contribution. For example, the difference between *no mask mandatory* and *any mask mandatory* are presumably different from the difference between *any mask mandatory* and *surgical mask mandatory*. The number of positive tests is also dummy coded, this is because the levels represent certain moments during the pandemic; therefore, the differences between levels are not always the same size. Either dummy or effect coding can be used; in this case, dummy coding is used. As every level contributes to a better (or worse) level of security, dummy coding gives an easy interpretation. The first level is the reference level. Effects coding could also be used, but in this case, dummy coding gives an easier interpretation.

Table 6.1: Dummy coded rating variables

Variable	Level	Coding		
Face mask policy		Any	Surgical	FFP2
	No face mask mandatory	0	0	0
	Face mask mandatory (any)	1	0	0
	Surgical mask mandatory	0	1	0
	FFP2 mask mandatory	0	0	1
Cleaning policy		Increased	Weekly	Daily
	Same cleaning policy as before COVID-19	0	0	0
	Increased cleaning policy (touch points)	1	0	0
	Weekly disinfection whole vehicle	0	1	0
	Daily disinfection whole vehicle	0	0	1
Ventilation/air conditioning		Ventilation	Airco no HEPA	Airco with HEPA
	No ventilation and air conditioning	0	0	0
	Only ventilation	1	0	0
	Air conditioning without HEPA filters	0	1	0
	Air conditioning with HEPA filters	0	0	1
Vaccination/recovery/testing requirements		3G	2G	2G+
	No mandatory requirements	0	0	0
	Either testing, vaccination or recovery required (3g-rule)	1	0	0
	Only vaccination or recovery (2g-rule)	0	1	0
	Always testing, and vaccination or recovery (2G+-rule)	0	0	1
Infection rate		10.000	25.000	100.000
	100 positive tests per day (summer 2020/June 2021)	0	0	0
	10.000 positive tests per day (October/November 2020/July 2021)	1	0	0
	25.000 positive tests per day (November 2021)	0	1	0
	100.000 positive tests per day (fictitious extreme number)	0	0	1
Travel advice		Yellow	Orange	Red
	Green advice	0	0	0
	Yellow advice	1	0	0
	Orange advice	0	1	0
	Red advice	0	0	1
On-board crowding/load factor		Ratio		
	25 % of seats occupied			
	50 % of seats occupied			
	75% of seats occupied			
	100 % of seats occupied			
Vaccination rate		Ratio		
	15 % fully vaccinated people (Bulgaria)			
	30 % fully vaccinated people (Romania)			
	70 % fully vaccinated people (Netherlands and EU average)			
	90 % fully vaccinated people (Portugal)			

To answer the second research question, the model includes socio-demographic factors. Therefore, socio-demographic variables are included in the regression. Background questions are not included as these are only important for the main choice experiment. Most of the socio-demographic variables are ordinal scale. The questions in the survey are very detailed. Including all levels as dummy variables would lead to a substantial increase in parameters. Therefore some of the levels are combined. The socio-demographic variables (and the dummy variables) are shown in table 6.1. For every variable, it is discussed how these levels are combined:

- **Gender**: Male is the reference level, female and other are level one. Only seven respondents out of 1139 (valid) responses identified themselves as other. Therefore the impact is very low; as a consequence, it is combined with women.

- **Age**: For age, the same categories from CBS are used: younger than 20 years, 20-40 years, 40-65 years, 65-80 years and older than 80 years. Two levels are combined, 65-80 and older than 80 years. The group older than 80 was very small and is therefore combined with 65-80 years.

- **(Gross) Income**: Again, the same categories are used as CBS: less than €10.000, €20.000-€30.000, €30.000-€40.000, €40.000-€50.000, €50.000-€100.000, €100.000-€200.000 and €200.000 and more. The categories are combined to less than €20.000, €20.000-€40.000, €40.000-€100.000 and more than €100.000. The groups respond to low income, around 'modaal' income, above 'modaal' income and very high income. Also, a substantial part of the respondents responded with 'I do not want to tell'. Therefore, this level is also included as a dummy as there is no logical way of combining this level, but this level gives not any information.

- **Education**: Very detailed levels are used: primary school, LBO, MAVO, VMBO, MBO 'oude stijl', Mbo1, Mbo2/Mbo3/Mbo4, MULO, HAVO, VWO, HBO (bachelor), HBO (master), WO (bachelor), WO (master), WO doctor, Other. The HBO and WO levels are combined to HBO and WO. All the other levels are combined to MBO or lower. The reason for this is HBO, WO and MBO are the three primary levels of education in the Netherlands.

- **Work status**: The levels are: school, student, salaried employment, working for the government, entrepreneur, freelance, volunteer, unemployed, housewife/houseman, retired, incapacitated, other, and I do not want to tell. The levels are combined to not working, student & school, retired and working. There were only three respondents that responded with 'I do not want to tell' and only one with 'other'. These are combined with not working for the sake of simplicity and minimal impact.

Table 6.2: Socio-demographic variables dummy coded

Socio-Demographics		Dummy coded			
Gender		Female			
	Male	0			
	Female + I do not want to tell	1			
Age		20-40	40-65	65-80	
	20	0	0	0	
	20-40	1	0	0	
	40-65	0	1	0	
	65-80	0	0	1	
Income		€20.000-€40.000	€40.000-€100.000	€100.000	I do not want to tell
	€20.000	0	0	0	0
	€20.000-€40.000	1	0	0	0
	€40.000-€100.000	0	1	0	0
	€100.000	0	0	1	0
	I do not want to tell	0	0	0	1
Education		HBO	WO	Other	
	MBO or lower	0	0	0	
	HBO	1	0	0	
	WO	0	1	0	
	Other	0	0	1	
Work status		Student	Retired	Working	
	Not working	0	0	0	
	Student & school	1	0	0	
	Retired	0	1	0	
	Working	0	0	1	

6.1.2. Expectations contribution attributes

Every factor contributes differently to the value of the rating. Every level of the factors is an increase (or decrease) in the level of protection against COVID-19. **Face mask policy, cleaning policy, ventilation/air conditioning, travel requirements and vaccination rate**: are expected to contribute to a decrease in perceived risk. Therefore the attributes are expected to have a negative sign. It is expected that the different levels do have different contributions (so different values). **Infection rate, travel advice and load factor** are expected to have a positive sign; these factors contribute to an increase in perceived risk. Every level leads to a higher risk, so as a consequence are also expected to contribute to perceived risk.

Including socio-demographic variables does help to investigate if certain groups in society rate perceived risk differently. Commodari (2017) studied the influence of socio-demographic and psychological variables on perceived risk of influenza. The study found that age and education significantly influenced perceived risk. Older respondents did perceive risk as higher, and people with a 'lower' level of education perceived their risk as higher as well. Therefore it is expected that an increase in **age** level does contribute to a higher level of perceived risk. For **education** the opposite is expected, so a higher level of education contributes to a lower level of risk. Risk in general is perceived higher by women than by men (Finucane et al., 2000). Therefore it is expected for **gender** that women perceive risk as higher than men. Hence a positive contribution is expected. For **work status and income** no clear expectation can be found in literature. Therefore no expectations about these variables are made.

6.1.3. Results linear regression

The linear regression findings are reviewed in this section. At first, only the main attributes of the rating experiment are estimated. After this, also socio-demographic variables are included. In table 6.3 all included variables are shown. All of the parameters are significant on the 5% significance level. In table 6.4 the insignificant parameters are shown. The *stepwise* method was used. With the process, insignificant parameters have been removed stepwise, and then the model is re-estimated. This procedure is done until all parameters are statistically significant.

Table 6.3: Included variables linear regression

Model	Main effects			Main & Socio-demographics		
Main attributes	Value	t	p-value	Value	t	p-value
(Constant)	3.039	52.919	0.000	2.815	40.913	0.000
ADVICE3	0.698	15.412	0.000	0.702	15.620	0.000
VACC	-0.006	-12.866	0.000	-0.007	-13.203	0.000
CROW	0.006	11.306	0.000	0.006	11.604	0.000
INFECT1	0.253	5.917	0.000	0.252	5.953	0.000
INFECT2	0.136	2.134	0.033	0.123	1.954	0.051
INFECT3	0.494	12.575	0.000	0.495	12.692	0.000
AIRCO1	-0.388	-5.411	0.000	-0.405	-5.684	0.000
AIRCO2	-0.193	-4.467	0.000	-0.198	-4.617	0.000
AIRCO3	-0.286	-7.419	0.000	-0.289	-7.560	0.000
REQUIRE1	0.138	2.627	0.009	0.144	2.757	0.006
REQUIRE3	-0.267	-7.424	0.000	-0.272	-7.608	0.000
MASK1	-0.261	-5.397	0.000	-0.268	-5.576	0.000
MASK3	-0.137	-3.889	0.000	-0.140	-4.005	0.000
Socio-demographic attributes				Value	t	p-value
GENDER				0.088	3.283	0.001
HBO				0.137	4.191	0.000
WO				0.127	3.783	0.000
INCOME_20_40				0.112	2.634	0.008
INCOME_40_100				0.168	4.005	0.000
AGE_20_40				-0.095	-2.915	0.004
DO_NOT_SAY_INCOME				0.259	5.473	0.000
R^2 = 0.129				R^2 = 0.143		
Adjusted R^2 = 0.127				Adjusted R^2 = 0.140		

Table 6.4: Variables not included in the model

Model	Main effects			Main & Socio-demographics		
Main attributes	Value	t	p-value	Value	t	p-value
MASK2	-0.036	-1.536	0.125	-0.029	-1.240	0.215
CLEAN1	-0.004	-0.304	0.761	-0.002	-0.132	0.895
CLEAN2	-0.023	-1.074	0.283	-0.020	-0.920	0.358
CLEAN3	0.003	0.216	0.829	0.002	0.139	0.890
REQUIRE2	0.022	0.791	0.429	0.019	0.674	0.500
ADVICE1	-0.003	-0.126	0.900	-0.001	-0.056	0.955
ADVICE2	-0.021	-0.941	0.347	-0.020	-0.886	0.376
Socio-demographic attributes				Value	t	p-value
AGE_40_65				-0.022	-1.581	0.114
AGE_65_AND_OLDER				0.022	1.530	0.126
INCOME_MORE_100				0.025	1.374	0.170
NOT_WORKING				-0.006	-0.382	0.702
STUDENT_SCHOOL				0.008	0.584	0.559
RETIRED				-0.009	-0.709	0.479
OTHER_EDU				0.012	0.801	0.423

Given the R^2 for both models is compared, it can be concluded that the model including socio-demographics outperforms the model with just main variables. The model with socio-demographics explains 14.3% of the variance of the dependent variable. This value is not very high, but this could be explained by the following. In society it was seen that a lot of people had different opinions on COVID-19 (Milosh et al., 2020). Moreover, this could be from polarisation around opinions on COVID-19 (Arnold-Forster, 2021). Linear regression is used for the estimation of the model. However, some may believe that this scale is ordinal, necessitating the estimation of an ordered logit model. Following this assumption, the perceived risk rating attribute should be considered ordinal. As a result, this attribute must be added with a set of dummy variables in the choice model; then, interpolation for undecided intermediate values would not be allowed. Therefore it is assumed that the five-point perceived risk scale is of interval measurement level. This allows the estimation of linear parameters for the perceived risk attribute (Molin et al., 2017). The following equation is the result of the results of the linear regression model:

$$
\begin{gathered}
PR_{COVID-19} = C + \beta_{RA} * RA + \beta_{VR} * VR + \beta_{LF} * LF \\
+\beta_{INFECT1} * INFECT1 + \beta_{INFECT2} * INFECT2 \\
+\beta_{INFECT3} * INFECT3 + \beta_{VL} * VL + \beta_{NH} * NH \\
+\beta_{HP} * HP + \beta_{3G} * 3G + \beta_{2G+} * 2G + +\beta_{AM} * AM \\
+\beta_{FFP2} * FFP2 + \beta_{GEN} * GEN + \beta_{HBO} * HBO \\
+\beta_{WO} * WO + \beta_{INC_{20-40}} * income_{20-40}
\end{gathered}
$$

C = constant, RA = Red travel advice, VR = vaccination rate, LF = load factor, INFECT1= 10.000 infections. INFECT2 = 20.000 infections, INFECT3 = 100.00 infections, VL = ventilation only, NH = air conditioning no HEPA filter, HP = air conditioning with HEPA filter, 3G = 3G policy, 2G+ = 2G+ policy, AM = any mask, FFP2 = FFP2 mask mandatory, GEN = gender, HBO = HBO education level, WO = university education level, INC_{20-40} = income between €20.000 and €40.000.

6.1.4. Interpretation of parameters

The results of linear regression give several implications. In total, eight main attributes are included in the regression model, two of them are ratio scales and six of them are ordinal scales. All the ordinal scale variables are dummy coded. In total, there were 20 main parameters estimated. Also, five socio-demographic attributes are included that are also dummy coded, ensuring a total of 14 parameters. For the main attributes, all parameters do have the expected sign, except for REQUIRE1, which means 3G policy (i.e., either testing, recovery or vaccination proof needed to travel). In this case, a negative sign would be expected as this policy decreases the probability of someone infected when travelling (in comparison to the base level, with no travel requirements). From this result, one could conclude that the respondents do not believe that the 3G policy helps in decreasing their perception of risk.

All other main parameters do have the expected sign, and there are also insignificant main parameters. The constant is 2.8; this is the value if all parameters are set to the base level. So when all parameter are set to their base level, respondents rate perceived risk at 2.8 (so that is around the mean value of 3). This means that respondents, on average, rated their perceived risk when base attributes are considered as a little under *medium*. Are significant parameters are highly significant, except for INFECT2 (i.e., 25.000 infections per day). For every attribute, it is discussed how to interpret the parameters.

Main parameters

- **Travel advice**: This variable is dummy coded, ADVICE1 means yellow advice, ADVICE2 means orange advice, and ADVICE3 means red advice. Both the yellow and orange advice parameters are not significant, so these levels do turn not out to be different from the base level green advice. However, red travel advice has the largest positive effect on perceived risk, with a value of 0.702.

- **Vaccination rate**: This is a ratio variable with a contribution of -0.007 for every percentage point increase in vaccination rate in the country of destination. For example a vaccination rate of 50% gives the following parameter: $50 \times -0.007 = -0.35$. When travelling to a country with a vaccination rate of 90% (Portugal), perceived risk is decreased with -0.63 rating points.

- **Load factor**: This is also a ratio variable, with (almost) the same but opposite contribution of vaccination rate. The value of this parameter is 0.006. A load factor of 75% would lead to $70 \times 0.006 = 0.42$ increase in rating points. A load factor of 100% would lead to an increase of 0.6 points on perceived risk.
- **Infections**: Dummy coded variable, with INFECT1 meaning 10.000 positive tests per day, INFECT2 25.000 positive tests per day and INFECT3 being 100.000 positive tests per day. There is some counter-intuitive outcome, as 10.000 positive tests per day contribute more (with a value of 0.252) to perceived risk than 25.000 positive tests per day (with a value of 0.123). A reason could be that respondents find it hard to imagine what the difference in levels means. INFECT2 is also just significant (or just insignificant) on the 5% level; thus, it is not very significant (p-value of 0.051). The highest level, i.e., 100.000 positive tests per day, has the highest contribution of the dummy variables. This is in line with expectations. It also has the second-highest contribution of the dummy coded attributes, with a value of 0.495.
- **Ventilation/air conditioning**: All dummy variables turn out to be significant. AIRCO1 (only ventilation) has the highest contribution to the decrease in perceived risk with a value of -0.405. This is in line with expectations as there was a huge focus from society on ventilation. Therefore it could be expected that people do think this is important. AIRCO2 (air conditioning without HEPA filter) has the lowest contribution of the dummies (with a value of -0.198); again, this could be expected, as air conditioning without HEPA filters has a lower level of protection against viruses than air conditioning with HEPA filters. AIRCO3 (air conditioning with HEPA filter) has a higher contribution than the previous level. The value of air conditioning with HEPA filters is -0.289.
- **Travel requirements**: As explained earlier, the first dummy variable REQUIRE1 (3G policy), has a positive sign with the value of 0.144, which is not in line with expectations. A possible explanation for this could be the fact that when first introducing the 3G policy last summer 2021, there was an exponential increase in infections. People could therefore believe that this policy is not working in order to reduce the infections and, as a consequence, contributes to an increase in perceived risk. REQUIRE2 (2G policy) is not significant. This can be explained as the 2G policy was never introduced, and there was a lot of resistance. Also, the effectiveness of both the 3G and 2G has been questioned and is reduced (Mouter et al., 2021). 2G+ is an extra level of security in comparison to 2G, with people also needing to test even with a vaccination or recovery proof, it turns out to be significant. This level shows a negative contribution to perceived risk, with a value of -0.272.
- **Face mask policy**: MASK1 (any face mask required) and MASK3 (at least FFP2 face mask required) are significant. The level 'any face mask required' has a higher contribution (value of -0.268) than 'at least FFP2' (value of -0.140). So the need to put on any face mask is more important to reduce perceived risk than having at least an FFP2 mask, according to the respondents. A reason for FFP2 being of less importance than any face mask can be partly due to the ignorance about the difference in levels of safety of the different types of masks. MASK2 (at least a surgical type II mask required) did not turn out to be significant. This is probably because this type of mask got attention very late (around December 2021/January 2022).
- **Cleaning policy**: None of the dummy variables turned out to be significant. This shows that there is no difference from the base level 'same cleaning policy as before COVID-19' and therefore does not contribute in reducing perceived risk.

Socio-demographic attributes

- **Gender**: As in line with the expectations, gender turns out to be significant with a value of 0.088. This means that being women or 'other' other increases perceived risk with 0.088 rating points. This is not a very high value in comparison to other attributes, but the value is significant, so there is a difference between men and women.
- **Education**: The level HBO and WO are both significant and positive. This means in comparison to the base (MBO or lower), people with education HBO and WO perceive the risk of COVID-19 as higher than

people with MBO or lower education level. The value for HBO is 0.137 and for WO 0.127, so people with HBO perceive risk as a bit higher than WO. The group OTHER_EDU is not significant, but does not give information anyway.

- **Income**: For this attribute, there were no expectations. The levels 'income between €20.000 and €40.000' and 'income between €40.000 and €100.000' and 'I do not want to tell' are significant. The first level has a value of 0.112, the second level 0.168 and the last 0.259. Having an income between €20.000 and €100.000 contributes to a higher perceived risk. However, the level 'income of more than €100.000' is not significant. The level 'I do not want to tell' is significant; however, it gives no information. The reason that it is included, as it does not make sense to combine this level with other levels.
- **Age**: It is expected that a higher age contributes to a higher perceived risk. This cannot be concluded from the results. The age group '20-40' years contributes to a lower perceived risk (in comparison to the base 'younger than 20 years). The other two dummy variables did not turn out to be significant. So the age groups '40-65 years' and '65 years and older' do not contribute to an increase or decrease in perceived risk.
- **Work status**: All of the parameters are highly insignificant. So work status does not influence perceived risk.

6.1.5. Examples of combination of attributes

A few examples of combinations of levels are discussed to show what the result of the rating is according to different combinations of the variables. These combinations are labelled to show potential real-world situations. As can be seen, perceived risk does not differ substantially from the mean level 3-medium risk.

Woman of age 35 is travelling on a busy Thalys in November 2021 to France

The levels are, in this case: any face mask required, air conditioning with HEPA filters, 3G policy, 25.000 positive tests per day, 75% occupancy and a vaccination rate of 70%. Gender is woman, age is 35 (so age group 20 to 40 years), study HBO and income between €20.000 and €40.000. For the sake of simplicity, all zeros are kept out of the formula.

$PR_{COVID-19} \approx 2.7$

Man of 60 on a full flight to Bulgaria in July 2021

The levels are any face mask required, air conditioning with HEPA filters, no travel requirements, 10.000 positive tests per day, 100% occupancy and vaccination rate of 15%, gender is man, age is 60, MBO or lower and income of €40.000-€100.000.

$PR_{COVID-19} \approx 3.1$

Women of 19 on a half occupied ICE train to Frankfurt in winter 2021

The levels are: at least FFP2, air conditioning with HEPA filters, 2G+ policy, 25.000 positive tests, 50% occupancy and vaccination rate of 75%, gender is women, age is 19, WO education level.

$PR_{COVID-19} \approx 2.4$

Table 6.5: Attribute level of examples

	Travel advice	Vaccination rate	Load factor	Infections	Ventilation/airconditioning	Requirements	Face mask policy	Gender	Education	Income	Age
Thalys to Paris	Yellow	70%	75%	25.000	HEPA	3G	Any	Woman	HBO	€20.000 - €40.000	35
Flight to Bulgaria	Yellow	15%	100%	10.000	HEPA	None	Any	Man	MBO	€40.000 - €100.000	60
ICE to Frankfurt	Yellow	75%	50%	25.000	HEPA	2G+	FFP2	Woman	WO	€20.000	19

Table 6.6: Values attributes of examples

	Travel advice	Vaccination rate	Load factor	Infections	Ventilation/air conditioning	Requirements	Face mask policy	Gender	Education	Income	Age	Rating points
Thalys to Paris	0	0.49	0.45	0.123	-0.289	0.144	-0.268	0.088	0.137	0.112	-0.095	2.727
Flight to Bulgaria	0	-0.105	0.6	10.000	-0.289	0	-0.268	0	0	0.168	0	3.173
ICE to Frankfurt	0	-0.525	0.3	25.000	-0.289	-0.272	-0.14	0.088	0.127	0	0	2.356

6.2. Main (mode) choice experiment: Discrete choice modelling theory

To study mode choice for long-distance travel in Europe, an MNL model is estimated. Based on four attributes travel time, travel cost, travel comfort and perceived risk, respondents made a choice between train, car and plane. First, a base model is estimated. Then, socio-demographic and travel behaviour interactions are added to the model.

6.2.1. Estimation of the model

The data that is collected is used to estimate a (discrete) choice model. In order to get to the final model, several separate models are estimated. To answer the remaining sub research questions, it is important to study which parameters are significant and what the values are. In order to include context, socio-demographic and travel behaviour interactions are included in the model. To estimate most of the socio-demographic and travel behaviour attributes, the data needs to be prepared. All nominal and ordinal variables are dummy coded, as is done in the rating experiment. Part of the variables is already dummy coded for the rating experiment. The preparation of the data is done in Microsoft Excel. To estimate the choice models, PandasBiogeme (in Python) is used (Bierlaire, 2020).

6.2.2. Expectations

In this section, the expectation of the contribution of parameters are discussed. Main attributes, socio-demographic attributes and travel behaviour attributes is discussed.

Main attributes

The four main attributes are expected to have the following signs. *Travel time, travel cost & perceived risk* are expected to have a negative sign. This based on intuition and previous main choice studies (van Goeverden, 2009; Román et al., 2010). *Travel comfort* is expected to have a positive sign, as an increase in class increases the comfort of travelling, thus utility. It is expected that travel time contributes to perceived risk; hence the interaction between these attributes should be significant. The levels of the main attributes can be found in table 4.2, 4.3, 4.4 & 4.5.

Socio-demographic interactions

For the included socio-demographic attributes, the expectations of the interaction with the main attributes is discussed. The levels can be found in table 6.2.

- **Gender**: This attribute is interacted with *perceived risk*, *travel time* and *travel cost*. As earlier explained in the rating experiment (regression analysis), *perceived risk* is higher for women than for men (Finucane et al., 2000). This result was also confirmed by the regression analysis showing a positive parameter for gender, meaning that women perceive COVID-19 risk as higher than men. For cost and time, there is no expectation of the effect.

- **Education**: This attribute is interacted with *perceived risk*. It is expected that 'higher' educated people perceive risk lower, but the opposite is found in the regression analysis. As a consequence, it is expected that 'higher' educated people perceive risk as higher.

- **Income**: This attribute is interacted with *travel cost*. Higher income is expected to contribute to less weight to the cost parameter (Ohnmacht and Scherer, 2010). For perceived risk, no expectation is there.

- **Age**: This attribute is interacted with *travel cost*, *travel time*, *travel comfort* and *perceived risk*. It is expected that a higher age contributes to a higher perceived risk; however, from the regression analysis, this could not be concluded. Thus, it is expected that age does not have a relation to perceived risk. For cost, it is expected that an increase in age contributes to a lower weight, the same counts for travel time. For comfort, it is expected that older people have a higher weight to this attribute (Ohnmacht and Scherer, 2010).

- **Work status**: This attribute is not expected to have any effect on the main attributes, as, in the regression analysis, there were no significant attributes found.

Travel behaviour attributes

For the included travel behaviour questions (attributes), the expectations of the interaction with main attributes is discussed. All of the attributes are dummy coded; see table 6.7.

- **Worry COVID-19**: This attribute is about the fact if people worry more or less about the omicron-variant in comparison to the delta-variant. This attribute is interacted with *perceived risk.* It is expected that respondents that worry less have a lower weight to perceived risk.
- **Payment for trip**: This attribute is about the payment of the trip. This attribute is interacted with *travel cost.* It is expected that people are less sensitive to the cost when respondents get paid by school or work or someone else (in comparison with the base level' payment by myself').
- **Travel company**: This attribute is about with whom people are travelling mainly. This attribute is interacted with *travel cost* and *perceived risk.* There is no expectation about the interaction with cost. For perceived risk, it is expected that respondents weigh perceived risk higher if they are travelling with other people. This is in line with Karlsson et al. (2021).
- **Travel purpose**: This question was about the main reason for travelling to destinations in Europe. This attribute is interacted with *travel cost* and *travel time.* It is expected that the reason work and education give both a lower weight to cost and time (Ohnmacht and Scherer, 2010).
- **Travel frequency**: This attribute is about how often respondents are travelling. This is interacted with *perceived risk.* There is no expectation for this interaction.
- **Mode preference**: This attribute is about the mode-preference of the respondent for both the 400-600km and 800-1200km distance class. It is interacted with the *ASC* of a mode. It is expected that respondents with a preference for a certain mode have a positive sign for the ASC of the same model and a negative sign if the mode preference is not the same as the ASC of this mode.

Table 6.7: Travel behaviour attributes with dummy coding

Travel behaviour attributes		Dummy coded variables			
Worry COVID-19		Same	More		
	Less	0	0		
	Same	1	0		
	More	0	1		
Payment trip		Someone else	Work/education		
	Myself	0	0		
	Someone else	1	0		
	Work, education pays	0	1		
Travel company		With friends	With friends	With family	Other
	Alone	0	0	0	0
	With partner	1	0	0	0
	With family/own household	0	1	0	0
	With friends	0	0	1	0
	Other	0	0	0	1
Travel purpose		VFR	School	Work	Other
	Leisure	0	0	0	0
	VFR	1	0	0	0
	School	0	1	0	0
	Work	0	0	1	0
	Other	0	0	0	1
Travel frequency		1-3 per year	4-7 per year	More than 7 times a year	
	No travel	0	0	0	
	1-3 per year	1	0	0	
	4-7 per year	0	1	0	
	More than 7 times a year	0	0	1	
Mode preference 400-600km		Train	Plane	Car	
	Bus + other	0	0	0	
	Train	1	0	0	
	Plane	0	1	0	
	Car	0	0	1	
Mode preference 800-1200km		Train	Plane	Car	
	Bus + other	0	0	0	
	Train	1	0	0	
	Plane	0	1	0	
	Car	0	0	1	

6.2.3. Including socio-demographic & parameter interactions

This section is about the steps that are taken in order to include socio-demographic interactions and travel behaviour questions. Together this leads to the final MNL model. The model with only main variables acts as the base model. There are several different MNL models estimated, with the base model always as a starting point. To the base model, socio-demographics and travel behaviour interactions are added. The following approach was taken. In section 6.2.2 the expectations of the interactions are discussed. Then, one by one, different interactions are added to the model. For every difference (with only one extra interaction), it is examined if the parameters are statistically significant. The 5% significance level is used. The PandasBiogeme codes and results of the estimation reports can be found in Appendix D.

Socio-demographic interactions

This section discusses the steps of including socio-demographic interactions. An explanation with the Log-likelihood and ρ^2 values of every different model, including the values of all parameters and interactions, can be found in Appendix C.

- **Gender**: The interaction between gender and cost is not significant; hence there is no relationship between gender and cost. For gender and time, the same can be found, with a highly insignificant interaction between gender and time. The opposite can be found for gender and perceived risk, with a highly significant parameter between perceived risk and gender. The negative sign is also in line with the literature and the regression model. Women have a higher weight.
- **Education**: Three extra parameters are estimated for the interaction between education and perceived risk; this is due to the dummy coding. All three parameters are significant. The negative sign is in line with the negative with the regression model, with HBO and WO (in comparison to MBO or lower) having positive contributions to perceived risk. In this model, there is a higher weight on perceived risk.
- **Income**: Four extra parameters are estimated to test the interaction between income and cost. All parameters are significant. All of the levels are significant with a positive sign, which means that for these levels, there is given less weight to cost. Every higher level in age contributes to a less negative weight on cost. The level 'do not say' is also significant, but it gives no information anyway.
- **Age**: Three extra parameters are estimated for the interaction with cost. All parameters are significant. In comparison to the base' younger than 20 years, all parameters have a positive sign; hence there is a lower weight to cost. This is in line with reasoning as 'older people are in general less sensitive to cost than younger people. The interaction between age and perceived risk also gives three extra parameters. None of the parameters turns out to be significant. So there is no relation between age and perceived risk. The last interaction between age and travel comfort gives three extra parameters as well. None of the parameters turns out to be significant; there is no relationship between age and travel comfort.
- **Work status**: Four extra parameters are estimated for the interaction between work status and cost. The levels' student' and 'not working' in comparison to the base level 'working' have significant parameters. Student shows a counter-intuitive sign, being positive (so less weight to cost) than working people. However, people that do not work have a higher weight to cost (which is in line with expectations).

Travel behaviour interactions

This section discusses the steps of including travel behaviour interactions. An explanation with the Log-likelihood and ρ^2 values of every different model, including the values of all parameters and interactions, can be found in Appendix D.

- **Worry COVID-19**: Two extra parameters are estimated in order to test the interaction between the worry of the omicron-variant in comparison to the delta-variant (base level). Both parameters are insignificant, so no relation can be found.
- **Payment**: Three extra parameters are estimated to test the interaction between payment (who pays for the trip) and cost. Only' work or education pays' is significant with a positive sign; hence a lower weight to cost can be found for this attribute in comparison to 'payment by yourself' (base level). The other parameters are insignificant.
- **Travel company**: Four extra parameters are estimated for the interaction between travel company and cost. All parameters are significant except for 'other', but this gives no information anyway. All parameters are negative, so they contribute to a higher weight of cost. This means that travelling with family,

friends or partner gives a higher weight to cost than travelling alone. For the interaction between the travel company and perceived risk, again, four parameters are estimated. Again only 'other' is not significant, which does not give information anyhow. All the other levels are significant, and all have a negative signs. This is in line with reasoning, with people perceiving risk as higher when not travelling alone (source) but with family, friends or partner.

- **Travel purpose**: Four extra parameters are estimated for the interaction between the purpose (reason) of travelling and cost. All parameters are insignificant except for the interaction with 'travelling for work' (in comparison to travelling for leisure). This parameter has a positive sign; hence a lower weight can be found for people that are travelling for work in comparison to the reference. This makes sense as the VoT values for business are higher than for leisure (KiM, 2016). For the interaction between travel purpose and time again, four extra parameters are estimated. All parameters are again insignificant except for 'travelling for work'. In line with expectation, this parameter is negative, so a higher weight to travel time (more negative value) can be found when people travel for work.
- **Travel frequency**: Travel frequency interacts with perceived risk; therefore, three extra parameters are estimated. All parameters are significant. A counter-intuitive outcome is that people travelling 1 to 3 times per year have a bigger weight to perceived risk than people who do not travel at all. Partly this can be explained that a huge part of the sample chooses to travel 1 to 3 times per year. However, if people are travelling, perceived risk is likely to decrease.
- **Mode preference and ASC for plane and train**: The interaction between mode preference and the ASC for plane, train and car contributes to nine extra parameters. The base level of the dummy variables is bus. All parameters are significant except for the interaction of respondents having a preference for car with the ASC of plane and respondents that have a preference for train with the ASC of plane. All parameters do have the expected sign. Respondents that stated to prefer car have a positive weight to the ASC of car and a negative weight to the ASC of train. Respondents that have a preference for train have negative weight to the ASC of plane and train and positive to train. Respondents that have a preference for plane have a strong positive weight to the ASC of plane and a negative weight for both the ASC of train and car.

Model with all interaction included

In order to get to the model final model with only significant parameters and interactions, first, all significant interactions of the separate models are added into one 'big' model. Then a stepwise backwards elimination is used. The elimination steps to get to the final model is discussed. All parameter value tables can be found in Appendix C:

- **All significant interactions included**: All of the interactions are included with the base model. ρ^2 is 0.219. Final log-likelihood = -7820.572, compared to null log-likelihood of -10010.56. Not all parameters are significant. There are 47 parameters in total.
- **Elimination 1**: First all parameters with a p-value of 0.3 and higher are deleted. In total, 14 parameters have been removed, so there remain 33 parameters. Some other parameters now also do not turn out to be significant. ρ^2 slightly decreased to 0.218. The final log-likelihood decreased slightly to -7825.372, but this is always a consequence of removing parameters.
- **Elimination 2**: Now all parameters with a p-value of 0.1, 0.2 and higher are deleted. Only one parameter has been deleted, so there remain still 32 parameters. The ρ^2 remains the same with a value of 0.218. The final log-likelihood decreased by less the 1 to -7826.299.
- **Elimination 3**: Then all parameters with a p-value of 0.05 or higher are removed. Again two parameters have been removed, so there are 30 parameters. ρ^2 is 0.218, final LL -7828.69.
- **Elimination 4**: Some parameters became insignificant on the 5% level; these parameters are removed. One parameter is removed. So in total, 29 parameters. ρ^2 is 0.218, final LL -7830.035.
- **Elimination 5**: Some parameters became insignificant on the 5% level; these parameters are removed. At last, two parameters are removed. In total, there remained 27 parameters. The final ρ^2 is 0.218. The final log-likelihood -7831.553. All parameters are significant on the 5% level.

- **Final model 6**: The preference for car is also added to the utility function of car. This gives in total three extra parameters. So as a consequence, there are now 30 parameters. ρ^2 increased to 0.24! Final LL is -7605.543. However, some parameters turn now out to be insignificant.
- **Final model**: The parameters 4_7 and 8_MORE are not significantly different from each other, as both values are very similar and the standard error captures the value of the other parameter. Therefore these variables are combined. Some parameters were not significant now, and are removed. There are 25 parameters, and this is the final model. ρ^2 is still 0.24. Final LL sightly decreased to -7606.821.

6.2.4. Results MNL model

The results of the different MNL models are presented in the table 6.8. It may be deduced from the findings that all parameters have the expected sign. All remaining parameters are significant at the 5 percent level, since all insignificant parameters have been eliminated. As a result of including the interaction between risk and time, there is no separate risk parameter. The time parameter is different for train and car. However, the time parameter for planes has become insignificant and is thus removed from the model.

Table 6.8: Base model, base model with main attribute interactions & final model

Model	Base model			Main interaction			Final model		
Parameter	Value	t	p-value	Value	t	p-value	Value	t	p-value
ASC_PLANE	-0.359	-5.3	1.16e-07	-1.51	-11.3	0	-2.23	-14.1	0
ASC_TRAIN	0.35	6.4	1.53e-10	-0.197	-2.52	0.0118	-0.414	-4.16	3.15e-05
B_COMFORT	0.281	8.7	0	0.281	8.3	0	0.346	10.4	0
B_COST	-0.00306	-21.2	0	-0.00326	-23	0	-0.00973	-8.02	1.11e-15
B_TIME	-0.167	-26.5	0						
B_RISK	-0.22	-22	0						
B_TIME_C				-0.168	-23.1	0	-0.194	-24.8	0
B_TIME_T				-0.0991	-9.42	0	-0.114	-10.1	0
B_TIME_RISK_P				-0.0437	-9.9	0	-0.011	-3.55	0.000392
B_TIME_RISK_T				-0.039	-17.1	0	-0.0107	-4.32	1.57e-05
B_GENDER_PR							-0.0217	-10.5	0
B_EDU_HBO_PR							-0.0142	-4.26	2.05e-05
B_EDU_WO_PR							-0.0237	-6.63	3.27e-11
B_EDU_OTHER_PR							-0.018	-4.99	6.05e-07
B_AGE_COST_20_40							0.00325	2.6	0.00932
B_AGE_COST_40_65							0.0064	5.18	2.23e-07
B_AGE_COST_65_AND_OLDER							0.00723	5.88	4.17e-09
B_PAYMENT_WORKEDU_COST							0.00223	3.18	0.00146
B_COMPANY_PR_FRIENDS							-0.00746	-2.62	0.00878
B_COMPANY_PR_OTHER							0.0197	4.06	4.86e-05
B_PURPOSE_WORK_TIME							-0.0596	-3.98	6.82e-05
B_PREF_CAR_C							0.496	6.3	2.97e-10
B_PREF_CAR_T							-0.268	-3.33	0.000881
B_PREF_PLANE_P							1.04	14.5	0
B_PREF_TRAIN_C							-0.184	-2.43	0.0152
B_PREF_TRAIN_T							0.404	5.49	3.99e-08
ρ^2	0.126			0.133			0.24		

6.2.5. Comparison of all models

The comparison of the different estimated models can be found in 6.9. The difference between the base model and the base model with the main interaction is not very big. After including all significant socio-demographic and travel behaviour interactions, there is a huge decrease in log-likelihood and thus a huge increase in ρ^2.

Table 6.9: Comparison of models

Model	Parameters	ρ^2	ρ^2-bar	Initial log-likelihood	Final log-likelihood
Base	6	0.126	0.126	-10010.56	-8745.12
Base interaction	8	0.133	0.132	-10010.56	-8682.56
Final MNL	25	0.24	0.238	-10010.56	-7606.82

The likelihood ratio statistic (LRS) test is only useful when a more complicated model B can be obtained by constraining model A. This is the case with these three models. The calculation of both LRS values can be found in the following equations.

$$LRS = -2 * (LL_{MNLbase} - LL_{MNLbaseinteraction} = -2 * (-8745.12 + 8682.56) = 125.12 \quad (6.1)$$

The χ^2 value for adding two parameters is 7.378 for the 5% significance level and 9.210 for the 1% significance level. This means that the probability that the 'MNL base interaction' model fits better than the 'MNL base' by coincidence is less than 1%. So it can be concluded that this model is the better fitting model. However, when all the different interactions are included, this gives the following value.

$$LRS = -2 * (LL_{MNLbaseinteraction} - LL_{MNLfinal} = -2 * (-8682.56 + 7606.82) = 2151.48 \quad (6.2)$$

The χ^2 value for adding 17 parameters is 30.191 for the 5% significance level and 33.409 for the 1% significance level. Therefore the addition of the extra parameters (interactions) is justified.

6.2.6. Contribution to utility main variables & interactions

The final model has been established, so the final parameter estimations have been determined. The range of the contribution of the main variables is represented in table 6.10. Moreover, there are several graphs to visualise the contribution of socio-demographic and travel behaviour interaction with the main attributes.

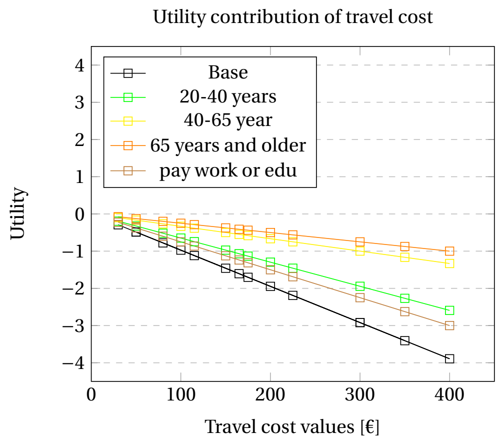
Table 6.10: Utility contribution of parameters

Parameter	Min. value	Max. value	Min. Utility contribution	Max. utility contribution
Travel cost train	€30	€350	-0.2919	-3.4055
Travel cost plane	€50	€400	-0.4865	-3.892
Travel cost car	€80	€200	-0.7784	-1.946
Travel time train	3h	12h	-0.342	-1.368
Travel time car	4.5h	16h	-0.873	-3.104
Travel comfort train	2^{nd} class	1^{st} class	0	0.346
Travel comfort plane	Economy	Business	0	0.346
Perceived risk train (dependent on time)	PR = 1 & TT = 3h	PR = 5 & TT = 12h	-0.0321	-0.642
Perceived risk plane (dependent on time)	PR = 1 & TT = 3h	PR = 5 & TT = 12h	-0.033	-0.33
Interaction gender on perceived risk	PR = 1 & TT = 3h	PR = 5 & TT = 12h	-0.033	-1.962
Interaction HBO education level on perceived risk	PR = 1 & TT = 3h	PR = 5 & TT = 12h	-0.033	-1.512
Interaction WO education level on perceived risk	PR = 1 & TT = 3h	PR = 5 & TT = 12h	-0.033	-2.082
Interaction 'other' education level on perceived risk	PR = 1 & TT = 3h	PR = 5 & TT = 12h	-0.033	-1.74
Interaction age 20 to 40 years on travel cost	€30	€400	-0.2919	-2.592
Interaction age 40 to 65 years on travel cost	€30	€400	-0.2919	-1.332
Interaction age 65 years and older on travel cost	€30	€400	-0.2919	-1
Interaction work or education pays for trip on travel cost	€30	€400	-0.2919	-3
Interaction traveling with friends on perceived risk	PR = 1 & TT = 3h	PR = 5 & TT = 12h	-0.033	-1.1076
Interaction traveling with 'other' on perceived risk	PR = 1 & TT = 3h	PR = 5 & TT = 12h	-0.033	0.522
Interaction traveling for work on travel time	3h	16h	-0.342	-4.0576
Effects of car preference on ASC car	0	1	0	0.496
Effects of car preference on ASC train	0	1	0	-0.268
Effects of plane preference on ASC plane	0	1	0	1.04
Effects of train preference on ASC car	0	1	0	-0.184
Effects of train preference on ASC train	0	1	0	0.404

Travel cost

Travel cost is generic for all modes with a parameter value of -0.00973. For travel cost, the only two interactions that turned out to be significant were 'age' and 'payment by education or work', see graph 6.1. All levels of age are significant. The base is younger than 20 years old. For every increase in age (group), the contribution to utility is bigger. Being at the 'age of 20 to 40 years has a positive contribution of 0.00325 utility points, so for this group, the travel cost parameter becomes $-0.00973 + 0.00325 = -0.00648$. For '40 to 65 year' the positive utility contribution is 0.0064 utility points, so the parameter becomes $-0.00973 + 0.0064 = -0,00333$. For the last group, '65 years and older, the positive utility contribution is 0.00723, so the cost parameter for this group becomes $-0.00973 + 0.00723 = -0,00250$. If the trip is paid by the educational institution or work, the contribution is 0.00223, which is lower than all age interactions. The parameter for travel cost becomes $-0.00973 + 0.00223 = -0.00750$.

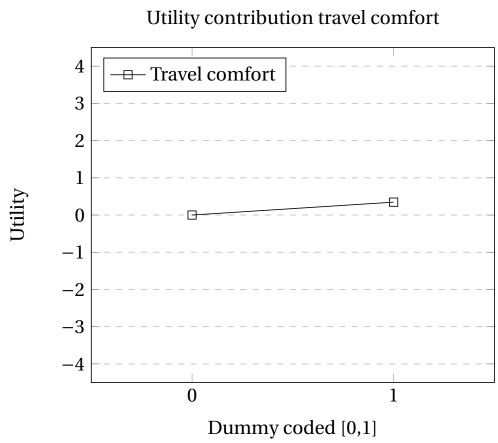
Figure 6.1: Graph with utility contribution of travel cost



Travel comfort

Comfort does have a positive contribution of 0.346. As this parameter did not have significant interactions, no relationships with socio-demographic and travel behaviour variables are found. No difference is made between train and plane, so the contribution from 2^{nd}/ economy class to 1^{st}/ business class is always 0.346 utility points, see graph 6.2.

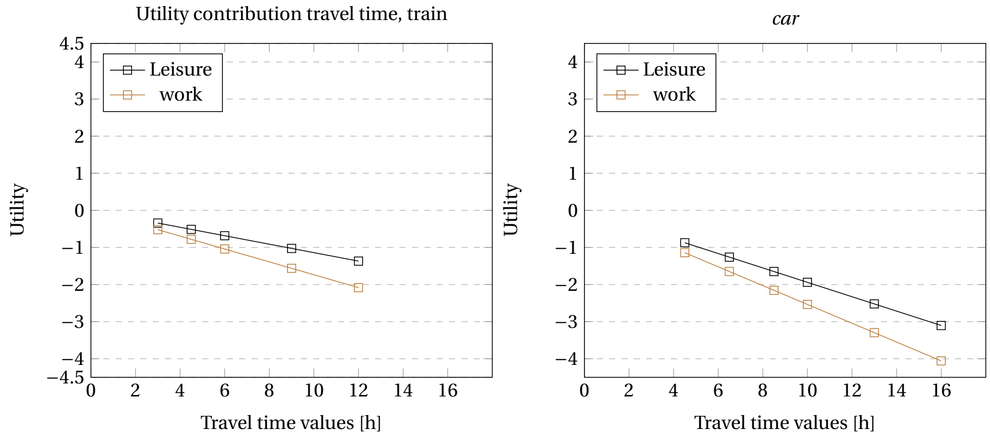
Figure 6.2: Graph with utility contribution of travel comfort



Travel time

Both travel time parameters are different for car and for train. The parameter for car is more negative than for train, with a value of -0.194. The value of train is -0.114. Respondents perceive time more negative in the car, than on the train. This makes sense as time in car is often perceived more negative than for train (KiM, 2020). The parameter of travel time for plane is not significant. This means that within this sample, respondents are indifferent to travel time when travelling by plane. The only significant interaction with time is travel purpose work. This interaction contributes to a more negative time parameter by -0.0596 utility points for every *extra hour* of travel time for both plane and car. The parameter for car becomes $-0.194-0.0596 = -0.253$. For train it becomes $-0.114-0.0596 = -0.1736$, see graph 6.3.

Figure 6.3: Graph with utility contribution on travel time



Perceived risk *dependent of time*

The main effect for this parameter is already an interaction between perceived risk and time. So all parameters are for every risk level and for every hour of travel. The parameters are almost identical; the value for plane is -0.011 and for train -0.0107. As both values are almost identical, only the value of plane is used from now on. Due to the late discovery of this small change, it is chosen not to change the estimation again, as all parameters would (slightly) be different. Gender, education, and travelling with friends turned out significant. The utility contribution of gender is negative, with a value of -0.0217. Therefore the parameter for men is the same as the base parameters; for women, it becomes $-0.011-0.0217 = -0.0327$. Education level 'MBO or lower' is again the base. HBO and WO both contribute to extra perceived risk, WO has the highest contribution of the two. The parameter for HBO becomes $-0.011-0.0142 = -0.0252$ and for WO $-0.011-0.0237 = -0.0347$. The last interaction is about travel company; travelling alone is the base, and travelling with friends gives the following parameter $-0.011-0.0237 = -0.01846$, see graph 6.4

Figure 6.4: Graph with utility contribution of perceived risk

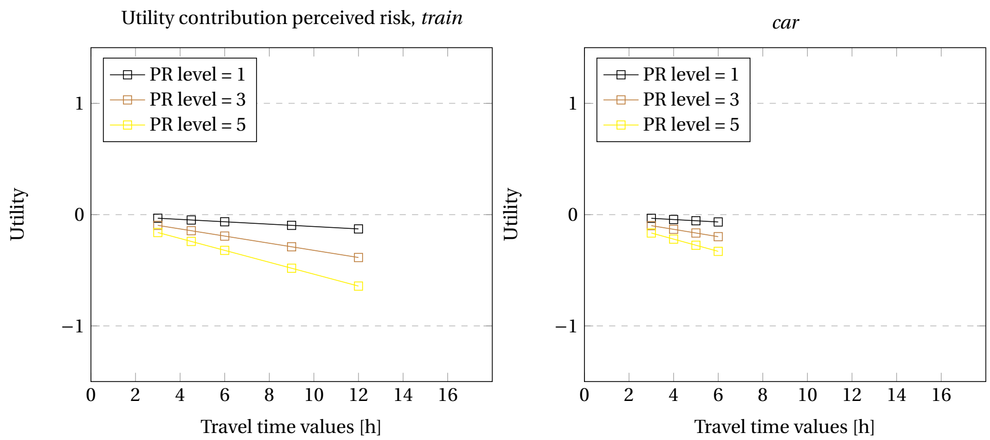


Figure 6.5: Graph with utility contribution of gender interaction with perceived risk, perceived risk level = 3

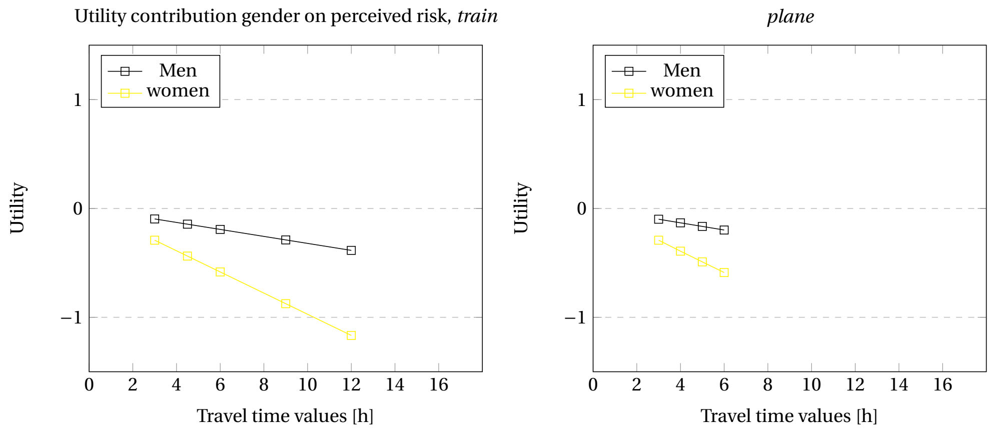
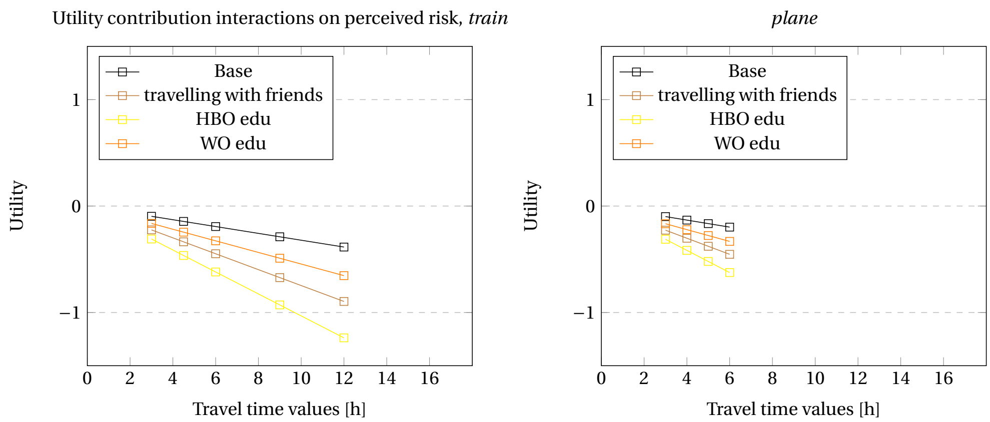


Figure 6.6: Graph with utility contribution of different interactions, perceived risk level = 3



6.2.7. Interpretation of parameters

The interpretation of the parameters provides insightful information on choice behaviour. Because of the different interactions, there is a huge amount of combinations possible, that changes the parameters and as a consequence, the interpretation as well. In order to keep the interpretation understandable, the 'average respondent' is used.

Average respondent

Several interactions contribute to the different main parameters. It is highlighted that the WtP for perceived risk, particularly in terms of travel costs, can vary significantly across people of varied ages, payment, education level and travel companies, and possibly other unobserved background variables not represented in this study. Including all of these different combinations would lead to a dramatic increase in Value-of-Time (VoT) and WtP values. The average respondent is based on the average (and when not possible) on the most common value. Therefore; these are the following assumptions:

- **Age**: The average value found of the sample for age is '1.9'; this is the consequence of the coding used. Younger than 20 years is coded as 0, 20-40 years is coded as 1, 40-65 years is coded as 2, and 65 years and older as 3. The value of 1.9 is thus 20 + 0.9*(40-20)=38. So this corresponds to an age of 38. Therefore the average respondent is 38 years old.
- **Education level**: The average value found for the sample is 1.05. MBO or lower is coded as 0, HBO is coded as 1, and WO is coded as 2. Therefore the value is very close (just above) to HBO. For this reason, HBO is used as the education level.
- **Gender**: Results are shown for both. Both genders are about 50% of the sample, so both genders are taken into account. At the same time, gender has the highest (absolute) contribution to perceived risk. Gerhold (2020) also found that women perceive risk (in regard to COVID-19) higher than men.
- **Trip purpose**: Leisure. An average for this value is not possible, as the levels are nominal. Most of the respondents gave the answer leisure (69%, see chapter 5).
- **Payment**: Payment is nominal as well. Almost 90% of the respondents gave this answer; therefore, payment is made by the respondent itself.
- **Travel company**: Also nominal. Only the interaction effect with friends is significant. It is assumed that respondents travel alone; the effect with friends is unimportant for this research.

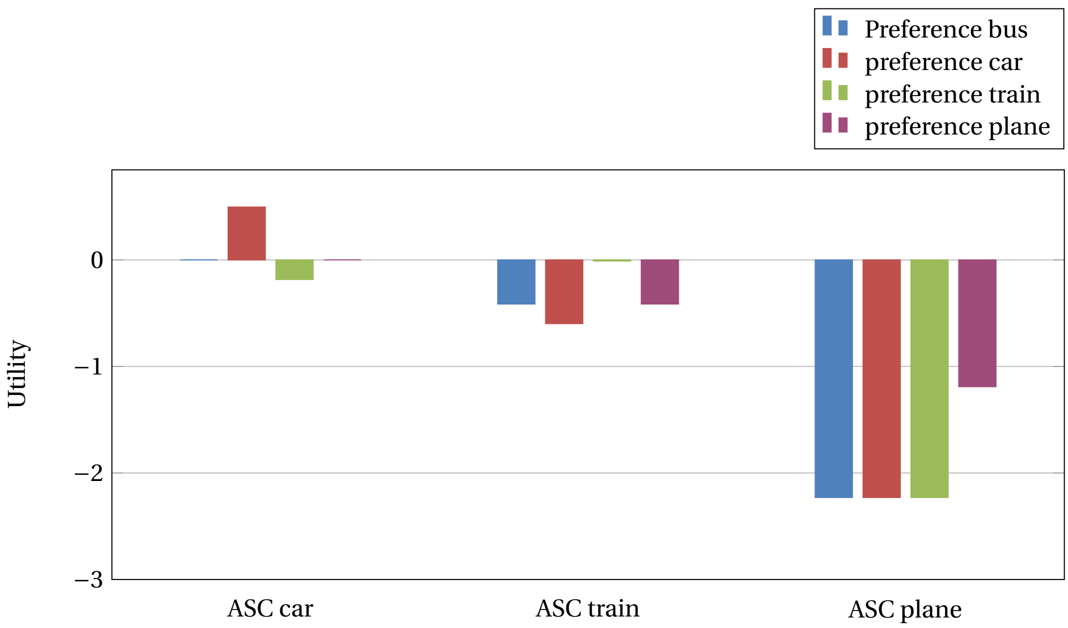
Table 6.11: The average respondent

Average respondent	
Age	38
Education	HBO
Gender	Men & women
Trip purpose	Leisure
Payment	Self
Travel company	Alone

Alternative specific constant

The Alternative Specific Constant (ASC) is the utility when all attributes have a value of zero. As a result, they perform a similar function to the constant in a regression model, which likewise captures the average effect if all of the factors are not included (Berkeley, 2002). In this case, travel time and travel cost would be zero; this is not possible for alternatives in real world situations. For this main choice experiment, the ASC of the car is set to zero. The ASC can also be interpreted as the preference respondents have for a certain mode, but this preference is not captured in the parameters. All of the interaction effects are included in figure 6.7. The ASC for plane is -2.23 and for train -0.414. This means that respondents do prefer the car over train and plane (if all parameters are zero). Train is again preferred over plane. If the interaction on the preference stated by the respondents is taken into account, the following conclusions can be taken. Respondents that stated to have plane as their preferred mode have 1.04 positive utility points, and therefore the ASC for plane for them becomes less negative. The parameter becomes $-2.23+1.04=-1.19$ (shown in purple for ASC plane). No significant relationship with respondents preferring plane on the ASC of train and car can be found (so the ASC stays 0 for car and -0.414 for train, also shown in purple). Respondents that do prefer car have a positive contribution to the utility of car with 0.496 (so this parameter becomes 0.496, shown in red for ASC car), and a negative contribution to the ASC of train with about half of the utility points with a value of -0.268, so this parameter becomes then $-0.414-0.268=-0.682$ (shown in red for ASC train). No significant parameter was found on the ASC of the plane, hence the ASC for plane stays -2.23 (shown in red for ASC plane). Respondents that stated to prefer train have a positive contribution on the ASC of train with 0.404 utility points (so this parameter becomes $-0.414+0.404=-0.01$, shown in green for ASC train). For ASC car, a negative contribution is found with only -0.184 utility points (so becomes -0.184, shown in green for ASC car). No relation on the ASC for plane is significant, so the ASC of plane stays -2.23 (shown in green for ASC plane). People that stated to prefer bus, all have the base utilities from the different ASC's, all shown in blue. This means ASC car stays 0, ASC train stays -0.414 and the ASC of plane stays -2.23.

Figure 6.7: Interaction effect on ASC for different mode preferences



Trade-off perceived risk and travel cost: value of risk (VoR)
The value of risk (VoR) for one level decrease in perceived risk is stated in the equation below.

$$VoR_{in\ travel\ cost} = \frac{\frac{\delta U}{\delta PR}}{\frac{\delta U}{\delta TC}} \tag{6.3}$$

Value of Risk (Willingness to pay for decrease in perceived risk) in terms of travel cost can be calculated with equation 6.4. *Note that the values are for every hour travel time and for every level of risk.*

$$VoR = \frac{\beta_{TT*PR} + \beta_{gender} * gender + \beta_{HBO} * HBO + \beta_{WO} * WO + \beta_{company_{friends}} * friends}{\beta_{TC} + \beta_{age_{20-40}} * Age_{20-40} + \beta_{age_{40-65}} * Age_{40-65} + \beta_{age_{>65}} * Age_{>65} + \beta_{pay_{eduwork}} * Pay_{edu-work}} \tag{6.4}$$

Using the average respondent values, the VoR values is calculated in equation 6.5 & 6.6.

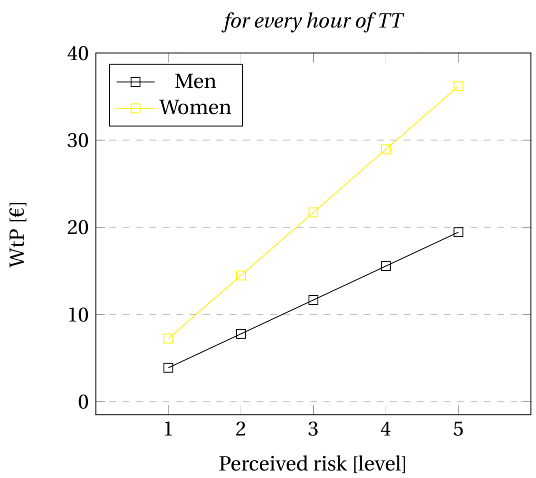
$$VoR_{men} = \frac{-0.011-0.142}{-0.00973+0.00325} = 3.89\ euro \tag{6.5}$$

$$VoR_{women} = \frac{-0.011-0.142-0.0217}{-0.00973+0.00325} = 7.24\ euro \tag{6.6}$$

From the calculation, it can be found that the VoR for a level in decrease of perceived risk for every hour is €3.89 for men. This value is close to an earlier research from van de Wiel (2021); in this research a value of €4.64

was found. However in this research, no distinction was made between men and women. For 3 hours of travel time, men are willing to pay €11.67 euro for every decrease in the level of risk. This is three times €3.89. The interaction between perceived risk and gender is significant. This results in a value for women that is (almost) double with €7.24. For a trip of 3 hours, this becomes then €21.71. There are no quadratic components specified; hence both perceived risk and travel time have a linear contribution. Therefore all of the values are multiplication of perceived risk and travel time. As a consequence the VoR value is the same for every decrease in risk level and every increase in travel time. Again, a different combination of socio-demographic and travel behaviour interactions would lead to different VoR values. The resulting VoR difference between men and women is in line with earlier research (Gustafsod, 1998; Finucane et al., 2000). These papers found significant difference between gender in risk perception. Moreover, Gerhold (2020) found that women have on average 1.5 times higher fear for risk than men; in this research the effect is almost double. The VoR values can be found in figure 6.8.

Figure 6.8: Value of perceived risk in terms of travel cost



Trade-off perceived risk and travel comfort

The value of risk in terms of travel comfort can be calculated with formula 6.8. The results show that men are willing to give up 0.072 comfort points for every hour of travel time to reduce one level of perceived risk, and women 0.134 comfort points for every hour of travel time to reduce one level of perceived risk. The VoR in comfort for every risk level is shown in figure 6.9.

$$VoR_{in\,travel\,comfort} = \frac{\frac{\delta U}{\delta PR}}{\frac{\delta U}{\delta CF}} \tag{6.7}$$

$$VoR_{in\,comfort} = \frac{\beta_{TT*PR} + \beta_{gender} * gender + \beta_{HBO} * HBO + \beta_{WO} * WO + \beta_{company_{friends}} * friends}{\beta_{CF} * comfort} \tag{6.8}$$

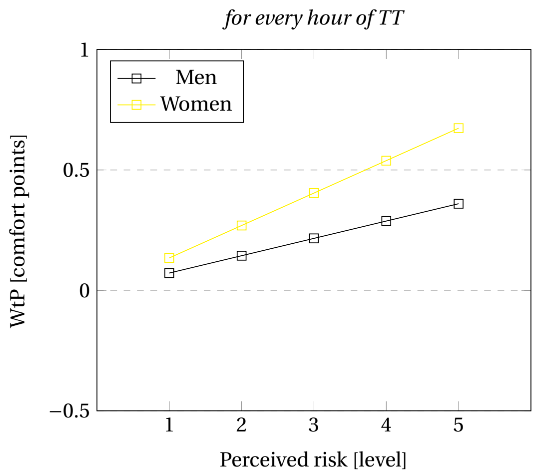
$$VoR_{in\,comfort,men} = \frac{-0.011 - 0.142}{0.346} = 0.072 \tag{6.9}$$

$$VoR_{in\,comfort,women} = \frac{-0.011 - 0.142 - 0.0217}{0.346} = 0.134 \tag{6.10}$$

The trade-off (TO) for a full reduction/increase in class, this means one full comfort 'point' (so 1^{st} class $\longleftrightarrow$ 2^{nd} class or business $\longleftrightarrow$ economy) is calculated with equation 6.11. Economy/2^{nd} class is coded as 0, business/1^{st} class is coded is 1.

$$TO_{men} = \frac{1-0}{0.072} * PR, \; TO_{women} = \frac{1-0}{0.134} * PR \tag{6.11}$$

Figure 6.9: Value of perceived risk in terms of travel comfort



The results for the different perceived risk levels are shown in table 6.12. For every different risk level, this table shows the trade-off in travel time to one full comfort class difference. The results show that for low levels of perceived risk, travel comfort is 'worth' a substantial hours of travel time. When perceived risk is very high (e.g. 5), for men a full difference in travel class is 'worth' only 2 hours en 47 minutes; for women this is 1 hour and 29 minutes. Moreover, the difference between levels is of importance as well (as this is possible in real life). To illustrate this; for a reduction from 5 to 1 in level of perceived risk; $\frac{1}{0.36-0.072} = 3.47$ hours of travelling for men. For women this is $\frac{1}{0.670-0.134} = 1.87$ hours of travelling for women. For every hour, a decrease in risk from $5 \rightarrow 1$, is $(0.072 * 5) - (0.072 * 1) = 0.288$ comfort points. For women this value is $(0.134 * 5) - (0.134 * 1) = 0.536$ comfort points. For longer travel times, the value is equivalent to 0.288 times the amount of hours travelled for men; for women 0.536 times the amount of hours travelled.

Table 6.12: Trade off comfort travel time

Men			Women		
PR level	Comfort [points]	TT [hours]	PR level	Comfort [points]	TT [hours]
1	0.072	13.889	1	0.134	7.463
2	0.144	6.944	2	0.268	3.731
3	0.216	4.630	3	0.402	2.488
4	0.288	3.472	4	0.536	1.866
5	0.36	2.778	5	0.670	1.493

Value of time (VoT)

The value of time in transportation (economics) is the potential cost of the time that a passenger spends on their route. In essence, this is the amount a passenger is ready to pay to save time, or the amount they would take as compensation for lost time. The amount of time that passengers save is one of the key justifications for transportation upgrades (Kouwenhoven et al., 2014). The economic advantages of a transportation project may be defined using a set of temporal values and compared to the expenses (thus forming the basis of cost-benefit analysis). Travel time savings (or increases) in particular are part of the shift in consumer surplus for a transportation investment. The Value-of-Time equation for both train and car is shown in equation 6.13. The calculation are shown in equation 6.14 & 6.15. The VoT for train has a value of €17.59 for train and €29.94 for car.

$$VoT = \frac{\frac{\delta U}{\delta TT}}{\frac{\delta U}{\delta TC}} \tag{6.12}$$

$$\text{VoT}_{train/car} = \frac{\beta_{TT_{train/car}} + \beta_{travelpurpose} * work}{\beta_{TC} + \beta_{age_{20-40}} * Age_{20-40} + \beta_{age_{40-65}} * Age_{40-65} + \beta_{age_{>65}} * Age_{>65} + \beta_{pay_{eduwork}} * Pay_{edu-work}} \tag{6.13}$$

$$VoT_{train} = \frac{-0.1140}{-0.0097 + 0.0033} = 17.59\ euro/hour \tag{6.14}$$

$$VoT_{car} = \frac{-0.194}{-0.0097+0.0033} = 29.94\ euro/hour \tag{6.15}$$

The different VoT values found for this thesis in comparison to other studies is shown in table 6.13. The fact that car has a higher VoT is in line with KiM *Netherlands Institute for Transport Policy Analysis* (KiM, 2020). The parameter of time for plane is not significant; hence, no VoT can be specified for plane. In this study, the VoT for car is just over 2 euros more. For Kouwenhoven et al. (2014) there is a very small difference found between the modes. Shires and De Jong (2009) did not find differences in VoT for train and car. It must be noted that both KiM (2020) and Kouwenhoven et al. (2014) focused on short distance travel, while Shires and De Jong (2009) focused on long-distance travel. For this thesis, a possible reason for the fact that travel time in train gets a lower weight than for car (for long-distance travel), is that in the train people are able to move, have a drink or food and entertain themselves. This is a possible reason that travel time in train is perceived less negative. However, this is not confirmed in literature. The values are a bit higher than the values found by KiM, this is a consequence of long-distance travel. Shires and De Jong (2009) confirmed with their study "An international meta-analysis of values of travel time savings" that the VoT for long-distance travel is higher than for short distance (Shires and De Jong, 2009). Moreover, in a study done by Börjesson and Eliasson (2014, VoT values for long-distance trips were found to be twice as high in comparison to short distance. Actually, the average value of both train and car is very close to the value found by Shires and De Jong (2009), €23.77 to €24.00, see table 6.13.

Table 6.13: VoT results of different studies

	This thesis	KiM (2020)	Van Kouwenhoven (2014)	Shires and De Jong, (2009)
VoT Train	17.59	13.22	9.25	24.00
VoT Car	29.94	15.58	9.00	
Average	23.77	14.40	9.13	24.00

WtP for comfort

The equation for Willingness to pay for comfort for both train and plane can be found in equation 6.17. The calculation is shown in equation 6.18.

$$WtP_{comfort} = \frac{\frac{\delta U}{\delta CF}}{\frac{\delta U}{\delta TC}} \tag{6.16}$$

$$\text{WtP}_{comfort} = \frac{\beta_{comfort} * comfort}{\beta_{TC}+\beta_{age_{20-40}} * Age_{20-40}+\beta_{age_{40-65}} * Age_{40-65}+\beta_{age_{>65}} * Age_{>65}+\beta_{pay_{eduwork}} * Pay_{edu-work}} \tag{6.17}$$

$$WtP_{comfort} = \frac{0.346}{-0.0097+0.0033} = 53.40\ euros \tag{6.18}$$

The result show the willingness to pay for an upgrade in class (so 2^{nd} class → 1^{st} class or economy → business) is €53.40. No mode-specific β turned out to be significant, so there is no distinction between classes in train and plane. Balcombe et al. (2009) found a value of about €120, but they agreed that this value is on the high side. A big amount of the passengers travelling in business or 1^{st} class is flying for business purposes (BusinessAM, 2020). It is therefore of interest to show the willingness to pay for business/educational travellers (respondents who's trip is payed by company of educational institution), the value is shown in equation 6.19.

$$WtP_{comfort,eduwork} = \frac{0.346}{-0.0097+0.0033+0.00222} = 81.22\ euros \tag{6.19}$$

This shows that the WtP for an upgrade is substantially higher, with a value of €81.22. On the other hand, most of the time, the difference in fare between economy class and business class is often higher (KLM, 2022). For example a return ticket with KLM for Amsterdam-London route from 6^{th} to 9^{th} September 2022 costs in economy €131 and €418 in business class, which is difference of €287 (KLM, 2022). Travelling by Eurostar on the same dates costs €164 in 2^{nd} class and €227 in 1^{st} class, so a difference of €64 (NS, 2022). The results implicate that the difference in class is about of the same order size as the WtP for comfort in this thesis. For KLM however, the implication might be to make business class cheaper in order to attract more passengers. Of course these as just examples, but it shows that for the train WtP values are plausible.

6.2.8. Combination results of both experiments

Now that both the perceived risk rating experiment and the main (mode) choice experiment have been estimated, it is possible to estimate the outcomes of both. In the rating experiment, perceived risk was the dependent variable, but in the main (mode) choice experiment, it was the independent variable. For the rating experiment, gender and education level had a positive contribution to perceived risk. This gives 'face validity as in the main choice experiment, gender and education both had a negative contribution on utility when interacting with perceived risk. Furthermore, an income of €20.000 to €40.000 and €40.000 to €100.000 contributed to perceived risk and age 20 to 40 years had a negative contribution. In the main choice, perceived risk is dependent on time.

Using the (absolute) linear regression coefficients of the perceived risk rating experiment, both experiments can be integrated. Then these values are combined with the WtP values that are just calculated. Because of the dummy coding, all different dummy variables have an independent contribution to the value of perceived risk. No interactions were taken into account within the regression analysis. For example the WtP for a decrease in load factor from 50% to 25% gives the following WtP values: $WtP_{men,\, LF\,50\%\rightarrow 25\%} = 15.08 - 7.54 =$ €7.54 for a travel time of 12 hours for men; $WtP_{women,\, LF\,50\%\rightarrow 25\%} = 28.07 - 14.04 =$ €14.03. This difference is due to the significant interaction between gender and perceived risk. In table 6.14, the WtP values in monetary terms can be found for each of the perceived risk rating attributes. Again, as there are no quadratic components, the increased WtP values are always linear by increasing the level of perceived risk and hours of travel time. So the difference between every increase in travel time (e.g., from 3 hours to 6 hours, or from 9 hours to 12 hours) is always €11.67 for men and €21.71 for women. This can be found in the first rows of the table. The rest of the table corresponds to the different (significant) factors from the perceived risk rating experiment. It must be noted that a negative value of the risk factor does mean it decreases perceived risk, and a positive value of the risk factor means a positive contribution to perceived risk. The table can be interpreted in different ways. The values in the table show the WtP values of the different risk factors in relation to a decrease in the level of risk by one for the mentioned travel time in the column. To calculate the WtP values for differences between levels of the risk factors, the difference of the coefficients from the linear regression model has to be multiplied by the hours of travel time and the perceived risk level. For example WtP for a decrease from 100% load factor to 50% load factor when traveling 6 hours with a perceived risk level of 3 gives: WtP = 3.889 * 6 * 3 * (0.646 – 0.485) = €22.63. For women this WtP would be WtP = 7,2376 * 6 * 3 * (0.646 – 0.485) = €42.11. So that is almost double the value of men. The same values could be derived from the table by using the parameter for load factor for 6 hours and multiplying this by 3 (for perceived risk level of 3), so WtP = 7.54 * 3 = €22.62. It does not make sense for all risk factors to calculate the WtP for differences in levels simply because these factors can not be affected. For example, the number of infections cannot be changed in a moment. But in this case, it is worth looking at the table the other way around. So, for example; with 10.000 infections per day. the WtP for women to reduce 1 level of risk when travelling for 12 hours would be €21.90; for men this value is €11.76. When there are 100.000 infections per day, women are willing to pay even €42.96 for a travel of 12 hours; men €23.09. All the positive parameters from the perceived risk rating experiment lead to a positive WtP; hence for a decrease in level of perceived risk, so people are 'willing' to pay more for a level of decrease in risk when this factor is there. An example is red travel advice. In comparison to the base (no travel advice), men are willing to pay €32.74 for a decrease in level of risk when travelling for 12 hours; women €60.93. If men are travelling for only 6 hours, this value is half of the 12-hour value, so €16.37; for women, €30.47.

Lastly, there are also some socio-demographic variables. They were added to the regression model as separate variables. The WtP values are for HBO educated people, and only the second part is women. Therefore, the part of gender for men is empty in the table. Being a woman leads to an additional WtP for 1 decrease in the level of perceived risk of €1.92 when travelling 3 hours. Moreover, having an income of €40.000 and €100.000 as a woman leads to an additional WtP for 1 decrease in level of perceived risk of €1.92 when travelling 3 hours. So the interpretation of the table can be two-sided. On the one hand, the WtP for an increase or decrease of the different risk attributes from the rating experiment can be calculated. On the other hand, given a certain risk attribute level, the WtP for a decrease in level of risk can be calculated (for each different travel time). Note that negative values of WtP values, in this case, has to be interpreted as follows. An example: if there is a vaccination rate of 70%, a man that is travelling for 3 hours would be willing to pay €10.63 less for a decrease of 1 risk level, compared to a vaccination rate of 0%. It can be concluded that the highest values for the WtP can be found for red travel advice and load factor for the longest travel time of 12 hours, this is €32.74 for men;

even €60.93 for women. For a load factor of 100%, this value is €56.15 for women; for men, it is €30.17. At last, infections are an important factor as well.

Table 6.14: WtP values for the different risk factors

			Men				Women			
		PR level difference	$\Delta_{level} = -1$				$\Delta_{level} = -1$			
		Travel time	3	6	9	12	3	6	9	12
		Value	€11.67	€23.33	€35.00	€46.67	€21.71	€43.43	€65.14	€86.85
Travel advice	Red travel advice	0.702	€8.18	€16.37	€24.55	€32.74	€15.23	€30.47	€45.70	€60.93
Vaccination rate	Parameter	-0.007	€-0.08	€-0.15	€-0.23	€-0.30	€-0.14	€-0.28	€-0.42	€-0.57
	15	-0.098	€-1.14	€-2.28	€-3.42	€-4.55	€-2.12	€-4.24	€-6.36	€-8.48
	30	-0.195	€-2.28	€-4.55	€-6.83	€-9.11	€-4.24	€-8.48	€-12.72	€-16.95
	70	-0.455	€-5.31	€-10.63	€-15.94	€-21.26	€-9.89	€-19.78	€-29.67	€-39.56
	90	-0.586	€-6.83	€-13.66	€-20.50	€-27.33	€-12.72	€-25.43	€-38.15	€-50.86
Load factor	Parameter	0.006	€0.08	€0.15	€0.23	€0.30	€0.14	€0.28	€0.42	€0.56
	25	0.162	€1.89	€3.77	€5.66	€7.54	€3.51	€7.02	€10.53	€14.04
	50	0.323	€3.77	€7.54	€11.31	€15.08	€7.02	€14.04	€21.06	€28.07
	75	0.485	€5.66	€11.31	€16.97	€22.63	€10.53	€21.06	€31.58	€42.11
	100	0.646	€7.54	€15.08	€22.63	€30.17	€14.04	€28.07	€42.11	€56.15
Infection rate	10.000 infections per day	0.252	€2.94	€5.88	€8.82	€11.76	€5.47	€10.95	€16.42	€21.90
	25.000 infections per day	0.123	€1.44	€2.88	€4.32	€5.76	€2.68	€5.36	€8.04	€10.72
	100.000 infections per day	0.495	€5.77	€11.54	€17.31	€23.09	€10.74	€21.48	€32.22	€42.96
Air conditioning	Only ventilation	-0.405	€-4.72	€-9.44	€-14.17	€-18.89	€-8.79	€-17.58	€-26.36	€-35.15
	Airco no HEPA	-0.198	€-2.31	€-4.62	€-6.92	€-9.23	€-4.29	€-8.59	€-12.88	€-17.18
	Airco with HEPA	-0.289	€-3.38	€-6.75	€-10.13	€-13.51	€-6.28	€-12.57	€-18.85	-25.14
Travel requirements	3G-policy	0.144	€1.68	€3.36	€5.03	€6.71	€3.12	€6.25	€9.37	€12.49
	2G+-policy	-0.272	€-3.17	€-6.35	€-9.52	€-12.70	€-5.91	€-11.81	€-17.72	€-23.63
Face mask policy	Any face mask	-0.268	€-3.12	€-6.24	€-9.37	€-12.49	€-5.81	€-11.62	€-17.43	-23.25
	At least FFP2	-0.140	€-1.64	€-3.27	€-4.91	€-6.55	€-3.05	€-6.09	€-9.14	€-12.18
	Socio-demographic									
Gender	Women	0.088		inapplicable			€1.92	€3.83	€5.75	€7.66
Education	HBO	0.137	€1.60	€3.21	€4.81	€6.41	€2.98	€5.97	€8.95	€11.94
Income class	Income €20.000 to €40.000	0.112	€1.31	€2.62	€3.92	€5.23	€2.43	€4.87	€7.30	€9.74
	Income €40.000 to €100.000	0.168	€1.96	€3.92	€5.88	€7.85	€3.65	€7.30	€10.95	€14.60
Age	Age 20 to 40 years	-0.095	€-1.11	€-2.22	€-3.32	€-4.43	€-2.06	€-4.12	€-6.18	€-8.25

6.3. Summary

Main points of chapter

- The largest contributor to perceived risk is red travel advice, together with a 100% load factor.
- The largest contributor to decreasing perceived risk is a vaccination rate of 90%. Also, ventilation has a large contribution to decreasing risk.
- Cleaning is not important for decreasing perceived risk
- 'Having any mask' is more important than 'at least an FFP2 mask' needed.
- The WtP for 1 level of decrease in perceived risk for every hour of travel time is €3.89 for men and €7.24 for women
- ASC plane and ASC train are both negative, so car is more preferred.
- Beta comfort is positive, beta time and beta cost are both negative. Travel time is more negative for car than for train. For plane, time is insignificant.
- Perceived risk is significant as an interaction with time. This means that for longer travel times, COVID-19 risk is perceived higher. Perceived risk is also negative and different for train and plan; however, there is a very small difference see table 6.8.
- Several interactions are significant with perceived, also on the ASCs and travel cost.
- A preference for car has a positive contribution to the ASC of car and a small negative contribution to train. A preference for train has a small negative contribution on the ASC of car and a positive contribution to the ASC of train. A preference for plane has a positive contribution to the ASC of plane.
- For travel time, only purpose work is significant. Respondents that travel for work are more sensitive to travel time.
- For travel cost, both age and payment by work or educational institution are significant. Higher age results in less weight to travel cost, the same counts if the trip is payed by work the educational institution.
- For perceived risk, educational level and gender are significant. Having at least HBO or WO education (in comparison to MBO) results in more weight to perceived risk. Women have a higher weight to perceived risk as well.

This chapter analysed the data collected from the survey. For perceived risk rating experiment, the largest contributor to perceived risk is *red travel advice*. However, the other levels *yellow and orange travel advice* do not contribute to perceived risk. Moreover, a high load factor leads to a high contribution to perceived risk. Almost the same (but opposite) effect can be found for vaccination rate, with a high contribution to decreased perceived risk. Furthermore, the same but opposite effects can be found for *100.000 infections per day* and *only ventilation*. Actually, *only ventilation* is more important than having air conditioning with HEPA filters. Probably, this is a consequence of the measures that were taken by the government, with a huge focus on ventilation. Three discrete choice models were estimated for the main choice experiment. It was chosen to only use the MNL model. The model, including all significant interactions, increased the ρ^2 from 0.126 to 0.24. This means 24% of the variance is explained by the model. Risk perception is almost the same for train as for plane. With the combination of both models, the WtP values for a decrease in level of risk are calculated. The VoR is €3.89 for men and €7.24 for women, which is based on the average respondent. The marginal WtP values vary from is €3.89 for men and €7.24 for women for 1 hour of travel to €46.67 for men and €86.85 for women for a 12-hour journey. The WtP in comfort points for perceived risk is for men 0.072 per hour; for women this is 0.134 per hour. People are WtP most when given a red travel advice, also a high load factor and very high infection rate (100.000 positive tests per day) are important. Also the VoT is calculated, with a value of €17.59 per hour for car; for train this is €29.94. The WtP for comfort is calculated as well, this value is €53.40 for an increase in travel class. When work or the educational institution pays, this value increases to €81.22.

7

Influence perceived risk on market share

This research takes a broad approach to long-distance travel within Europe. To give an example of what the influence is on modal split, it is chosen to go for three different routes. First, a short route that is popular by train is used, this is the Amsterdam - London route. At the same time, the plane and car options are viable and popular as well. The second route is Amsterdam - Berlin. This route has all three modes as viable options as well. At third, a longer route is chosen to see the difference between a short and longer routes. In order to look at the influence of perceived risk on the modal split, real-world values for travel time, travel cost and comfort are being used for these routes. Then, the perceived risk levels are varied so that the influence is discussed. Potential revenue differences due to load factor differences is discussed per route as well.

Table 7.1: The average respondent

Age	38
Education	HBO
Gender	Men & women
Trip purpose	Leisure
Travel company	Alone

7.1. Route Amsterdam - London

The Amsterdam - London route is a very busy route; in 2019, it was the busiest route from Amsterdam Schiphol Airport, with almost 2 million passengers per year (Schiphol, 2020). The direct Eurostar route is relatively new, being introduced back in 2018 (Treinenweb, 2018). However, it was already possible to go the London by Eurostar; first a Thalys train to Brussels had to be taken. At the same time, it is possible to go to London by car, by taking the Eurotunnel or Ferries. For the calculations, Rome2Rio, ViaMichelin, Google Maps, Skyscanner and NS international are used (Rome2Rio, n.d.; ViaMichelin, n.d.; Google, n.d.; Skyscanner, n.d.; NS, n.d.).

Table 7.2: Values example Amsterdam - London

	Amsterdam - London					
	Eurostar		NS + easyJet + Gatwick Express		Hatchback E95	
	TT train	TC train	TT plane	TC plane	TT car	TC car
Access	0	0	0.25 h	€ 5	0	0
Wait	0	0	2 h	0	0	€ 15 (toll)
Main	4 h	€ 125	1.25 h	€ 50	7.5 h	€ 65
Egress	0.25 h	0	0.50 h	€ 20	0	0
Total	**4.25 h**	**€125**	**3.75 h**	**€ 75**	**7.5 h**	**€80**

7.1.1. Policy implications train load factor

The WtP values in table 6.14 show what the different risk factors mean given a certain travel time for the reduction of perceived risk by one. For this Amsterdam - London route, the values are used to give policy implications in regard to the different risk factors that contribute to perceived risk. Load factor is chosen as this can be influenced by the train operator. Realistic values for travel time, travel cost and travel time are shown in table 7.2. Load factor is the most important mode-related risk factor. The Amsterdam - London route is served by Eurostar, the Eurostar e320 train has a maximum capacity of 900 passengers, and serves the route 3 times per day (Eurostar, n.d.). So the total capacity is 2700 passengers per day (one way). The duration is 4.25 hours. From table 6.14 the WtP per percentage point reduction in load factor can be calculated. For men this is 0.08/3 hours*4.25 hours = €0.113. For women this is 0.14/3*4.25 = €0.1983. Assuming the calculated price of a one-way ticket of €125, the total revenue for one day and one way can be calculated; this is 125*2700 = €337.500. Considering a reduction of 25% load factor, but with an increasing ticket price due to the WtP for reduction in load factor of 25*0.113= €2.825 per passenger for men and 25*0.1983= €4.9575 per passenger for women, the resulting revenue is (assuming gender equally): 2700*0.5*0.75*(125+2.825)+2700*0.5*0.75*(125+4.9575) = €261.005. When a 50% load factor is considered, but the ticket price increases by €5.65 for men and €9.91 for women, the resulting revenue is: 2700*0.5*0.5*(125+5.65)+2700*0.5*0.5*(125+9.915) = €179.253. Considering a reduction of 75% load factor, but with an increasing ticket price of €8.475 for men and €14.8725 for women, the resulting revenue is: 2700*0.5*0.25*(125+8.475)+2700*0.5*0.25*(125+14.8725) = €92.255. The results are shown in table 7.3. It can be concluded that the WtP for reduction in load factor can not account for the revenue loss due to this reduction in load factor. However, the WtP increases revenue in comparison with the same load factor without the WtP.

Table 7.3: Revenues example Amsterdam - London

Eurostar				
	Revenue		WtP reduction LF	
Load factor	With WtP	Without WtP	Men	Women
25%	92255	84375	8.48	14.87
50%	179256	168750	5.65	9.92
75%	261005	253125	2.83	4.96
100%	337500	337500	0	0.00

7.1.2. Impact perceived risk

Table 7.4: Values example Amsterdam - London, utilities & predicted market shares

Men		Values route Amsterdam - London					
Situation		Utility			Market share		
PR train	PR plane	Train	Plane	Car	Train	Plane	Car
1	1	-1.81	-2.81	-1.97	45%	17%	38%
1	2	-1.81	-2.91	-1.97	46%	15%	39%
1	3	-1.81	-3.00	-1.97	46%	14%	40%
1	4	-1.81	-3.09	-1.97	47%	13%	40%
1	5	-1.81	-3.19	-1.97	47%	12%	41%
2	1	-1.92	-2.81	-1.97	42%	17%	40%
2	2	-1.92	-2.91	-1.97	43%	16%	41%
2	3	-1.92	-3.00	-1.97	44%	15%	41%
2	4	-1.92	-3.09	-1.97	44%	14%	42%
2	5	-1.92	-3.19	-1.97	45%	13%	43%
3	1	-2.03	-2.81	-1.97	40%	18%	42%
3	2	-2.03	-2.91	-1.97	40%	17%	43%
3	3	-2.03	-3.00	-1.97	41%	16%	43%
3	4	-2.03	-3.09	-1.97	42%	14%	44%
3	5	-2.03	-3.19	-1.97	42%	13%	45%
4	1	-2.13	-2.81	-1.97	37%	19%	44%
4	2	-2.13	-2.91	-1.97	38%	18%	44%
4	3	-2.13	-3.00	-1.97	39%	16%	45%
4	4	-2.13	-3.09	-1.97	39%	15%	46%
4	5	-2.13	-3.19	-1.97	40%	14%	47%
5	1	-2.24	-2.81	-1.97	35%	20%	45%
5	2	-2.24	-2.91	-1.97	36%	18%	46%
5	3	-2.24	-3.00	-1.97	36%	17%	47%
5	4	-2.24	-3.09	-1.97	37%	16%	48%
5	5	-2.24	-3.19	-1.97	37%	14%	48%
		Average			41%	16%	43%

Women		Values route Amsterdam - London					
Situation		Utility			Market share		
PR train	PR plane	Train	Plane	Car	Train	Plane	Car
1	1	-1.91	-2.89	-1.97	43%	16%	41%
1	2	-1.91	-3.07	-1.97	44%	14%	42%
1	3	-1.91	-3.24	-1.97	45%	12%	43%
1	4	-1.91	-3.42	-1.97	46%	10%	43%
1	5	-1.91	-3.60	-1.97	47%	9%	44%
2	1	-2.10	-2.89	-1.97	39%	18%	44%
2	2	-2.10	-3.07	-1.97	40%	15%	45%
2	3	-2.10	-3.24	-1.97	41%	13%	46%
2	4	-2.10	-3.42	-1.97	42%	11%	47%
2	5	-2.10	-3.60	-1.97	42%	10%	48%
3	1	-2.30	-2.89	-1.97	34%	19%	47%
3	2	-2.30	-3.07	-1.97	35%	16%	49%
3	3	-2.30	-3.24	-1.97	36%	14%	50%
3	4	-2.30	-3.42	-1.97	37%	12%	51%
3	5	-2.30	-3.60	-1.97	38%	10%	52%
4	1	-2.50	-2.89	-1.97	30%	20%	50%
4	2	-2.50	-3.07	-1.97	31%	17%	52%
4	3	-2.50	-3.24	-1.97	32%	15%	53%
4	4	-2.50	-3.42	-1.97	32%	13%	55%
4	5	-2.50	-3.60	-1.97	33%	11%	56%
5	1	-2.70	-2.89	-1.97	26%	21%	53%
5	2	-2.70	-3.07	-1.97	27%	18%	55%
5	3	-2.70	-3.24	-1.97	27%	16%	57%
5	4	-2.70	-3.42	-1.97	28%	14%	58%
5	5	-2.70	-3.60	-1.97	29%	12%	59%
		Average			36%	14%	50%

For this case, realistic values are calculated with the aforementioned websites. The values of travel time, travel cost, and perceived risk are added to the utility functions. Then all combinations of perceived risk are varied in the table so that every combination of perceived between train and plane is in the table. This results in a different utility for every row for every mode. With the utilities, the choice probabilities can be calculated with the following formula:

$$P(i \mid C) = \frac{e^{V_i}}{\sum_{j=1}^{j \in C} e^{V_j}} \tag{7.1}$$

It has to be noted that the numbers sometimes do not fully add up to 100% due to rounding. The results show that in general that perceived risk does not have a big influence on the market share of the modes for the Amsterdam - London route. To highest changes in perceived risk level between train and plane are used to illustrate the most extreme possible market share shift. When train risk level remains 1 and risk level for plane increases from 1 to 5, the difference in market share for train is only 2% point increase for men; for women this is 4% point. For plane there is a decrease of 5% point for men; for women this is 7% point. For car there is an increase of 3% point for men; for women this is 3% as well. When the train increases from 1 to 5 and the plane stays at level 1, a fairly different picture can be seen. In this case, train loses 10% point market share for men; for women this is 17%. Plane has 3% point decrease in market share for men; for women this is 5% point. Car increases by 7% for men; for women this is 12% point. This shows train changes in perceived risk weigh more than for plane. Partly, this is because train has a longer travel time. However, due to the highly negative ASC for plane, differences in utilities for plane are smaller than for train. Therefore, other attributes like perceived risk, have a lower impact. Thus, changes in perceived risk are greater for train than for plane. Women have a higher weight to perceived risk, therefore market shares differences are higher for women. For smaller differences in perceived risk, these market share changes are even smaller. The perceived risk levels are complex concepts, its score is determined by a lot of factors. Therefore there is no real such thing as risk level 1. So, these extreme changes are not realistic in real life. Moreover, four out of the eight factors are destination specific. In this case, these factors do not change between modes for the same OD-pair. To make this clear, the maximum difference in perceived risk points for the same OD-pair will be (only considering mode-related attributes):

$$PR_{low} = 2.815 - 0.405 - 0.268 = 2.14$$
$$PR_{high} = 2.815 + (0.006 * 100) = 3.46$$
$$\Delta PR = 3.46 - 2.14 = 1.32$$

In this case, when the perceived risk of train is 2 and plane 3 (to approximate risk point 2.14 and 3.46) to train 1 and plane 3, there is only a 1 to 2% point market share difference for men; for women this is 2 to 3% point. Thus, it can be concluded that the influence of perceived risk is very moderate for this route. The maximum (possible) difference will only have a small impact on the market share for this Amsterdam - London route. The impact of one of the four mode-related risk attributes is even smaller. In conclusion, perceived risk has a small impact on market share for the Amsterdam - London route. However, this example gives more insights. It can be seen from the table that the market shares from the train are, in general, on the high side. Even though travelling by train is a little bit slower and a bit more expensive than plane, train has huge part of the market share. The reason for this is the very high (negative) value for the ASC for plane. As a consequence, the shorter travel time and cheaper fare can not compensate for this utility. So because of the strong aversion for plane in this dataset, the train has a lot of gains in market share. The difference in travel time between train and car is 1 hour and 45 minutes. However, the weight for travel time in car is substantially higher, with a value of -0.194 for car and -0.114 for train. This means that the weight is 70% higher, see next equation: $\frac{-0.194+0.114}{-0.114} = 0.70$. Therefore, the 1 hour and 45 minutes travel time difference has a quite substantial impact, so this contributes even more to the market share of train. Nevertheless, there is no perceived risk for car; at the same time, the ASC of train is negative (in comparison to car), so car has the highest (average) market share.

7.2. Route Amsterdam - Berlin

The route Amsterdam - Berlin is an important route for both NS international and Schiphol, being the 10^{th} busiest route from Schiphol in 2019 (including both the Tegel and Schönefeld airports) (Schiphol, 2020). Plane is the fastest option. By train, there is a direct connection by Intercity train from Amsterdam to Berlin Hauptbahnhof (NS, n.d.). The third option is car, with a travel that is almost the same as by train. The values for the attributes are shown in table 7.5. Again calculations are done using Rome2Rio, ViaMichelin, Google Maps, Skyscanner and NS international are used (Rome2Rio, n.d.; ViaMichelin, n.d.; Google, n.d.; Skyscanner, n.d.; NS, n.d.).

Table 7.5: Values example Amsterdam - Berlin

	Amsterdam - Berlin					
	NS international IC		NS + KLM + RE17 train		Hatchback E95	
	TT train	TC train	TT plane	TC plane	TT car	TC car
Access	0	0	0.25 h	€5	0	0
Wait	0	0	2 h	0	0	0
Main	6.25 h	€ 100	1.25 h	€ 50	6.5 h	€ 160 (fuel)
Egress	0	0	0.75 h	€ 5	0	0
Total	**6.25 h**	**€100**	**4.25 h**	**€100**	**6.5 h**	**€160**

7.2.1. Policy implications train load factor

The NS international intercity service between Amsterdam and Berlin is served 5 times per day (NS, n.d.). This train has a capacity of 600 passengers per train. This gives a total of potential 3000 passengers per day one way. With a ticket price of €100, the total revenue that can be generated is 3000*100 = €300.000. The WtP for a reduction in load factor of 1% for men is 0.08/3 hours*6.25hours = €0.167 per passenger; for women this is 0.14/3*6.25 = €0.292 per passenger. The calculations steps the same as for the London route, but now the values are used for the Berlin route. Table 7.6 shows again that the increased WtP for reduction in load factors also can not compensate for the loss in revenue due to this reduction in load factor. However, the differences in loss are smaller than for the London route.

Table 7.6: Revenues example Amsterdam - Berlin

Thalys + TGV Amsterdam - Barcelona				
	Revenue		WtP reduction LF	
Load factor	With WtP	Without WtP	Men	Women
25%	106641	75000	12.5	21.875
50%	204688	150000	8.333	14.583
75%	294141	225000	4.167	7.292
100%	300000	300000	0	0

7.2.2. Impact perceived risk

Table 7.7: Values example Amsterdam - Berlin, utilities & predicted market shares

Men		Values route Amsterdam - Berlin						Women		Values route Amsterdam - Berlin					
Situation		Utility			Market share			Situation		Utility			Market share		
PR train	PR plane	Train	Plane	Car	Train	Plane	Car	PR train	PR plane	Train	Plane	Car	Train	Plane	Car
1	1	-1.61	-2.73	-2.30	55%	18%	27%	1	1	-1.74	-2.82	-2.30	52%	18%	30%
1	2	-1.61	-2.83	-2.30	56%	16%	28%	1	2	-1.74	-3.02	-2.30	54%	15%	31%
1	3	-1.61	-2.94	-2.30	57%	15%	28%	1	3	-1.74	-3.22	-2.30	55%	13%	32%
1	4	-1.61	-3.05	-2.30	58%	14%	29%	1	4	-1.74	-3.42	-2.30	57%	11%	33%
1	5	-1.61	-3.15	-2.30	58%	12%	29%	1	5	-1.74	-3.62	-2.30	58%	9%	33%
2	1	-1.76	-2.73	-2.30	51%	19%	30%	2	1	-2.03	-2.82	-2.30	45%	21%	35%
2	2	-1.76	-2.83	-2.30	52%	18%	30%	2	2	-2.03	-3.02	-2.30	47%	17%	36%
2	3	-1.76	-2.94	-2.30	53%	16%	31%	2	3	-2.03	-3.22	-2.30	48%	15%	37%
2	4	-1.76	-3.05	-2.30	54%	15%	31%	2	4	-2.03	-3.42	-2.30	50%	12%	38%
2	5	-1.76	-3.15	-2.30	55%	14%	32%	2	5	-2.03	-3.62	-2.30	51%	10%	39%
3	1	-1.92	-2.73	-2.30	47%	21%	32%	3	1	-2.32	-2.82	-2.30	38%	23%	39%
3	2	-1.92	-2.83	-2.30	48%	19%	33%	3	2	-2.32	-3.02	-2.30	40%	20%	41%
3	3	-1.92	-2.94	-2.30	49%	18%	33%	3	3	-2.32	-3.22	-2.30	41%	17%	42%
3	4	-1.92	-3.05	-2.30	50%	16%	34%	3	4	-2.32	-3.42	-2.30	42%	14%	43%
3	5	-1.92	-3.15	-2.30	51%	15%	35%	3	5	-2.32	-3.62	-2.30	43%	12%	45%
4	1	-2.07	-2.73	-2.30	43%	22%	34%	4	1	-2.62	-2.82	-2.30	31%	26%	43%
4	2	-2.07	-2.83	-2.30	44%	21%	35%	4	2	-2.62	-3.02	-2.30	33%	22%	45%
4	3	-2.07	-2.94	-2.30	45%	19%	36%	4	3	-2.62	-3.22	-2.30	34%	19%	47%
4	4	-2.07	-3.05	-2.30	46%	17%	37%	4	4	-2.62	-3.42	-2.30	35%	16%	49%
4	5	-2.07	-3.15	-2.30	47%	16%	37%	4	5	-2.62	-3.62	-2.30	36%	13%	50%
5	1	-2.23	-2.73	-2.30	39%	24%	37%	5	1	-2.91	-2.82	-2.30	25%	28%	47%
5	2	-2.23	-2.83	-2.30	40%	22%	38%	5	2	-2.91	-3.02	-2.30	27%	24%	49%
5	3	-2.23	-2.94	-2.30	41%	20%	38%	5	3	-2.91	-3.22	-2.30	28%	21%	51%
5	4	-2.23	-3.05	-2.30	42%	19%	39%	5	4	-2.91	-3.42	-2.30	29%	17%	53%
5	5	-2.23	-3.15	-2.30	43%	17%	40%	5	5	-2.91	-3.62	-2.30	30%	15%	55%
				Average	49%	18%	33%					Average	41%	17%	42%

The impact of perceived risk on market share for the Amsterdam - Berlin route is again illustrated using the most extreme cases in differences of perceived risk level for both plane and train. When perceived risk is 1 for train and plane increases from 1 to 5, market share of train increases by 3% point for men; for women this is 6% point. Plane decreases by 6% point for men; for women this is 9% point. Car increases by 2% point for men; for women this is 3% point as well. The other way around, when train is increasing from 1 to 5 while plane stays at 1, results in a decrease of market share for train of 16% point for men; for women this is 27% point. Plane increases 6% point for men; for women it increase by 10% point. Car increases by 10% point for men; for women this is 17% point. The results show that the differences in market share for women are higher than the differences for men, due to the higher weight of perceived risk for women. Moreover, due to the higher travel time of train, the market share losses are higher for train than for plane. In this case, when the perceived risk of train is 3 and plane 2 (to approximate risk point 3.46 and 2.14) to train 2 and plane 3 the maximum market share difference possible is around 1-4% point for men; for women this is 2-7% point. This shows that for the Amsterdam - Berlin route there is a higher possible impact.

7.3. Route Amsterdam - Barcelona

The route Amsterdam - Barcelona is the second busiest route from Schiphol airport, with almost 1.4 million passengers in 2019 (Schiphol, 2020). For the train, it is possible to travel from Amsterdam to Barcelona by changing trains in Paris. The first part will be travelling on Thalys and the second part on the TGV. In the summertime, there is an extra possibility of travelling to Barcelona by train. In this case, the Thalys train will take one to Valence (France), and one has to switch trains to Barcelona. This trip is either on TGV or AVE (Spanish high-speed train). The third option is the car, but it has a long travel time of around 15 hours (ViaMichelin, n.d.). For this case, it is again discussed how perceived risk influences market share. Once more, for the calculations, Rome2Rio, ViaMichelin and Google Maps are used (Rome2Rio, n.d.; ViaMichelin, n.d.; Google, n.d.; Skyscanner, n.d.; NS, n.d.). Moreover, the websites of SNCF and Renfe are used for the calculation of travel cost of train (SNCF, n.d.; SNCF/Renfe, n.d.).

Table 7.8: Values example Amsterdam - Barcelona

	Amsterdam - Barcelona					
	Thalys + TGV		NS + Transavia		Hatchback E95	
	TT train	TC train	TT plane	TC plane	TT car	TC car
Access	0	0	0.25 h	€ 5	0	0
Wait	0	€ 5 (transfer)	2 h	0	0	€ 50 (toll)
Main	3.25 h + 6.75 h	€ 35 + €145	2.25 h	€ 65	15 h	€ 170 (fuel)
Egress	0.25 h	0	0.50 h	€ 5	0	0
Total	**10.25 h**	**€185**	**5 h**	**€75**	**15 h**	**€220**

7.3.1. Policy implications train load factor

The first part of this route is the Thalys to Paris, then the TGV service from Paris to Barcelona. The TGV service to Barcelona is offered two times per day. The Thalys train has a capacity of around 400 passengers (NS, n.d.). The TGV services to Barcelona uses TGV duplex train with a higher capacity. So the maximum number of passengers per day is 400*2 = 800 passengers. With a ticket price of €185, this generates 185*800 = €148.000. The WtP per percentage point reduction in load factor for men is 0.08/3 hours*10.25hours = €0.273 per passenger; for women this is 0.14/3*10.25 = €0.478 per passenger. For this longer route, the WtP values are substantially higher than for the London and Berlin routes. However, due to the high ticket price for this route, the WtP for reduction in load factor cannot compensate the loss in revenue. The differences between revenues for lower load factors decreases. This means that for lower load factors the difference between when people are willing to pay for reduction and when they are not willing to pay for this gets smaller. This is a logical consequence as for lower load factors, the need for a further reduction is lower.

Table 7.9: Revenues example Amsterdam - Barcelona

Thalys + TGV Amsterdam - Barcelona				
	Revenue		WtP reduction LF	
Load factor	With WtP	Without WtP	Men	Women
25%	42638	37000	20.50	35.88
50%	81517	74000	13.67	23.92
75%	116638	111000	6.83	11.96
100%	148000	148000	0	0

7.3.2. Impact perceived risk

Table 7.10: Values example Amsterdam - Barcelona, utilities & predicted market shares

Men		Values route Amsterdam - Barcelona					
Situation		Utility			Market share		
PR train	PR plane	Train	Plane	Car	Train	Plane	Car
1	1	-3.04	-2.84	-4.34	40%	49%	11%
1	2	-3.04	-2.97	-4.34	43%	46%	12%
1	3	-3.04	-3.09	-4.34	45%	43%	12%
1	4	-3.04	-3.22	-4.34	48%	40%	13%
1	5	-3.04	-3.35	-4.34	50%	37%	14%
2	1	-3.29	-2.84	-4.34	34%	54%	12%
2	2	-3.29	-2.97	-4.34	37%	51%	13%
2	3	-3.29	-3.09	-4.34	39%	47%	14%
2	4	-3.29	-3.22	-4.34	41%	44%	15%
2	5	-3.29	-3.35	-4.34	43%	41%	15%
3	1	-3.55	-2.84	-4.34	29%	58%	13%
3	2	-3.55	-2.97	-4.34	31%	55%	14%
3	3	-3.55	-3.09	-4.34	33%	52%	15%
3	4	-3.55	-3.22	-4.34	35%	49%	16%
3	5	-3.55	-3.35	-4.34	37%	46%	17%
4	1	-3.80	-2.84	-4.34	24%	62%	14%
4	2	-3.80	-2.97	-4.34	26%	59%	15%
4	3	-3.80	-3.09	-4.34	28%	56%	16%
4	4	-3.80	-3.22	-4.34	30%	53%	17%
4	5	-3.80	-3.35	-4.34	32%	50%	19%
5	1	-4.06	-2.84	-4.34	19%	66%	15%
5	2	-4.06	-2.97	-4.34	21%	63%	16%
5	3	-4.06	-3.09	-4.34	23%	60%	17%
5	4	-4.06	-3.22	-4.34	25%	57%	19%
5	5	-4.06	-3.35	-4.34	26%	54%	20%
			Average		34%	52%	15%

Women		Values route Amsterdam - Barcelona					
Situation		Utility			Market share		
PR train	PR plane	Train	Plane	Car	Train	Plane	Car
1	1	-3.26	-2.95	-4.34	37%	50%	13%
1	2	-3.26	-3.19	-4.34	41%	45%	14%
1	3	-3.26	-3.42	-4.34	46%	39%	16%
1	4	-3.26	-3.65	-4.34	50%	33%	17%
1	5	-3.26	-3.89	-4.34	53%	28%	18%
2	1	-3.74	-2.95	-4.34	27%	59%	15%
2	2	-3.74	-3.19	-4.34	30%	53%	17%
2	3	-3.74	-3.42	-4.34	34%	47%	19%
2	4	-3.74	-3.65	-4.34	38%	41%	21%
2	5	-3.74	-3.89	-4.34	42%	36%	23%
3	1	-4.21	-2.95	-4.34	18%	65%	16%
3	2	-4.21	-3.19	-4.34	21%	60%	19%
3	3	-4.21	-3.42	-4.34	24%	54%	22%
3	4	-4.21	-3.65	-4.34	27%	48%	24%
3	5	-4.21	-3.89	-4.34	31%	42%	27%
4	1	-4.69	-2.95	-4.34	12%	70%	18%
4	2	-4.69	-3.19	-4.34	14%	65%	21%
4	3	-4.69	-3.42	-4.34	17%	60%	24%
4	4	-4.69	-3.65	-4.34	19%	54%	27%
4	5	-4.69	-3.89	-4.34	21%	48%	31%
5	1	-5.17	-2.95	-4.34	8%	74%	18%
5	2	-5.17	-3.19	-4.34	9%	69%	22%
5	3	-5.17	-3.42	-4.34	11%	64%	25%
5	4	-5.17	-3.65	-4.34	13%	58%	29%
5	5	-5.17	-3.89	-4.34	14%	52%	33%
			Average		26%	53%	21%

For this route, perceived risk has a bigger influence on the market shares. This could be expected as perceived risk is dependent on time. When perceived risk is 1 for train and increases from 1 to 5 for plane, market share of train increases by 10% point for men; for women this is 16% point. Plane decreases by 12% point for men; for women plane decreases by 22% point. Car increases by 3% point for men; for women this is 5% point. The other way around, when train is increasing from 1 to 5 while plane stays at 1, results in a decrease of market share for train of 21% point for men; for women this is 29% point. For men plane increases by 17% point; for women this is 24% point. Car increases by 4% point for men; for women this is 5% point. Due to the long travel time of the car, the differences for car in market share are relatively small. Now, the market share losses of train are higher for train than for plane, due to the high travel time difference of plane and train; hence, perceived risk has a higher impact. However, these extreme changes are not realistic in real life. With 1.32

perceived risk level difference, the realistic effect on market share of perceived risk is maximum of 4-6% point for men; for women 3-16% point. The results show that for this longer route, perceived risk becomes more important. However, realistically, not all four mode-related attributes will be different. Airplanes (almost) always have HEPA-filters, trains as well. Furthermore, face mask policies for train and plane are often the same. Only load factor can be be significantly different between modes for the same OD-pairs. The most extreme difference in load factor results in a difference of perceived risk level of 0.646, this is approximately half of 1.32. In this case the maximum possible market share shift is 2-8% point.

7.3.3. Impact socio-demographics & preferences

This section will be about the impact of age and mode preferences on market shares on the Amsterdam - Barcelona route.

Impact age on modal split

In this case, the reference is the age of 20 to 40 years, as the average respondent is 38 years. Respondents younger than 20 years are more sensitive to cost. In this case, both train and car are more expensive than plane, so market share for plane increases by almost 10% point. Car has a smaller decrease in market share than train. For the 40 to 65 years group, respondents are less sensitive to costs than 20 to 40 years. As a consequence, market share for train increases by 6% point, car by 3% point and a decrease for plane by 9% point. For respondents older than 65, the cost sensitivity is even smaller, so a further increase in market share for train can be seen, with 8% point difference from the reference, for car 4% point difference from the reference. Plane decreases by 12% point. So it can be concluded that for an increase in age, there is an increase in market share for both train and car, and a decrease for plane.

Table 7.11: Impact age, utilities & predicted market shares

Situation	TT train	TC train	CF train	PR train	TT plane	TC plane	CF plane	PR plane	TT car	TC car	Train	Plane	Car	Train	Plane	Car
Reference	10.25	185	0	1	5	75	0	1	15	220	-3.04	-2.84	-4.34	40%	49%	11%
Younger than 20 years	10.25	185	0	1	5	75	0	1	15	220	-3.64	-3.09	-5.05	34%	58%	8%
20 to 40 years	10.25	185	0	1	5	75	0	1	15	220	-3.04	-2.84	-4.34	40%	49%	11%
40 to 65 years	10.25	185	0	1	5	75	0	1	15	220	-2.45	-2.61	-3.64	46%	40%	14%
65 years and older	10.25	185	0	1	5	75	0	1	15	220	-2.30	-2.54	-3.46	48%	37%	15%

Impact mode-preference on modal split

Within the survey, there was a question about the preference for a mode. The base (for these dummy coded variables) is bus. Not all interactions were significant. There was no effect found of preference for plane on the ASC for train and car. Also, there were no effects found of preference for car and train on the ASC of plane. It can be seen that a preference for train results in a increase in market share for train by 11% point, mainly on the cost of plane, which decreases by 8% point. Car decreases only by 3% point. Respondents with a preference for train only have a small negative effect on the utility of cars so; therefore, the effect on market share for car is moderate. A preference for plane results in a huge increase in market share for plane, by 24% point, mainly at the expense of train, which decreases by 19% point. Car decreases only by 5% point. A preference for car leads to a moderate increase in market share of 7% point for car. This is because the dis-utility of the long travel time and high costs can not be compensated by the preference utility. Plane stays about the same with a 1% point increase, and train decreases with 8% point. So it can be concluded that a preference for a certain mode has a huge impact on the resulting market shares.

Table 7.12: Impact mode-preference, utilities & predicted market shares

Situation	TT train	TC train	CF train	PR train	TT plane	TC plane	CF plane	PR plane	TT car	TC car	Train	Plane	Car	Train	Plane	Car
Reference	10.25	185	0	1	5	75	0	1	15	220	-3.04	-2.84	-4.34	40%	49%	11%
Preference train	10.25	185	0	1	5	75	0	1	15	220	-2.63	-2.84	-4.52	51%	41%	8%
Preference plane	10.25	185	0	1	5	75	0	1	15	220	-3.04	-1.80	-4.34	21%	73%	6%
Preference car	10.25	185	0	1	5	75	0	1	15	220	-3.30	-2.84	-3.84	32%	50%	18%

7.4. Summary

Main points of chapter

- Three routes were chosen to show impact on mode choice, Amsterdam - London, Amsterdam - Berlin & Amsterdam - Barcelona
- Women have a higher weight to perceived risk, hence the market share differences due to perceived risk is higher.
- Perceived risk has a small influence on market share, for longer travel times, perceived risk becomes more important.
- The WtP for reduction in load factor can not compensate the loss of revenue due to this reduction.
- Mode-preference has a quite substantial effect on modal split, age has a moderate effect on modal split.

For the routes Amsterdam - London, Amsterdam - Berlin and Amsterdam - Barcelona, it is discussed how perceived risk does influence market share. It can be concluded that for the short distance route Amsterdam - London, perceived risk only has a small impact on changes in the modal split. For the Amsterdam - Berlin and Amsterdam - Barcelona route, this effect is larger, but it has to be noticed that the difference in perceived risk for the rating experiment can only be 1.32 point for the same OD-pair. As a consequence, perceived risk has a small impact for this route as well (and even smaller for the Amsterdam - London route and Amsterdam - Berlin route). Mode preference has quite a substantial effect on modal split, age a moderate effect. The WtP values for reduction in load factor can not compensate for the loss in revenue in all of the three routes. So also for the longer routes, the effect of perceived risk is not strong enough.

8 Conclusion

This chapter provides the conclusion of this research. In section 8.1, the sub-questions are answered. Then in section 8.2, the main question is answered.

8.1. Key findings

COVID-19 did have (and in some countries still has) a huge impact on travel behaviour. Almost all countries took measures to reduce the number of infections, like bans on entering other countries, the closing of restaurants/bars and measures during the trip (like face masks or negative tests). One of the elements of travel behaviour is mode choice. It is determined by a lot of factors that are often interrelated to each other to a smaller or larger extent. It is often the consequence of a complex choice process that occurs automatically or deliberately in everyday life and incorporates both objective and subjective influences. Quite some research on the effects of mode choice has been done, mainly focusing on daily (short-distance) travel behaviour. Nevertheless, for long-distance travel, there has been little done research on the effects of COVID-19 (in terms of perceived risk).

To research the effects of perceived COVID-19 infection risk, this study was conducted. In total, 1147 respondents took part in this research. To study whether perceived COVID-19 risk influences mode choice, the survey consisted of two main parts, a part about socio-demographics and a part about travel behaviour. The first part of the survey was the rating experiment. In this part, respondents rated their perceived risk of infection with COVID-19 based on eight risk factors; four of these factors were destination-related, and the other four were mode-related. After this, respondents were faced with the main (mode) choice experiment. In this experiment, respondents chose between train, plane and car based on travel time, travel cost, travel comfort (travel class) and perceived risk (now a given value). To analyse the data, the discrete choice modelling theory was used. At last, an adapted variant of the Hierarchical Information Integration (HII) theory was used to combine both of the results. All of these steps were taken in order to answer the research questions. In the following sections, every sub-question is answered.

How do the different risk factors influence the perceived risk variable?. In total, eight factors were established: travel advice, vaccination rate, load factor, infections, air conditioning, travel requirements, face mask policy and cleaning policy. From the results, it can be concluded that the yellow and orange travel advice does not influence perceived risk. Nevertheless, red travel advice is the most important factor in the contribution to perceived risk. Vaccination rate and load factor do have about the same contribution to perceived risk, but the signs are opposite. Vaccination rate has a negative sign, load factor a positive sign. The number of infections has a counter-intuitive outcome, with a higher contribution for 10.000 positive tests per day than for 25.000 positive tests. A possible reason could be that respondents find it hard to imagine what this number means; moreover, the 25.000 level was almost insignificant with a p-value of 0.051. Nevertheless, the attribute of 100.000 positive tests has the highest contribution to the number of infections. For the ventilation/air conditioning variables, ventilation has the highest negative contribution. This can be explained as there was a focus on ventilation by the government to reduce the number of infections. After this, air conditioning with HEPA filters did have a smaller contribution than ventilation. Air conditioning without HEPA has the lowest contribution of these variables. Regarding travel requirements, the 3G policy has a positive contribution to perceived risk, which is counter-intuitive. This could be explained, as with the (first) introduction of the 3G-policy, there was a dramatic increase in number of infections. The 2G policy has no effects, but the 2G+ policy does decrease perceived risk. When looking at the face mask policy, having any mask is more important in the decrease of perceived risk than having at least an FFP2 mask. A surgical mask is not of importance. None of the cleaning variables have a contribution to perceived risk.

How do socio-demographic variables influence the perceived risk variable within the rating experiment?. For this question, gender, age, income, education and work status were added to the model. Work status does not have a contribution to perceived risk and thus does not influence perceived risk. Gender (woman = 1, man = 0) has a positive contribution to perceived risk; hence, being woman results in a higher (average) rating for perceived risk. The same counts for education levels' HBO' and 'WO'. Moreover, having an gross income between €20.000 and €100.000 per year also contributes to perceived risk. When looking at age, most of the age groups did not have a contribution to perceived risk; however, being between 20 and 40 years old leads to a decrease in perceived risk.

How do socio-demographic and travel behaviour variables influence travel cost, travel time, travel comfort and perceived risk in the main choice experiment?. To test whether socio-demographic and travel behaviour variables influence the main variables in the main choice experiment, interactions were estimated. One by one, a different interaction was estimated with the main variables. Then all of the interactions were added to the base model. After this, all interactions that were not significant were removed using a stepwise backwards elimination. Travel cost becomes less important with increasing age; this means that the cost parameter becomes less negative with increasing age. Moreover, if the payment is made by the educational institute or work, people become less sensitive to cost as well. For travel time, only travel purpose has an influence. Travelling for work contributes to a more negative weight on travel time, so travel time becomes more important. For travel comfort, no significant interactions were found. At last, for perceived risk, both gender and travel company turned out to be significant. Being a woman contributes to a higher weight to perceived risk; this was also found in the rating experiment. At last, travelling with friends contributes to perceived risk.

How are travel cost, travel time and travel comfort traded off against people's perception of COVID-19 risk?. Perceived risk is dependent on time, due to the significant interaction between perceived risk and travel time. The value of risk for a decrease in level of risk for a man is €3.89 per decrease in level per hour; for a woman, this is €7.24 per decrease of level per hour. The difference in gender was found due to a significant interaction between perceived risk and gender. The levels of risk go from 1-very low to 5-very high. As perceived risk is dependent on time, there is no trade-off in terms of travel time. Longer travel times leads to a higher perceived risk. These values are found for the average respondent. Different values are found for different ages, travel purposes, payment of the trip and education levels. In terms of travel comfort, the value of risk for men is 0.072 comfort 'points' for one level decrease of perceived risk per hour; for women this 0.134. An upgrade in travel class is worth 13.89 hours for men when perceived risk level is 1; for women this is 7.46 hours. For perceived risk level 5, an full increase in travel class is 'worth' only 2.78 hours of travel for men; for women 1.49 hours. This shows that perceived risk becomes more important at higher levels. The value of time found is 17.59 euro/hour for train, and 29.94 euro/hour for car. The willingness to pay for an increase in travel class is 53.40 euro. When work or the educational institution pays for the trip, the WtP becomes 81.22 euro.

8.2. Main research question

To what extent does risk perception of infection with COVID-19 influence mode choice for long- distance trips in Europe?. To test to what extent risk perception of COVID-19 influences market shares, examples with real-world values for the main choice experiment are established. The first example is the **Amsterdam - London** route. This route is the most popular destination for Schiphol in 2019. The maximum shift in market share is found when perceived risk for train increases from 1 to 5 and the plane stays at level 1. In this case, train loses 10% point market share for men; for women this is 17%. Plane has 3% point decrease in market share for men; for women this is 5% point. Car increases by 7% point for men; for women this is 12% point. Risk perception consists of several factors in real life, there is no such thing as risk levels in the real world. For this research, the rating experiment is based on eight factors. Only four of these factors are mode-related, the other four are destination-related. For the same OD-pair, these four factors thus do not change. As a consequence, the maximum level for risk is 3.46 points (calculated from the rating experiment). The minimum is 2.14 risk points, so the maximum difference is only 1.32 risk points. With this maximum difference, the realistic market share difference is 1-3% point. For the route **Amsterdam - Berlin** the maximum market share differences for this route are a bit higher. The change in perceived risk for train from level 1 to 5 while plane stays at 1, results in a decrease of market share for train of 16% point for men; for women this is 27% point. Plane increases 6% point for men; for women it increase by 10% point. Car increases by 10% point for men; for women this is 17% point. With the 1.32 perceived risk level difference possible, the maximum market share possible is around 2-3% point. For the **Amsterdam - Barcelona** route perceived risk has the biggest influence of the three routes on the market shares. This could be expected as perceived risk is dependent on time. The maximum market share difference for train is 21% point decrease for men; for women this is 29% point. For men plane increases by 17% point; for women this is 24% point. Car increases by 4% point for men; for women this is 5% point. With the 1.32 difference, market share differences are at maximum 4-16% point. For the three routes, the maximum possible market share difference show for London a small effect of perceived risk. For the Amsterdam - Berlin and Amsterdam - Barcelona routes, this effect becomes larger. However, realistically, not all four mode-related attributes will be different. Airplanes (almost) always have HEPA-filters, trains as well. Furthermore, face mask policies for train and plane are often the same. Only load factor can be be significantly different between modes for the same OD-pair. The most extreme difference in load factor results in a difference of perceived risk level of 0.646 (this is a 100% load factor difference), this is approximately half of 1.32. In this case the maximum possible market share shift is 2-8% point. **To conclude, risk perception of infection with COVID-19 has a moderate influence on market share.**

9
Limitations & suggestions for further research

This research does contain several limitations and simplifications but are essential when considering the results. First, the attributes will be discussed, then the experimental set-up, after this the collection of the data and at last, the application of the model. In the end, suggestions for further research will be mentioned.

Attributes
At the beginning of this report, several attributes were selected for both the rating experiment and the main choice experiment. The rating experiment is based on eight factors that together contribute to perceived risk. However, for respondents, other attributes could be important as well. For example, spacing between seats onboard the vehicle could be important. Travel cost, travel time (and comfort to a lesser extent) are the most important variables for mode choice. Because of the main choice experimental set-up, with perceived risk next to the other main choice attributes, travel cost, travel time and travel comfort, an overestimation of the importance of perceived risk is possible. However, from the conclusion, it can be seen that perceived risk, in general, has not a big influence on mode choice. For the main choice experiment, in an early phase, several attributes were eliminated and not taken into account. These attributes were access and egress time, transfers, reliability, frequency, waiting time and departure/arrival time. Adding these to the main choice experiment could theoretically lead to a better model, and thus less context is necessary. However, the choice experiment would probably be too complex for respondents. Moreover, as a more general approach is taken rather than going for a case study, it would almost be impossible to incorporate all these attributes. The main goal of this research was to investigate risk perception on COVID-19 and how this influences mode choice; therefore, having only four main attributes seems to be reasonable.

Experiment
For this experiment, it is chosen to separate the survey into two parts. To measure perceived risk, the rating experiment was presented to the respondents. It could also be chosen to include risk perception attributes only to the main choice experiment and not include the rating experiment. The rating of perceived risk was an integer, on a 5-points scale, based on the chosen attributes. The experiment is on an ordinal scale, whereas the rating is estimated on a ratio scale. There is some debate about using linear regression to estimate the rating because the respondents gave the answers on an ordinal scale. Further research is needed to investigate if is it fair to use a linear regression on a ratio scale when answers are given on an ordinal scale.

The way perceived risk is established in the rating experiment can be further elaborated. Even though the attributes in the rating experiment are precisely chosen, it is possible that these factors do not influence the potential decision in the respondent's life. It is possible that respondents would not even consider these attributes in real life, but because these attributes were presented to them, it is incorporated into the judgement of the respondent. These context framing concerns in experimental design are to some degree inevitable, but they certainly enhance the mismatch between model and reality. Another important limitation is the use of perceived risk variable in both the rating experiment and the main choice experiment. Respondents took part in both parts (rating and main choice); it is, therefore, possible for the respondent to rate perceived for all trips to 1-very low. As the perceived risk rating variable returns in the main (mode) choice experiment, the respondents could not be familiar with what it means if this value has a high value (so for example 5-very high). The related risk factors associated with such an unknown rating must therefore be imagined by the respondent, which cannot be supported by science. A solution would be presenting less extreme values for perceived risk in the main experiment and thereby reducing the range. However, this does not hold as well, as then less variation could be captured. The solution for this 'problem' is the use of pivot designs, in which the responses to the rating experiment are used to adjust the mode choice experiment possibilities in the subsequent experiment. In other words, the attribute levels in the main (mode) choice experiment should be chosen based on the ratings provided in the rating experiment. Besides that, the focus of the research is on how risk factors influence mode choice. However, half of the risk factors were destination-related. These destination-related attributes do not change for the same OD-pair. This is one of the reasons why there is not much influence of perceived risk. At the same time, only half of the attributes can be changed by the airlines or train companies. If more mode-related attributes were introduced, more policy measure recommendations could be taken.

Another limitation of the research is that it did not include an opting-out option. COVID-19 has several influences on travel behaviour. The main goal of this research was to test the effect of COVID-19 on the modal split for long-distance travel in Europe. Including an opting-out option (i.e., not travelling at all) would also allow for testing the reduction in ridership. However, it was chosen not to include this, as respondents would then be forced to choose a certain mode (so that trade offs between modes can be tested). Moreover, only train, car and plane were included as an option. In real life, more options are available, like the bus. Therefore, respondents may choose a certain mode, whereas, in real life, they would have chosen another mode. Nonetheless, this was an experimental set-up, and therefore it is not always the same as real life. Besides this, a survey is a snapshot. The moment of taking the survey was in February and the beginning of March. At that moment, there were still a lot of measures valid, while at this moment, there are no measures. As a consequence, choices are probably be different at this moment.

The most notable limitation is the use of the MNL model, which has some significant shortcomings. At first, it is assumed that the error term is of the Type 1 extreme value distribution. At second, the property assumption of Independence from Irrelevant alternatives (IIA) remains true. This indicates that the relative popularity of alternatives is independent of the popularity of others. This becomes a difficulty for this study since train choices and aircraft share unobserved (transit) characteristics. The third problem with the MNL model is that it overlooks the variability of the respondents' attribute weights. The fourth difficulty with MNL is the decision-maker's choices. The MNL implies independence from every choice, however in fact, the respondents' decisions are not independent of each other, resulting in the so-called 'panel' effect. Because of these reasons, the Mixed Logit model (ML model) would be better to estimate. The ML model is a model that can deal with these shortcomings stated above. By including an additional error term (σ), the model can account for correlations between error terms within the utility specification. The panel-ML model should probably used; this model corrects for the fourth (wrong) assumption of the MNL model mentioned above. Nevertheless, a model fit of 24% was found for the main choice experiment. Moreover, a latent class model could also be estimated, to look into different market share and also to test for heterogeneity.

There are also limitations regarding the modes themselves. It was assumed that people were travelling on legacy airlines such as KLM, Air France or Lufthansa and not with low-cost carriers like Ryanair or easyJet. In reality, these two airlines are one of the biggest in Europe and therefore are also very important for choosing modes for travelling within Europe. It was chosen not to go for these airlines, as sometimes very unrealistic prices are offered, like flying for €5. Often it is the result of discounts; in this case, it is even possible that people base their destination on the price rather than the destination itself. Moreover, for train, several options are available, such as high-speed train (HST) or night-train. In order to keep the survey simple and easily understandable, it was chosen to only go for conventional high-speed rail services or long-distance intercity trains.

Another possible shortcoming is the combination of attribute levels from the Ngene design. This design is orthogonal for the rating experiment and a D-efficient design for the main choice experiment. As a consequence of the design, sometimes respondents were presented with unrealistic combinations of values (e.g., a flight for 50€ in business class). It was possible to account for this by excluding certain value combinations, but then the design was not efficient anymore. The efficient design was more important than creating realistic combinations, so this is the reason to ignore the unrealistic values.

Collection of the data

The use of the NS panel to collect responses is an additional drawback of this research. While this contributed to a lot of responses (938) and, therefore, a lot of data for estimating the choice models, it also indicated that those who participated are mostly frequent train users in their daily lives and are likely among the most devoted train riders, given they are willing to spend their own time responding NS questionnaires. This suggests that the dataset may have a bias. This resulted in a larger market share for train. High income, high education, work status and age all have similar strong socio-demographic impacts. Normally, when conducting studies, there is often bias about the sample. This is due to the fact that students from the university are often in social groups with higher educated people (and most of the time, higher-income classes). Therefore bias arises about, for example, Willingness to Pay values. From the NS panel, it was expected that there was a better distribution of income and education. However, in the NS sample, there are also mostly higher incomes and higher educated people.

In this study, it is investigated if the perception of risk may substantially compete with other important attributes such as price, travel time, and comfort. Due to the lack of directly observable evidence, as there is limited research done on COVID-19 risk for long-distance travel, a stated-preference survey was conducted. However, there is always a discrepancy between what a person says he or she would do and what he or she will really do. Revealed preference is difficult to undertake since perceived risk is a complex variable with a lot of factors contributing to this. Consequently, a stated-preference study in survey format is the most practical alternative. The stated-preference experiment has some shortcomings. Either, it is possible that attributes are mentioned, and the respondents did not think about this. Alternatively, it is possible that attributes that might really play a role are not in the survey.

Model application

The collected data from this sample was very extensive. Because of this, it was not possible to test all the possible combinations of the outcomes. For example, the average respondent was introduced in order to calculate the Willingness to Pay values for the reduction of risk. However, if another combination of socio-demographic attributes and travel behaviour attributes were used, different values for WtP values would be found. A disadvantage in this case is there would be too much information. By using 'the average respondent', the interpretation of the results is easy and understandable, but it gives less information.

Suggestions for further research

Based on this research, several suggestions for further research can be made. First of all, because of the very broad approach that was taken, further research could focus on certain routes. In this case, it can be researched which factors are important for different case studies. Hence, more precise results can be used and could also be implemented. Moreover, the risk attributes were based mainly on literature. For further research, it is suggested also to include interviews and/or focus groups. In this case, it is likely to capture more of the relevant risk factors along the way. Crowding onboard was an important factor, and at the same time, it can be regulated by airlines or companies to a certain extent, so it is wise to explore the impact of various crowding levels more thoroughly. A similar research is suggested in another country, to compare the results. However, this is hard, as COVID-19 implications change a lot. However, revealed-preference data (now that the pandemic is at another stage) could be used to compare this with this data. In this case, the model can be calibrated with the revealed-preference data.

Bibliography

[1] Bh. Aaditya and T.M. Rahul. Psychological impacts of covid-19 pandemic on the mode choice behaviour: A hybrid choice modelling approach. *Transport Policy*, 108:47–58, 2021. ISSN 0967-070X. doi: https://doi.org/10.1016/j.tranpol.2021.05.003. URL `https://www.sciencedirect.com/science/article/pii/S0967070X21001359`.

[2] Borgar Aamaas, Jens Borken-Kleefeld, and Glen P. Peters. The climate impact of travel behavior: A german case study with illustrative mitigation options. *Environmental Science & Policy*, 33:273–282, 2013. ISSN 1462-9011. doi: https://doi.org/10.1016/j.envsci.2013.06.009. URL `https://www.sciencedirect.com/science/article/pii/S1462901113001366`.

[3] Mary Abboah-Offei, Yakubu Salifu, Bisi Adewale, Jonathan Bayuo, Rasheed Ofosu-Poku, and Edwina Beryl Addo Opare-Lokko. A rapid review of the use of face mask in preventing the spread of covid-19. *International journal of nursing studies advances*, 3:100013, 2021.

[4] Muhammad Abdullah, Charitha Dias, Deepti Muley, and Md. Shahin. Exploring the impacts of covid-19 on travel behavior and mode preferences. *Transportation Research Interdisciplinary Perspectives*, 8: 100255, 2020. ISSN 2590-1982. doi: https://doi.org/10.1016/j.trip.2020.100255. URL `https://www.sciencedirect.com/science/article/pii/S2590198220301664`.

[5] Paulus Teguh Aditjandra, Corinne Mulley, and John D Nelson. Neighbourhood design impact on travel behaviour: A comparison of us and uk experience. *Projections*, 2009.

[6] Daniel Kwasi Ahorsu, Chung-Ying Lin, Vida Imani, Mohsen Saffari, Mark D Griffiths, and Amir H Pakpour. The fear of covid-19 scale: development and initial validation. *International journal of mental health and addiction*, pages 1–9, 2020.

[7] Alberto García Álvarez. Energy consumption and emissions of high-speed trains. *Transportation research record*, 2159(1):27–35, 2010.

[8] Jillian Anable. 'complacent car addicts' or 'aspiring environmentalists'? identifying travel behaviour segments using attitude theory. *Transport policy*, 12(1):65–78, 2005.

[9] Theo A Arentze and Eric JE Molin. 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, 2013.

[10] Agnes Arnold-Forster. Polarisation, incivility, and scientific debate during covid-19—an essay by agnes arnold-forster. *bmj*, 374, 2021.

[11] Bilge Atasoy, Aurélie Glerum, and Michel Bierlaire. Attitudes towards mode choice in switzerland. *disP - The Planning Review*, 49(2):101–117, 2013. doi: 10.1080/02513625.2013.827518. URL `https://doi.org/10.1080/02513625.2013.827518`.

[12] Deutsche Bahn. Db konzern, n.d. URL `https://www.deutschebahn.com/de`.

[13] Kelvin Balcombe, Iain Fraser, and Liam Harris. Consumer willingness to pay for in-flight service and comfort levels: A choice experiment. *Journal of Air Transport Management*, 15(5):221–226, 2009.

[14] Raymond A Bauer. Consumer behavior as risk taking. In *Proceedings of the 43rd National Conference of the American Marketing Assocation, June 15, 16, 17, Chicago, Illinois, 1960*. American Marketing Association, 1960.

[15] Christiaan Behrens and Eric Pels. Intermodal competition in the london–paris passenger market: High-speed rail and air transport. *Journal of Urban Economics*, 71(3):278–288, 2012. ISSN 0094-1190. doi: https://doi.org/10.1016/j.jue.2011.12.005. URL `https://www.sciencedirect.com/science/article/pii/S0094119011000921`.

[16] Moshe Ben-Akiva and Michel Bierlaire. Discrete choice methods and their applications to short term travel decisions. In *Handbook of transportation science*, pages 5–33. Springer, 1999.

[17] Moshe Emanuel Ben-Akiva, Daniel McFadden, Kenneth Train, et al. *Foundations of stated preference elicitation: Consumer behavior and choice-based conjoint analysis*. Now, 2019.

[18] University of California Berkeley. Part i: Behavioral models, 2002. URL `https://eml.berkeley.edu/choice2/ch2.pdf`.

[19] Rajesh K Bhagat, MS Davies Wykes, Stuart B Dalziel, and PF Linden. Effects of ventilation on the indoor spread of covid-19. *Journal of Fluid Mechanics*, 903, 2020.

[20] Chandra R Bhat. Analysis of travel mode and departure time choice for urban shopping trips. *Transportation Research Part B: Methodological*, 32(6):361–371, 1998.

[21] Michel Bierlaire. A short introduction to pandasbiogeme. *A short introduction to PandasBiogeme*, 2020.

[22] Thierry Blayac and Patrice Bougette. Should i go by bus? the liberalization of the long-distance bus industry in france. *Transport Policy*, 56:50–62, 2017. ISSN 0967-070X. doi: https://doi.org/10.1016/j.tranpol.2017.03.004. URL `https://www.sciencedirect.com/science/article/pii/S0967070X16300750`.

[23] Maria Börjesson and Jonas Eliasson. Experiences from the swedish value of time study. *Transportation Research Part A: Policy and Practice*, 59:144–158, 2014.

[24] Brabants-Dagblad. Eerste prik tegen corona is gezet in veghel, gevaccineerde voelt zich 'supergoed', 2021. URL `https://www.bd.nl/brabant/eerste-prik-tegen-corona-is-gezet-in-veghel-gevaccineerde-voelt-zich-supergoed~a4a557e2/?referrer=https%3A%2F%2Fwww.google.nl%2F`.

[25] Richard Brown, Lynne Coventry, and Gillian Pepper. Covid-19 risk perceptions and their associations with related media consumption and personal experiences. *OSF Preprints*, pages 1–37, 2020.

[26] David Brownstone and Kenneth A. Small. Valuing time and reliability: assessing the evidence from road pricing demonstrations. *Transportation Research Part A: Policy and Practice*, 39(4):279–293, 2005. ISSN 0965-8564. doi: https://doi.org/10.1016/j.tra.2004.11.001. URL `https://www.sciencedirect.com/science/article/pii/S096585640400103X`. Connection Choice: Papers from the 10th IATBR Conference.

[27] Ralph Buehler. Determinants of transport mode choice: a comparison of germany and the usa. *Journal of Transport Geography*, 19(4):644–657, 2011. ISSN 0966-6923. doi: https://doi.org/10.1016/j.jtrangeo.2010.07.005. URL `https://www.sciencedirect.com/science/article/pii/S0966692310001171`.

[28] Ralph Buehler and Claudia Nobis. Travel behavior in aging societies: Comparison of germany and the united states. *Transportation research record*, 2182(1):62–70, 2010.

[29] BusinessAM. Wie business class vliegt, waant zich nu in economy, 2020. URL `https://businessam.be/wie-business-class-vliegt-waant-zich-nu-in-economy/`.

[30] Ennio Cascetta, Andrea Papola, Francesca Pagliara, and Vittorio Marzano. Analysis of mobility impacts of the high speed rome–naples rail link using withinday dynamic mode service choice models. *Journal of Transport Geography*, 19(4):635–643, 2011. ISSN 0966-6923. doi: https://doi.org/10.1016/j.jtrangeo.2010.07.001. URL `https://www.sciencedirect.com/science/article/pii/S0966692310001018`.

[31] CDC. Road travel toolkit for transportation partners, 2021. URL `https://www.cdc.gov/coronavirus/2019-ncov/travelers/transportation-toolkit.html`.

[32] Ji-Ming Chen. Novel statistics predict the covid-19 pandemic could terminate in 2022. *Journal of Medical Virology*, 2022.

[33] Yi-Tui Chen. The effect of vaccination rates on the infection of covid-19 under the vaccination rate below the herd immunity threshold. *International Journal of Environmental Research and Public Health*, 18(14):7491, 2021.

[34] Alex WH Chin, Julie TS Chu, Mahen RA Perera, Kenrie PY Hui, Hui-Ling Yen, Michael CW Chan, Malik Peiris, and Leo LM Poon. Stability of sars-cov-2 in different environmental conditions. *The Lancet Microbe*, 1(1):e10, 2020.

[35] Matteo Chinazzi, Jessica T Davis, Marco Ajelli, Corrado Gioannini, Maria Litvinova, Stefano Merler, Ana Pastore y Piontti, Kunpeng Mu, Luca Rossi, Kaiyuan Sun, et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (covid-19) outbreak. *Science*, 368(6489):395–400, 2020.

[36] Caspar Chorus. Choice behaviour modeling and the logit model, 2020. URL `https://brightspace.tudelft.nl/d2l/le/content/195054/viewContent/1371917/View`.

[37] Caspar G Chorus, John M Rose, and David A Hensher. Regret minimization or utility maximization: it depends on the attribute. *Environment and Planning B: Planning and Design*, 40(1):154–169, 2013.

[38] Derek K Chu, Elie A Akl, Stephanie Duda, Karla Solo, Sally Yaacoub, Holger J Schünemann, Amena El-harakeh, Antonio Bognanni, Tamara Lotfi, Mark Loeb, et al. Physical distancing, face masks, and eye protection to prevent person-to-person transmission of sars-cov-2 and covid-19: a systematic review and meta-analysis. *The lancet*, 395(10242):1973–1987, 2020.

[39] Paradee Chuaybamroong, Rotruedee Chotigawin, Sitthisuntorn Supothina, Pipat Sribenjalux, S Larpkiattaworn, and C-Y Wu. Efficacy of photocatalytic hepa filter on microorganism removal. *Indoor Air*, 20(3):246–254, 2010.

[40] Regina R. Clewlow, Joseph M. Sussman, and Hamsa Balakrishnan. The impact of high-speed rail and low-cost carriers on european air passenger traffic. *Transport Policy*, 33:136–143, 2014. ISSN 0967-070X. doi: https://doi.org/10.1016/j.tranpol.2014.01.015. URL `https://www.sciencedirect.com/science/article/pii/S0967070X14000304`.

[41] Elena Commodari. The role of sociodemographic and psychological variables on risk perception of the flu. *Sage Open*, 7(3):2158244017718890, 2017.

[42] Joyce M Dargay and Stephen Clark. The determinants of long distance travel in great britain. *Transportation Research Part A: Policy and Practice*, 46(3):576–587, 2012.

[43] Juan de Dios Ortúzar and Luis G Willumsen. *Modelling transport*. John wiley & sons, 2011.

[44] Mathijs de Haas, Roel Faber, and Marije Hamersma. How covid-19 and the dutch 'intelligent lockdown' change activities, work and travel behaviour: Evidence from longitudinal data in the netherlands. *Transportation Research Interdisciplinary Perspectives*, 6:100150, 2020. ISSN 2590-1982. doi: https://doi.org/10.1016/j.trip.2020.100150. URL `https://www.sciencedirect.com/science/article/pii/S2590198220300610`.

[45] André De Palma, Robin Lindsey, and Guillaume Monchambert. The economics of crowding in rail transit. *Journal of Urban Economics*, 101:106–122, 2017.

[46] Jonas De Vos. The effect of covid-19 and subsequent social distancing on travel behavior. *Transportation Research Interdisciplinary Perspectives*, 5:100121, 2020. ISSN 2590-1982. doi: https://doi.org/10.1016/j.trip.2020.100121. URL `https://www.sciencedirect.com/science/article/pii/S2590198220300324`.

[47] Astrid De Witte, Joachim Hollevoet, Frédéric Dobruszkes, Michel Hubert, and Cathy Macharis. Linking modal choice to motility: A comprehensive review. *Transportation Research Part A: Policy and Practice*, 49:329–341, 2013. ISSN 0965-8564. doi: https://doi.org/10.1016/j.tra.2013.01.009. URL `https://www.sciencedirect.com/science/article/pii/S0965856413000165`.

[48] Frédéric Dobruszkes, Catherine Dehon, and Moshe Givoni. Does european high-speed rail affect the current level of air services? an eu-wide analysis. *Transportation Research Part A: Policy and Practice*, 69:461–475, 2014. ISSN 0965-8564. doi: https://doi.org/10.1016/j.tra.2014.09.004. URL `https://www.sciencedirect.com/science/article/pii/S0965856414002158`.

[49] Yuanyuan Dong, XI Mo, Yabin Hu, Xin Qi, Fang Jiang, Zhongyi Jiang, and Shilu Tong. Epidemiological characteristics of 2143 pediatric patients with 2019 coronavirus disease in china. *Pediatrics*, 2020.

[50] Sarah Dryhurst, Claudia R Schneider, John Kerr, Alexandra LJ Freeman, Gabriel Recchia, Anne Marthe Van Der Bles, David Spiegelhalter, and Sander Van Der Linden. Risk perceptions of covid-19 around the world. *Journal of Risk Research*, 23(7-8):994–1006, 2020.

[51] Steffen E Eikenberry, Marina Mancuso, Enahoro Iboi, Tin Phan, Keenan Eikenberry, Yang Kuang, Eric Kostelich, and Abba B Gumel. To mask or not to mask: Modeling the potential for face mask use by the general public to curtail the covid-19 pandemic. *Infectious Disease Modelling*, 5:293–308, 2020.

[52] Eurostar. 25 dingen die je nog niet wist over eurostar, n.d. URL `https://www.eurostar.com/nl-nl/25jaar/25-weetjes-over-eurostar`.

[53] Armand Faganel. Quality perception gap inside the higher education institution. *International Journal of academic research*, 2(1):213–215, 2010.

[54] Eli P Fenichel, Nicolai V Kuminoff, and Gerardo Chowell. Skip the trip: Air travelers' behavioral responses to pandemic influenza. *PloS one*, 8(3):e58249, 2013.

[55] Melissa L Finucane, Paul Slovic, Chris K Mertz, James Flynn, and Theresa A Satterfield. Gender, race, and perceived risk: The'white male'effect. *Health, risk & society*, 2(2):159–172, 2000.

[56] Donna L Floyd, Steven Prentice-Dunn, and Ronald W Rogers. A meta-analysis of research on protection motivation theory. *Journal of applied social psychology*, 30(2):407–429, 2000.

[57] Nicola Gartland, David Fishwick, Anna Coleman, Karen Davies, Angelique Hartwig, Sheena Johnson, and Martie Van Tongeren. Transmission and control of sars-cov-2 on ground public transport: A rapid review of the literature to date. 2021.

[58] Nevine Labib Georggi and Ram M Pendyala. An analysis of long-distance travel behavior of the elderly and the low-income. Master's thesis, University of South Florida Tampa, FL, 2000.

[59] Lars Gerhold. Covid-19: risk perception and coping strategies. 2020.

[60] Moshe Givoni. Development and impact of the modern high-speed train: A review. *Transport Reviews*, 26(5):593–611, 2006. doi: 10.1080/01441640600589319. URL `https://doi.org/10.1080/01441640600589319`.

[61] Moshe Givoni and Frédéric Dobruszkes. A review of ex-post evidence for mode substitution and induced demand following the introduction of high-speed rail. *Transport Reviews*, 33(6):720–742, 2013. doi: 10.1080/01441647.2013.853707. URL `https://doi.org/10.1080/01441647.2013.853707`.

[62] Andreas Glöckner, Angela Rachael Dorrough, Tobias Wingen, and Simone Dohle. The perception of infection risks during the early and later outbreak of covid-19 in germany: Consequences and recommendations. 2020.

[63] Google. Google maps, n.d. URL `https://www.google.nl/maps`.

[64] Trisha Greenhalgh, Manuel B Schmid, Thomas Czypionka, Dirk Bassler, and Laurence Gruer. Face masks for the public during the covid-19 crisis. *Bmj*, 369, 2020.

[65] Guardian. Back to work: 'capacity of transport network will be down by 90%, 2020. URL `https://www.theguardian.com/world/2020/may/09/back-to-work-capacity-of-transport-network-will-be-down-by-90`.

[66] Per E Gustafsod. Gender differences in risk perception: Theoretical and methodological erspectives. *Risk analysis*, 18(6):805–811, 1998.

[67] Royal HaskoningDHV. Enhancing society together, n.d. URL `https://global.royalhaskoningdhv.com/`.

[68] David A Hensher and John M Rose. 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*, 41(5):428–443, 2007.

[69] Maogui Hu, Hui Lin, Jinfeng Wang, Chengdong Xu, Andrew J Tatem, Bin Meng, Xin Zhang, Yifeng Liu, Pengda Wang, Guizhen Wu, Haiyong Xie, and Shengjie Lai. Risk of Coronavirus Disease 2019 Transmission in Train Passengers: an Epidemiological and Modeling Study. *Clinical Infectious Diseases*, 72 (4):604–610, 07 2020. ISSN 1058-4838. doi: 10.1093/cid/ciaa1057. URL `https://doi.org/10.1093/cid/ciaa1057`.

[70] Elwood H Hunt and David R Space. The airplane cabin environment. *TSP*, 3(2):1–000, 1994.

[71] Italo. With the new hepa filters traveling with italo is even safer: Innovative ventilation system, n.d. URL `https://www.italotreno.it/en/the-train/hepa-filters`.

[72] Christopher I Jarvis, Kevin Van Zandvoort, Amy Gimma, Kiesha Prem, Petra Klepac, G James Rubin, and W John Edmunds. Quantifying the impact of physical distance measures on the transmission of covid-19 in the uk. *BMC medicine*, 18(1):1–10, 2020.

[73] Juan Luis Jiménez and Ofelia Betancor. When trains go faster than planes: The strategic reaction of airlines in spain. *Transport Policy*, 23:34–41, 2012. ISSN 0967-070X. doi: https://doi.org/10.1016/j.tranpol.2012.06.003. URL `https://www.sciencedirect.com/science/article/pii/S0967070X1200087X`.

[74] Maria Vredin Johansson, Tobias Heldt, and Per Johansson. The effects of attitudes and personality traits on mode choice. *Transportation Research Part A: Policy and Practice*, 40(6):507–525, 2006.

[75] R Johnson and B Orme. Sample size issues for conjoint analysis. *Getting started with conjoint analysis: strategies for product design and pricing research. Madison: Research Publishers LLC*, pages 57–66, 2010.

[76] Ankur Joshi, Saket Kale, Satish Chandel, and D Kumar Pal. Likert scale: Explored and explained. *British journal of applied science & technology*, 7(4):396, 2015.

[77] Günter Kampf, Daniel Todt, Stephanie Pfaender, and Eike Steinmann. Persistence of coronaviruses on inanimate surfaces and their inactivation with biocidal agents. *Journal of hospital infection*, 104(3): 246–251, 2020.

[78] Linda C. Karlsson, Anna Soveri, Stephan Lewandowsky, Linnea Karlsson, Hasse Karlsson, Saara Nolvi, Max Karukivi, Mikael Lindfelt, and Jan Antfolk. Fearing the disease or the vaccine: The case of covid-19. *Personality and Individual Differences*, 172:110590, 2021. ISSN 0191-8869. doi: https://doi.org/10.1016/j.paid.2020.110590. URL `https://www.sciencedirect.com/science/article/pii/S0191886920307819`.

[79] Aisha N Khatib, Anna-Maria Carvalho, Robert Primavesi, Kent To, and Vincent Poirier. Navigating the risks of flying during covid-19: a review for safe air travel. *Journal of travel medicine*, 27(8):taaa212, 2020.

[80] KiM. Een nieuwe kijk op de value of time!?, 2016.

[81] KiM. Kengetallen bereikbaarheid, 2020. URL `https://www.rwseconomie.nl/kengetallen/kengetallen-bereikbaarheid-map`.

[82] Chansung Kim, Seung Hoon Cheon, Keechoo Choi, Chang-Hyeon Joh, and Hyuk-Jin Lee. Exposure to fear: Changes in travel behavior during mers outbreak in seoul. *KSCE Journal of Civil Engineering*, 21 (7):2888–2895, 2017.

[83] KLM. Reizen en het coronavirus. URL `https://www.klm.nl/information/covid-19`.

[84] KLM. Uw reis naar london, 2022. URL `https://www.klm.nl/search/summary?pax=1:0:0:0:0:0:0:0&cabinClass=BUSINESS&activeConnection=1&connections=AMS:A:20220906@2030:KL1033:0:OS5VPBNH:BUSINESS%3ELHR:A-LHR:A:20220909@2025:KL1032:0:OS5VPBND:BUSINESS%3EAMS:A&bookingFlow=LEISURE`.

[85] Abhiteja Konda, Abhinav Prakash, Gregory A Moss, Michael Schmoldt, Gregory D Grant, and Supratik Guha. Aerosol filtration efficiency of common fabrics used in respiratory cloth masks. *ACS nano*, 14(5): 6339–6347, 2020.

[86] Marcel Korbee. Ialles wat je moet weten over luchtfiltersystemen en het schoonmaken van een vliegtuig, 2020. URL `https://www.skyscanner.nl/nieuws/luchtfiltersystemen-in-vliegtuigen-en-schoonmaken-vliegtuig`.

[87] Marco Kouwenhoven, Gerard C de Jong, Paul Koster, Vincent AC van den Berg, Erik T Verhoef, John Bates, and Pim MJ Warffemius. New values of time and reliability in passenger transport in the netherlands. *Research in Transportation Economics*, 47:37–49, 2014.

[88] Panchamy Krishnakumari and Oded Cats. Virus spreading in public transport networks, 05 2020.

[89] Maarten Kroesen. Side panelexpand side panel sen1721 introduction 2020-2021, 2020.

[90] Tracy L Lamb, Scott R Winter, Stephen Rice, Keith J Ruskin, and Austin Vaughn. Factors that predict passengers willingness to fly during and after the covid-19 pandemic. *Journal of air transport management*, 89:101897, 2020.

[91] Anja Leppin and Arja R Aro. Risk perceptions related to sars and avian influenza: theoretical foundations of current empirical research. *International journal of behavioral medicine*, 16(1):7–29, 2009.

[92] Zheng Li and David A Hensher. Crowding and public transport: A review of willingness to pay evidence and its relevance in project appraisal. *Transport Policy*, 18(6):880–887, 2011.

[93] Narisra Limtanakool, Martin Dijst, and Tim Schwanen. 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, 2006.

[94] Jun Liu, Steven E Moss, and Jonathan Zhang. The life cycle of a pandemic crisis: Sars impact on air travel. *Journal of International Business Research*, 10(2):63, 2011.

[95] George F Loewenstein, Elke U Weber, Christopher K Hsee, and Ned Welch. Risk as feelings. *Psychological bulletin*, 127(2):267, 2001.

[96] Jordan J Louviere. Hierarchical information integration: A new method for the design and analysis of complex multlattribute judgment problems. *ACR North American Advances*, 1984.

[97] Jianyun Lu, Jieni Gu, Kuibiao Li, Conghui Xu, Wenzhe Su, Zhisheng Lai, Deqian Zhou, Chao Yu, Bin Xu, and Zhicong Yang. Covid-19 outbreak associated with air conditioning in restaurant, guangzhou, china, 2020. *Emerging infectious diseases*, 26(7):1628, 2020.

[98] Tafadzwa Matiza. Post-covid-19 crisis travel behaviour: towards mitigating the effects of perceived risk. *Journal of Tourism Futures*, 2020.

[99] Daniel McFadden et al. Conditional logit analysis of qualitative choice behavior. 1973.

[100] Gaëtan Mertens, Lotte Gerritsen, Stefanie Duijndam, Elske Salemink, and Iris M Engelhard. Fear of the coronavirus (covid-19): Predictors in an online study conducted in march 2020. *Journal of anxiety disorders*, 74:102258, 2020.

[101] Maria Milosh, Marcus Painter, Konstantin Sonin, David Van Dijcke, and Austin L Wright. Political polarisation impedes the public policy response to covid-19. *The Economist*, 2020.

[102] Rolf Moeckel, Rhett Fussell, and Rick Donnelly. Mode choice modeling for long-distance travel. *Transportation Letters*, 7(1):35–46, 2015. doi: 10.1179/1942787514Y.0000000031. URL `https://doi.org/10.1179/1942787514Y.0000000031`.

[103] BE Moen, L Sjöberg, and T Rundmo. Explaining risk perception. an evaluation of the psychometric paradigm in risk perception research. *Editor: Torbjørn Rundmo, Trondheim, Norway*, 2004.

[104] Eric Molin. Introduction to experimental designs lecture 1 (sen1221 – part ii), 2020. URL https://brightspace.tudelft.nl/d2l/le/content/195054/viewContent/1371905/View.

[105] Eric Molin. Complex variables & non-linearity, 2020. URL https://brightspace.tudelft.nl/d2l/le/content/195054/viewContent/1371927/View.

[106] Eric Molin and Michel van Gelder. Freeway access to public transport: A hierarchical multimodal choice model. *Transportation Research Record*, 2076(1):106–113, 2008. doi: 10.3141/2076-12. URL https://doi.org/10.3141/2076-12.

[107] Eric Molin, Patricia Mokhtarian, and Maarten Kroesen. Multimodal travel groups and attitudes: A latent class cluster analysis of dutch travelers. *Transportation Research Part A: Policy and Practice*, 83:14–29, 2016. ISSN 0965-8564. doi: https://doi.org/10.1016/j.tra.2015.11.001. URL https://www.sciencedirect.com/science/article/pii/S0965856415002591.

[108] Eric Molin, Joey Blangé, Oded Cats, and Caspar Chorus. Willingness to pay for safety improvements in passenger air travel. *Journal of Air Transport Management*, 62:165–175, 2017. ISSN 0969-6997. doi: https://doi.org/10.1016/j.jairtraman.2017.04.002. URL https://www.sciencedirect.com/science/article/pii/S0969699716304756.

[109] Lidia Morawska and Junji Cao. Airborne transmission of sars-cov-2: The world should face the reality. *Environment international*, 139:105730, 2020.

[110] Pedro Moreira. Stealth risks and catastrophic risks: On risk perception and crisis recovery strategies. *Journal of Travel & Tourism Marketing*, 23(2-4):15–27, 2008.

[111] Taka Morikawa, Moshe Ben-Akiva, and Daniel McFadden. Discrete choice models incorporating revealed preferences and psychometric data. In *Advances in Econometrics*. Emerald Group Publishing Limited, 2002.

[112] Niek Mouter, Jose Ignacio Hernandez, and Anatol Valerian Itten. Public participation in crisis policymaking. how 30,000 dutch citizens advised their government on relaxing covid-19 lockdown measures. *PloS one*, 16(5):e0250614, 2021.

[113] Charles Musselwhite, Erel Avineri, and Yusak Susilo. Editorial jth 16–the coronavirus disease covid-19 and implications for transport and health. *Journal of Transport & Health*, 16:100853, 2020.

[114] Larissa Neuburger and Roman Egger. Travel risk perception and travel behaviour during the covid-19 pandemic 2020: A case study of the dach region. *Current Issues in Tourism*, 24(7):1003–1016, 2021.

[115] CDC Newsroom. Cdc updates covid-19 transmission webpage to clarify information about types of spread, 2020.

[116] Marta Nieto-García, Pablo A Muñoz-Gallego, and Óscar Gonzalez-Benito. The more the merrier? understanding how travel frequency shapes willingness to pay. *Cornell Hospitality Quarterly*, 61(4):401–415, 2020.

[117] NS. Uw reis naar london, 2022. URL https://www.nsinternational.com/nl/treintickets-v3/#/book/NXSCFMF/88b319fdded0c780f7bfc83ee0e7b77dc6137fadb8d70f8c6f59c449abc26ac3/overview.

[118] NS. Ns international, n.d.. URL https://www.nsinternational.com/.

[119] NS. Plattegronden rijtuigen thalys, n.d.. URL https://www.nsinternational.com/nl/treinen/thalys/plattegronden-rijtuigen-thalys.

[120] Timo Ohnmacht and Milena Scherer. More comfort, shorter travel time, or low fares? comparing rail transit preferences of commuters, holiday and leisure travelers, business travelers, and shoppers in switzerland. *Transportation research record*, 2143(1):100–107, 2010.

[121] Marcel Paulssen, Dirk Temme, Akshay Vij, and Joan L Walker. Values, attitudes and travel behavior: a hierarchical latent variable mixed logit model of travel mode choice. *Transportation*, 41(4):873–888, 2014.

[122] Lori Pennington-Gray, Ashley Schroeder, and Kiki Kyriaki Kaplanidou. Examining the influence of past travel experience, general web searching behaviors, and risk perceptions on future travel intentions. *IJSSTH*, 1(1):64–89, 2011.

[123] J Paul Peter and Michael J Ryan. An investigation of perceived risk at the brand level. *Journal of marketing research*, 13(2):184–188, 1976.

[124] Abraham Pizam and Aliza Fleischer. Severity versus frequency of acts of terrorism: Which has a larger impact on tourism demand? *Journal of Travel research*, 40(3):337–339, 2002.

[125] Alexander Reichert and Christian Holz-Rau. Mode use in long-distance travel. *Journal of Transport and Land Use*, 8(2):87–105, 2015. ISSN 19387849. URL `http://www.jstor.org/stable/26202714`.

[126] Yvette Reisinger and Felix Mavondo. Travel anxiety and intentions to travel internationally: Implications of travel risk perception. *Journal of travel research*, 43(3):212–225, 2005.

[127] Rijksoverheid. Reisadviezen, 2021. URL `https://www.nederlandwereldwijd.nl/reizen/reisadviezen`.

[128] Bongkosh Ngamsom Rittichainuwat and Goutam Chakraborty. Perceived travel risks regarding terrorism and disease: The case of thailand. *Tourism management*, 30(3):410–418, 2009.

[129] RIVM. 1,5 meter afstand houden en drukte vermijden, 2020. URL `https://www.rivm.nl/documenten/anderhalve-meter-afstand-houden-en-drukte-vermijden`.

[130] RIVM. Number of confirmed cases, 2021. URL `https://coronadashboard.government.nl/landelijk/positief-geteste-mensen`.

[131] RIVM. Werking en bescherming, 2021. URL `https://www.rivm.nl/covid-19-vaccinatie/vragen-achtergronden/werking`.

[132] RIVM. Aantal positief geteste mensen door de tijd heen (per 100.000 inwoners), 2022. URL `https://coronadashboard.rijksoverheid.nl/landelijk/positief-geteste-mensen`.

[133] Ronald W Rogers. A protection motivation theory of fear appeals and attitude change1. *The journal of psychology*, 91(1):93–114, 1975.

[134] Rome2Rio. Rome2rio: Discover how to get anywhere by plane, train, bus, ferry car, n.d. URL `https://www.rome2rio.com/`.

[135] Concepción Román, Raquel Espino, and Juan Carlos Martín. 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, 2010. ISSN 1755-5345. doi: https://doi.org/10.1016/S1755-5345(13)70030-7. URL `https://www.sciencedirect.com/science/article/pii/S1755534513700307`.

[136] Narendra Saini and Vandana Saini. Air travel in covid-19 pandemic. *Journal of Patient Safety and Infection Control*, 8(1):29–29, 2020.

[137] Schiphol. Traffic review 2019, 2020. URL `https://www.annualreportschiphol.com/trafficreview2019/summary`.

[138] Ashley Schroeder, Lori Pennington-Gray, Kiki Kaplanidou, and Fangzi Zhan. Destination risk perceptions among us residents for london as the host city of the 2012 summer olympic games. *Tourism Management*, 38:107–119, 2013.

[139] Amin Shaer and Hossein Haghshenas. Evaluating the effects of the covid-19 outbreak on the older adults' travel mode choices. *Transport Policy*, 112:162–172, 2021. ISSN 0967-070X. doi: https://doi.org/10.1016/j.tranpol.2021.08.016. URL `https://www.sciencedirect.com/science/article/pii/S0967070X21002559`.

[140] Ali Shamshiripour, Ehsan Rahimi, Ramin Shabanpour, and Abolfazl (Kouros) Mohammadian. How is covid-19 reshaping activity-travel behavior? evidence from a comprehensive survey in chicago. *Transportation Research Interdisciplinary Perspectives*, 7:100216, 2020. ISSN 2590-1982. doi: https://doi.org/10.1016/j.trip.2020.100216. URL https://www.sciencedirect.com/science/article/pii/S2590198220301275.

[141] S Shelat, O Cats, and S Van Cranenburgh. Avoiding the crowd: How do passengers trade-off time and crowding in the age of covid-19. Technical report, Working paper, 2020.

[142] Ye Shen, Changwei Li, Hongjun Dong, Zhen Wang, Leonardo Martinez, Zhou Sun, Andreas Handel, Zhiping Chen, Enfu Chen, Mark H Ebell, et al. Community outbreak investigation of sars-cov-2 transmission among bus riders in eastern china. *JAMA internal medicine*, 180(12):1665–1671, 2020.

[143] Jeremy D Shires and Gerard C De Jong. An international meta-analysis of values of travel time savings. *Evaluation and program planning*, 32(4):315–325, 2009.

[144] Skyscanner. Skyscanner, n.d. URL https://www.skyscanner.nl/.

[145] Paul Slovic. Perception of risk: Reflections on the psychometric paradigm. 1992.

[146] SNCF. Sncf, n.d.. URL https://www.sncf.com/fr.

[147] SNCF. Covid-19 and climate control, n.d.. URL https://www.sncf.com/en/passenger-offer/travel-by-train/covid-19-and-air-conditioning.

[148] SNCF/Renfe. Travel between france and spain, n.d. URL https://www.renfe-sncf.com/.

[149] Joseph B. Sobieralski. Covid-19 and airline employment: Insights from historical uncertainty shocks to the industry. *Transportation Research Interdisciplinary Perspectives*, 5:100123, 2020. ISSN 2590-1982. doi: https://doi.org/10.1016/j.trip.2020.100123. URL https://www.sciencedirect.com/science/article/pii/S2590198220300348.

[150] Sevil F Sönmez and Alan R Graefe. Influence of terrorism risk on foreign tourism decisions. *Annals of tourism research*, 25(1):112–144, 1998.

[151] SurveyMonkey. Het aantal respondenten dat u nodig hebt berekenen, 2021. URL https://help.surveymonkey.com/articles/nl_NL/kb/How-many-respondents-do-I-need.

[152] Song Tang, Yixin Mao, Rachael M Jones, Qiyue Tan, John S Ji, Na Li, Jin Shen, Yuebin Lv, Lijun Pan, Pei Ding, et al. Aerosol transmission of sars-cov-2? evidence, prevention and control. *Environment international*, 144:106039, 2020.

[153] Steven Taylor, Caeleigh A Landry, Michelle M Paluszek, Thomas A Fergus, Dean McKay, and Gordon JG Asmundson. Development and initial validation of the covid stress scales. *Journal of Anxiety Disorders*, 72:102232, 2020.

[154] D Thompson. Hygiene theater is a huge waste of time. *The Atlantic. July*, 27, 2020.

[155] Alejandro Tirachini and Oded Cats. Covid-19 and public transportation: Current assessment, prospects, and research needs. *Journal of Public Transportation*, 22(1):1, 2020.

[156] Alejandro Tirachini, David A Hensher, and John M Rose. Crowding in public transport systems: effects on users, operation and implications for the estimation of demand. *Transportation research part A: policy and practice*, 53:36–52, 2013.

[157] Treinenweb. Eurostar londen - amsterdam start 4 april, 2018. URL https://www.treinenweb.nl/nieuws/6875/eurostar-londen-amsterdam-start-4-april.html.

[158] Joy Troko, Puja Myles, Jack Gibson, Ahmed Hashim, Joanne Enstone, Susan Kingdon, Christopher Packham, Shahid Amin, Andrew Hayward, and Jonathan Nguyen Van-Tam. Is public transport a risk factor for acute respiratory infection? *BMC infectious diseases*, 11(1):1–6, 2011.

[159] European Union. Re-open eu, 2021. URL https://reopen.europa.eu/nl.

[160] Sander Van Cranenburgh, Caspar Chorus, and Bert Van Wee. Substantial changes and their impact on mobility: a typology and an overview of the literature. *Transport Reviews*, 32(5):569–597, 2012.

[161] Thijs van de Wiel. Impacts of covid-19 risk perceptions on train travel decisions: A hierarchical information integration analysis. 2021.

[162] Sander Van Der Drift, Luc Wismans, and Marie-José Olde Kalter. Changing mobility patterns in the netherlands during covid-19 outbreak. *Journal of Location Based Services*, pages 1–24, 2021.

[163] Neeltje Van Doremalen, Trenton Bushmaker, Dylan H Morris, Myndi G Holbrook, Amandine Gamble, Brandi N Williamson, Azaibi Tamin, Jennifer L Harcourt, Natalie J Thornburg, Susan I Gerber, et al. Aerosol and surface stability of sars-cov-2 as compared with sars-cov-1. *New England journal of medicine*, 382(16):1564–1567, 2020.

[164] Cees D van Goeverden. Explaining factors for train use in european long-distance travel. *Tourism and Hospitality Planning & Development*, 6(1):21–37, 2009.

[165] Ruben Van Loon and Jan Rouwendal. Vacation behaviour: Frequency, destination choice and expenditures. *Tourism Economics*, 19(5):1005–1026, 2013.

[166] Bert van Wee and Frank Witlox. Covid-19 and its long-term effects on activity participation and travel behaviour: A multiperspective view. *Journal of transport geography*, 95:103144, 2021.

[167] Thirumalaisamy P Velavan and Christian G Meyer. The covid-19 epidemic. *Tropical medicine & international health*, 25(3):278, 2020.

[168] ViaMichelin. Viamichelin: Routes, kaarten, verkeersinformatie, hotels, n.d. URL `https://www.viamichelin.nl/`.

[169] Kuo-Ying Wang. How change of public transportation usage reveals fear of the sars virus in a city. *PloS one*, 9(3):e89405, 2014.

[170] Mark Wardman. A comparison of revealed preference and stated preference models of travel behaviour. *Journal of transport economics and policy*, pages 71–91, 1988.

[171] Neil D Weinstein. The precaution adoption process. *Health psychology*, 7(4):355, 1988.

[172] Zhang Wen, Gu Huimin, and Raphael R Kavanaugh. The impacts of sars on the consumer behaviour of chinese domestic tourists. *Current Issues in Tourism*, 8(1):22–38, 2005.

[173] J Frank Yates and Eric R Stone. The risk construct. 1992.

A
Scientific paper

The influence of perceived COVID-19 risk on the modal-split for long-distance travel in Europe: a Hierarchical Information Integration and Stated-Preference study approach

M.M. van Dalen

Faculty of Civil Engineering & Geosciences, Delft University of Technology, Stevinweg 1, Delft, 2628 CN, The Netherlands.

Contributing authors: M.M.vanDalen@student.tudelft.nl;

Abstract

COVID-19 had (and still has) a huge impact on travel behaviour. Most research on the effects of COVID-19 on travel behaviour is based on daily travel behaviour. Long-distance travel is most of the time not included. Therefore, this paper studies the impact of perceived COVID-19 risk on mode-choice for long-distance travel in Europe. This paper uses a Stated-Preference approach which consists of two parts, the perceived risk rating experiment and the main mode choice experiment. In total 1147 responses were collected. In the rating experiment, respondents rated their perceived risk based on eight attributes. In the main choice experiment, respondents chose between train, plane and care based on travel time, travel cost, travel comfort and perceived risk. With this, the Value of Risk (VoR) for a decrease in perceived risk is derived, both expressed in travel cost and travel comfort. Moreover, the Willingness to Pay (WtP) for comfort and Value of Time (VoT) are derived. With the combination of both models, the WtP for risk attributes given a perceived risk level are derived. The combination of both models shows the Willingness to Pay values for the different risk attributes, given a certain perceived risk level. To test the influence of perceived risk on modal split, three routes with real world values are used. The results implicate a maximum of 5% market share difference possible. This shows that perceived COVID-19 has a moderate effect on modal split.

Keywords: Sated-Preference, Perceived COVID-19 risk, MNL model, Discrete Choice Modelling, Hierarchical Information Integration (HII), long-distance travel, risk

1 Introduction

The new coronavirus (SARS-CoV-2) (from now: COVID-19) is an ongoing global pandemic that is considered one of the worst post-World War II pandemics that affects the world surpassing both MERS and SARS outbreaks (Matiza, 2020). Based on rising infection rates in China and then to over the world, the WHO Emergency Committee declared a global health emergency on January 30, 2020 (Velavan and Meyer, 2020). Countries have taken several measures to reduce the number of infections; some of these measures were restrictions from entering other countries, the closing of restaurants/bars and measures during the trip (like a face mask or negative test). Because of these measures, COVID-19 has a huge impact on travel behaviour (De Vos, 2020).

One of the elements of travel behaviour is mode choice. It is determined by a lot of factors that are often interrelated to each other to a smaller or larger extent (De Witte, Hollevoet, Dobruszkes, Hubert, and Macharis, 2013). Quite some research is done on the effects of mode-choice, mainly focusing on daily (short-distance) travel behaviour (Buehler, 2011; Atasoy, Glerum, and Bierlaire, 2013; De Witte et al., 2013). However, long-distance travel behaviour is often excluded from the analysis. At the same time, over 50% of the passenger-kilometers travelled come from long-distance travel (Aamaas, Borken-Kleefeld, and Peters, 2013). Several studies researched the main drive for passengers to choose a certain mode for long-distance travel in Europe. The main finding of these studies was that travel time seems to be the most important factor (van Goeverden, 2009). Important other factors are travel cost, comfort, reliability, access & egress time and number of interchanges (Román, Espino, and Martín, 2010).

Several papers researched the effects of COVID-19 infection on mode-choice, but the focus was on short-distance travel. Abdullah, Dias, Muley, and Shahin (2020) found that people tend to use less public transport services and more private cars during the pandemic. There was a shift found to active modes as well. Moreover, sanitisation measures and social distancing characterise the perception of safety and, therefore, the willingness to use public transport services (Aaditya and Rahul, 2021). According to de Haas, Faber, and Hamersma (2020) in the Netherlands, around 80% of the people reduced their activities, with an increase from 6% of the people working at home to 39%. For public transport, there was a drop in usage of 90%. Shamshiripour, Rahimi, Shabanpour, and Mohammadian (2020) noticed that people tend to shift more to individual and active modes (e.g., walking or cycling) of travel or not travelling at all. Another possible reason for the reduction of travel by public transport is that people believe that public transport is an unhygienic place, with a high chance of infection with viruses (Troko et al., 2011).

At this stage, most of the COVID-19 measures are gone. Partly this is a consequence of the omicron-variant of the COVID-19 virus (Chen, 2022). Thus, lower hospitalisation rates can be found. The Dutch society accepted to 'live' with COVID-19 because COVID-19 will not fully be gone. By definition, a lot of different people are transported at the same time in PT systems; therefore, the virus can relatively easy be transmitted among travellers. To illustrate this, Krishnakumari and Cats (2020) found that, on average, a person interacts with 1200 other people on a single trip in the metro system network of Washington D.C. Now, it seems evident that measures were (and still are) taken in order to reduce the number of infections and mitigate the public health crisis.

To study the (risk) effects of COVID-19 on mode-choice, a new variable is introduced: perceived risk. Perceived risk is the risk people perceive during their trip of getting infected with COVID-19. It is often defined as the perceived likelihood of getting the disease times the perceived severity of the symptoms (Karlsson et al., 2021). For this study, however, perceived risk consists on mode-related and destination-related attributes. Perceived risk does also have an emotional dimension like worry and fear (Loewenstein, Weber, Hsee, and Welch, 2001). Potentially, people will change their travel behaviour (and thus mode-choice) due to these factors as well as other factors of their trip. The outcome of this study could be beneficial in transport planning and policy-making during health crises. Providers of PT services (like airlines or rail operators) can use the information the optimise their services and operations.

This work contributes to the scientific literature because it is one of the first to examine the effects of COVID-19 on modal-split for long-distance travel in Europe. This is studied using a mode choice experiment with Stated-Preference data. In addition, a perceived risk rating experiment is done to determine the factors that influence the perception of risk posed by a COVID-19 infection. Using a modified variant of the HII methodology, combining both models gives additional insights, such as the Value of Risk (Willingness to Pay for reduction in risk). On data acquired from a sample of 1147 (predominantly regular) train passengers recruited in the Netherlands, this method is used and model results are provided.

2 Methodology

Perceived risk due to COVID-19 is complicated to measure, with several elements/attributes that could contribute to this perceived risk. These attributes possibly weigh different for each individual. Therefore, perceived risk is a complex variable, its score is determined by other variables (Molin, 2020).

2.1 Hierarchical Information Integration (HII) theory

This research uses the Hierarchical Information Integration (HII) theory which was introduced by Louviere (1984). This theory is used when decision-makers are confronted with many attributes. Decision-makers categorise these attributes into 'decision constructs'. Decision-makers (respondents) trade-off attributes that belong to such a 'decision-construct' in the first (sub) experiment, the 'rating' experiment. Then in the 'bridging' experiment, decision-makers make a trade-off between the construct evaluations that are done in the 'rating' experiment (Molin, 2020). For this study, an adapted version of the classical HII experiment is used, with only one 'decision construct', to determine perceived risk. Estimating the 'rating' model allows for predicting the perceived risk. Then in the main choice experiment, the perceived risk attribute is shown among the other attributes that are defined. This is not a true 'bridging' experiment as in a conventional 'bridging' experiment more sub experiments for the decision constructs are used; in this research; however, only one. Molin and van Gelder, (2008) showed that this approach is successful.

2.2 Perceived risk rating experiment

The objective of the perceived risk rating experiment is to determine to what degree COVID-19 infection risk variables impact the perceived risk rating of a train or plane trip. To do this, respondents will be asked to score their perceived risk for various trip combinations (by either train or plane). This risk rating is affected by several variables. The determinants/attributes are based mainly on the papers of Dryhurst et al. (2020), Mertens, Gerritsen, Duijndam, Salemink, and Engelhard (2020), Tirachini and Cats, (2020) and Leppin and Aro (2009). All attributes do have four levels. The first four attributes are mode-related attributes. The last four attributes are destination-related attributes.

1. **On-board crowding**: This is specified as the percentage of seats occupied on-board of the vehicle/plane. The levels are: 25% of the seats occupied, 50% of the seats occupied, 75% of the seats occupied and 100% of the seats occupied.
2. **Face mask policy**: This is the policy on-board the vehicle whether or not or which mask is mandatory. The attribute levels are a representation of the masks available. The four levels are: no mask mandatory, any face mask mandatory, at least a surgical (type II) mask mandatory or at least an FFP2 mask mandatory. Every increase in the level of the type of mask gives better protection.
3. **Cleaning policy**: This is policy the train company or airlines has regarding cleaning. The four levels are: same cleaning policy as before COVID-19, increased cleaning policy (focus on touching points), weekly disinfection of the whole vehicle and daily disinfection of the whole vehicle. The levels are based on several policies that airlines and rail companies implemented during the pandemic.

4. **Air conditioning/ventilation**: This attributes is about the type of ventilation or air conditioning is on-board the vehicle or plane. The levels are no ventilation and air conditioning, only ventilation, air conditioning without HEPA filters and air conditioning with HEPA filters.
5. **Travel requirements**: Several policies are implemented within Europe, like 3G or 2G. The following levels were chosen, they reflect different policies within Europe: no mandatory requirements, either testing, vaccination or recovery proof required (3G-rule), only vaccination or recovery proof required (2G-rule), and vaccination or recovery + testing required or booster required (2G+-rule).
6. **Infection rate**: The infection rate levels do reflect different time moments during the pandemic. The levels are: 100 positive tests per day (reflects the situation of the summer of 2020 or in June 2021), 10.000 positive tests per day (reflects the situation of November 2020 and July 2021), 25.000 tests per day (reflects the situation of November 2021) and 100.000 positive tests per day (fictitious situation).
7. **Vaccination rate**: The levels reflect the vaccination rates in different European countries. The levels are from December 2021: 15% vaccination rate (level as in Bulgaria), 30% vaccination rate (level as in Romania), 70% vaccination rate (level as in the Netherlands and EU average), 90% vaccination rate (level as in Portugal).
8. **Travel advice**: These are the travel advice from the Dutch government (Rijksoverheid, 2021). The levels are: green, yellow, orange and red travel advice.

Ngene is used as the software to generate choice sets (the experimental design). There is no prior information available for the rating experiment. This is required in order to create efficient designs. An orthogonal design is chosen because it seeks an attribute level balanced design, this means that the attribute levels occur the same number of times in the option sets. It is decided to do an unlabeled experiment. The rating experiment will only be conducted for train and plane modes. To simplify, there is no differentiation inside the rating experiment. This suggests that people are either taking the train or flying. The rating experiment makes no difference between these modes. As a consequence, the choice sets may be constructed in a sequential order. Ngene generates an orthogonal design with 20 rows. Four blocks are used within the design and each respondent is given five questions to assess their level of perceived COVID-19 risk. An example of the presented rating experiment can be found in figure 1.

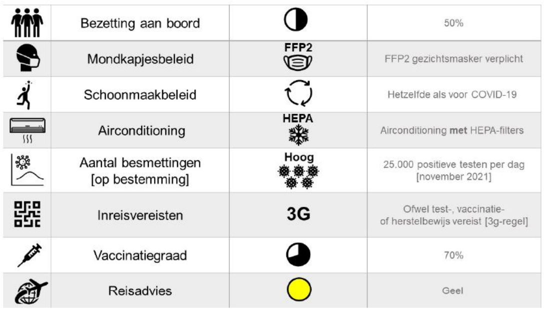


Fig. 1 Example of rating experiment

2.3 Main choice experiment

The main goal of the main (mode) choice experiment is to study how perceived risk is weighted against other factors, such as for example travel cost, while making mode choice decisions. In this case, perceived risk is an independent variable amongst the other main variables. The determinants of the main choice experiment are travel time, travel cost, travel comfort and perceived risk. All attributes do have three levels. The attributes of travel time and travel cost are varied for all modes. As there are two distance classes, in total there are six attribute levels per mode. Travel comfort and perceived risk are only varied for plane and train. Perceived risk and travel comfort are not varied for car. This is because it is assumed that respondents are not sharing their car with strangers. As a consequence, perceived risk in the car is always very low. For travel comfort, it is assumed that people 'own' the same car within the experiment. Therefore, the comfort of the car does not change; thus, the levels of comfort are not varied. The attribute levels can be found in table 1 & 2. For every attribute it is shortly explained why this is included in the main choice experiment.

Travel cost: One of the most important variables in travel behaviour research, and (almost) always included within stated (mode) choice experiments. Trip cost does refer to the cost of making a trip. This can be either the ticket price or the total price for driving the car (fuel + any additional cost). Travel cost is often used in choice experiments to retrieve the willingness to pay for improvements in one of the other attributes of interest. In the case of this research, trade-offs regarding COVID-19 risk are of interest regarding mode-choice for long-distance travel within Europe. From literature, it can be concluded that travel cost is one of the most important variables for mode-choice on long-distance travel (van Goeverden, 2009; Román et al., 2010; Dobruszkes, Dehon, and Givoni, 2014).

Travel time: This variable/attribute is also (almost) always taken into account in studies regarding mode-choice (Morikawa, Ben-Akiva, and McFadden, 2002; van Goeverden, 2009; Román et al., 2010). Some studies refer to the total travel time (thus including access, egress, transfer and waiting time). Other studies only refer to the in-vehicle time. In this study, the total travel time will be used. This includes in-vehicle time and transfer time (if this applies). **Access and egress** are incorporated in the total travel time in this study. This is done in order to keep alternatives simple and understandable for the respondents.

Travel comfort: Travel comfort is also an important factor regarding mode choice for long-distance travel. Román et al. (2010) included comfort as an attribute in their research. The willingness to pay increased if the level of comfort was lower. Furthermore, they found that increased levels of comfort in the aircraft did decrease the perception of time. Train companies and airlines do offer different levels of comfort by distinguishing travel classes. For plane, often, there is a choice between economy class and business class. For the train, this is mostly 2^{nd} class and 1^{st} class.

Perceived risk: Perceived risk is the last attribute that will be included in the main choice experiment. This attribute is directly connected with the rating experiment. In the rating experiment, respondents did rate the risk of getting infected with COVID-19 on their train or plane journey based on eight factors. Respondents rated their journey on a Likert scale with 1-very low, 2-low, 3-medium, 4-high and 5-very high. In the main choice experiment, perceived risk is an attribute amongst the other main attributes. As all levels in the main choice experiment do have three levels, perceived risk does have three levels as well. Therefore, the levels are 1-low, 3-medium and 5-high.

Table 1 Travel time & travel cost values

Travel time						
400-600 km				800-1200km		
# levels	Mode	Value attribute levels		# levels	Mode	Value attribute levels	
3	Train	180 min	3h	3	Train	360 min	6h
		270 min	4.5h			540 min	9h
		360 min	6h			720 min	12h
	Airplane	180 min	3h		Airplane	240 min	4h
		240 min	4h			300 min	5h
		300 min	5h			360 min	6h
	Car	270 min	4.5h		Car	600 min	10h
		390 min	6.5h			780 min	13h
		510 min	8.5h			960 min	16h

Travel cost					
400-600 km			800-1200km		
# levels	Mode	Value attribute levels	# levels	Mode	Value attribute levels
3	Train	30 euro	3	Train	50 euro
		165 euro			200 euro
		300 euro			350 euro
	Airplane	50 euro		Airplane	50 euro
		175 euro			225 euro
		300 euro			400 euro
	Car	80 euro		Car	100 euro
		115 euro			150 euro
		150 euro			200 euro

Table 2 Travel comfort & perceived risk levels

Travel comfort		
# levels	Mode	Value attribute levels
2	Train	2nd class
		1st class
2	Airplane	Economy
		Business
1	Car	Same level

Perceived risk		
# levels	Mode	Value attribute levels
3	Train	1-Very low
		2-Medium
		3-Very High
3	Airplane	1-Very low
		2-Medium
		3-Very High
1	Car	1-Very low

In this part, the included socio-demographic attributes and the travel behaviour questions are discussed.

Age: Is one of the most common used socio-demographic variables. Several papers and research did look include this into their research: Buehler and Nobis (2010), Hensher and Rose (2007), Paulssen, Temme, Vij, and Walker (2014), Román et al. (2010). Often different age groups do have different preferences for certain modes.

Gender: This socio-demographic variable is also a very common variable to include in stated-choice experiments. Almost all studies include this socio-demographic variable (Buehler and Nobis 2010; Hensher and Rose 2007; Paulssen et al. 2014; Román et al. 2010; Johansson, Heldt, and Johansson 2006). With this variable, it can be analysed if women and men do have do have different preferences regarding the variables in the main choice experiment, this could be for example in the preference for a certain mode.

Income: Also an important socio-demographic variable to take into account. It is not always clear how this income is asked in the survey. Some papers ask about dispensable income, while at the same time other papers ask for gross income (Buehler and Nobis 2010; Hensher and Rose 2007; Paulssen et al. 2014; Román et al. 2010; Johansson et al. 2006). For this study, the gross-income is used. As it is expected that higher income influences mode-choice, this variable is included in the model. It is expected that higher income will increase the willingness to pay for the attributes, like time and comfort for example.

Work status: Hensher and Rose (2007) included this variable in their research in order to check if the sample was representative, but they did not include this in the model specification. However, for this study it is included in the model specification as well.

Education level: Socio-demographic variable that is often included in models as well. Johansson et al. (2006) stated that they previously did not find any literature on including education level for long-distance travel; however, in his research, it turned out to be significant. Education level is also expected to have an influence on mode-choice and, therefore, is included in the model.

Travel frequency: This variable is sometimes included in research. Several research included this such as Román et al. (2010), Van Loon and Rouwendal (2013) & Nieto-García, Muñoz-Gallego, and Gonzalez-Benito (2020). For this research, this attribute will be included to test whether travel frequency of respondents influences perceived risk.

Preferred travel mode: Hensher and Rose (2007) included the preferred travel mode in their study. For this mode-choice research, it is interesting to see if the preferred travel mode influences the mode choice. It is expected that people will stick to their main preferred mode when making a choice. In this research, it will be questioned for both the 400-600km and 800-1200km distance classes.

Payment: This attribute is included in order test whether the value of time values changes when payment is done by the respondents or by someone else, or education/work. Kouwenhoven et al. (2014) found significant different in the VoT values for different purpose of work.

Trip purpose: Both Buehler and Nobis (2010) and Román et al. (2010) included trip purpose in their research, and for both studies this turned out to be an important factor. Often willingness to pay for business trips is higher (as the respondent does not pay by him or herself) than for leisure trips. Therefore this will also be taken into account in this research. However, this is not done in the same manner as in these studies. In this research, people are asked if they pay for themselves or if someone else is paying for their trip. In this case, it can still be analysed if the willingness to pay is higher if respondents do not have to pay for themselves.

Travel company: A study done by Mertens et al. (2020) showed that risk for family and loved ones, was one of the most important predictors for COVID-19 fear. Therefore it is included also in this research, to test whether travel company contributes to perceived risk.

For the main choice experiment, an efficient design is chosen. An efficient design results in fewer choice sets for the survey than an orthogonal design. In this case, there is prior information available; however, not for all four variables. Therefore a Bayesian D-efficient design is chosen, so the prior values can differ around the mean. Ngene finds an efficient design with only ten rows. It is chosen to go for a design with 12 rows as this number can be divided by three. The reason for this is the fact that 12 choice sets give more information. As the main choice is divided into two distance classes, this gives in total eight main choice questions for the respondents. An example of the main choice experiment is shown in figure 2.

Fig. 2 Example of main choice experiment question (in Dutch)

3 Characteristics of the sample

The intended target group is a representative sample of the Dutch population. Data was collected in several ways: by sharing the link of the survey on social media platforms, by sharing the link to colleagues and by distributing the survey to the customer panel of NS, the Dutch Railways. The survey was open for a response from the 8^{th} of February to the 8^{th} of March. Cooperation with NS was done. The survey was also distributed to their panel. In total, the survey was sent to 5000 people. In total, 938 respondents took part in the survey and fully completed the survey. This is a response rate of 18.7%. Moreover, the link to the survey was distributed among friends, family, and colleagues and through social media platforms like LinkedIn, Instagram, and the Royal HaskoningDHV C-Infra department). This resulted in a total of 209 completed responses. All in all, the total number of completed responses of this survey was 1147, which is way above the minimum needed respondents. The sample consists mainly of highly educated people and people with high income. For age and gender, the sample approximates the Dutch population. The consequence of the higher incomes and higher education is that presumably the Value of Time and Willingness to Pay values are overestimated. The frequencies, percentages and percentages from CBS can be found in table 3.

Table 3 Characteristics of the sample

Socio-demographic variable	Category	Frequency	Percentage sample	Percentage CBS
Age	0-19	27	2.4%	21
	20-40	261	23.1%	25% / 34%[1]
	40-65	418	36.9%	34% / 43%[1]
	65 to 80	371	32.8%	15% / 19%[1]
	80+	55	4.9%	5% / 6%[1]
	Total	1132		
Gender	Female	547	49.0%	50.3%
	Male	569	51.0%	49.7%
	Total	1116		
Income	€10.000	77	8.5%	13.6%
	€10.000-€20.000	85	9.4%	23.3%
	€20.000-€30.000	117	12.9%	18%
	€30.000-€40.000	156	17.2%	14.7%
	€40.000-€50.000	141	15.5%	10.9%
	€50.000-€100.000	247	27.2%	16.5%
	€100.000-€200.000	76	8.4%	2.6%
	€200.000 or more	10	1.1%	0.4%
	Total	909		
Education	Basisonderwijs	7	0.8%	8.3%
	Vmbo-b/k, mbo1	22	2.4%	10.7%
	Vmbo-g/t, vwo-onderbouw	65	7.0%	8.4%
	Mbo2, mbo3 en mbo4	119	12.9%	26.6%
	Havo, vwo	93	10.0%	9.5%
	Hbo-,wo-bachelor	258	27.9%	21.9%
	Hbo-,wo-master, doctor	323	34.9%	13%
	Do not know	39	4.2%	1.7%
	Total	926		

4 Model estimation

For the rating experiment, respondents rated their perceived risk of COVID-19 based on a Likert scale. A Likert scale is a method for interrogating data that is difficult to quantify and providing it with an ordinal level of measurement (Joshi, Kale, Chandel, and Pal, 2015). Therefore it is widely used in questionnaires and surveys. It was chosen to go for a five-point scale, as this is easy for respondents to rate. The rating is analysed using a linear regression model. The data is analysed with IBM SPSS statistics 26.0. The following regression formula is for the perceived risk rating.

$$\begin{gathered}
PR_{COVID-19} = C + \beta_{RA} * RA + \beta_{VR} * VR + \beta_{LF} * LF \\
+\beta_{INFECT1} * INFECT1 + \beta_{INFECT2} * INFECT2 \\
+\beta_{INFECT3} * INFECT3 + \beta_{VL} * VL + \beta_{NH} * NH \\
+\beta_{HP} * HP + \beta_{3G} * 3G + \beta_{2G+} * 2G + +\beta_{AM} * AM \\
+\beta_{FFP2} * FFP2 + \beta_{GEN} * GEN + \beta_{HBO} * HBO \\
+\beta_{WO} * WO + \beta_{INC_{20-40}} * income_{20-40}
\end{gathered}$$

C = constant, RA = Red travel advice, VR = vaccination rate, LF = load factor, INFECT1= 10.000 infections. INFECT2 = 20.000 infections, INFECT3

= 100.00 infections, VL = ventilation only, NH = air conditioning no HEPA filter, HP = air conditioning with HEPA filter, 3G = 3G policy, 2G+ = 2G+ policy, AM = any mask, FFP2 = FFP2 mask mandatory, GEN = gender, HBO = HBO education level, WO = university education level, INC_{20-40} = income between €20.000 and €40.000.

4.1 Results regression

The results of linear regression give several implications. In total, eight main attributes are included in the regression model, two of them are ratio scales, and six of them are ordinal scales. All the ordinal scale variables are dummy coded. In total, there are 20 main parameters estimated. Also, five socio-demographic attributes are included that are also dummy coded, ensuring a total of 14 parameters. For the main attributes, all parameters do have the expected sign, except 3G policy (i.e., either testing, recovery or vaccination proof needed to travel). In this case, a negative sign would be expected as this policy will decrease the probability of someone infected when travelling (in comparison to the base level, with no travel requirements).

All other main parameters do have the expected sign, and there are also insignificant main parameters. The constant is 2.8; this is the value if all parameters are set to be zero. In this case all levels are set to the base; respondents rate perceived risk at 2.8 (so that is around the mean value of 3). This means that respondents, on average, rated their perceived risk with all levels set to the base, as a little under 3-medium perceived risk. All significant parameters are highly significant, except for the level '25.000 infections per day'.

Main parameters

- **Travel advice**: Both the yellow and orange advice parameters are not significant, so these levels do turn not out to be different from the base level green advice. However, red travel advice has the largest positive effect on perceived risk with a value of 0.698.
- **Vaccination rate**: This is a ratio variable with a contribution of -0.007 for every percentage point increase in vaccination rate in the country of destination. For example a vaccination rate of 50% gives the following parameter: $50 \times -0.007 = -0.35$. When travelling to a country with a vaccination rate of 90% (Portugal), perceived risk is decreased with -0.63 rating points.
- **Load factor**: This is also a ratio variable, with (almost) the same but opposite contribution of vaccination rate. The value of this parameter is 0.006. A load factor of 75% would lead to $70 \times 0.006 = 0.42$ increase in rating points. A load factor of 100% would lead to an increase of 0.6 points on perceived risk.
- **Infections**: There is some counter-intuitive outcome, as 10.000 tests per day contribute more (with a value of 0.252) to perceived risk than 25.000 tests per day (with a value of 0.123). A reason could be that respondents

find it hard to imagine what the difference in levels means. INFECT2 is also just significant (or just insignificant) on the 5% level (p-value of 0.051). The highest level, i.e., 100.000 tests per day, has the highest contribution of the dummy variables. This is in line with expectations. It also has the second-highest contribution of the attributes, with a value of 0.495.

- **Ventilation/air conditioning**: All dummy variables turn out to be significant. Only ventilation has the highest contribution to the decrease in perceived risk with a value of -0.405. This is in line with expectations as there was a focus from society on ventilation. Therefore it could be expected that people do think this is important. Air conditioning without HEPA filter has the lowest contribution of the dummies (with a value of -0.198); again, this could be expected, as air conditioning without HEPA filters has a lower level of protection against viruses than air conditioning with HEPA filters. Air conditioning with HEPA filter has a higher contribution than the previous level. The value of air conditioning with HEPA filters is -0.289.
- **Travel requirements**: As explained earlier, the first dummy variable 3G policy, has a positive sign with the value of 0.144, which is not in line with expectations. A possible explanation for this could be the fact that when first introducing the 3G policy last summer 2021, there was an exponential increase in infections. 2G policy is not significant. This can be explained as the 2G policy was never introduced, and there was a lot of resistance. Also, there effectiveness of both the 3G and 2G has been questioned and is reduced (Mouter, Hernandez, and Itten, 2021). 2G+ is an extra level of security in comparison to 2G, with people also needing to test even with a vaccination or recovery proof, it turns out to be significant. This level shows a reduction of perceived risk, with a value of -0.272.
- **Face mask policy**: Any face mask required and at least FFP2 face mask required are significant. The level 'any face mask required' has a higher contribution (value of -0.268) than 'at least FFP2' (value of -0.140). So the need to put on any face mask is more important to reduce perceived risk than having at least an FFP2 mask, according to the respondents.
- **Cleaning policy**: None of the dummy variables turned out to be significant. This shows that there is no difference from the base level 'same cleaning policy as before COVID-19' and therefore does not reduce perceived risk.

Socio-demographic attributes

- **Gender**: As in line with the expectations, gender turns out to be significant with a value of 0.088. This means that being women & 'other' increases perceived risk with 0.088 rating points. This is not a very high value in comparison to other attributes, but the value is significant, so there is a difference between men and women.
- **Education**: The level HBO and WO are both significant and positive. This means in comparison to the base (MBO or lower), people with education HBO and WO perceive the risk of COVID-19 as higher than people with

MBO or lower education level. The value for HBO is 0.137 and for WO 0.127, so people with HBO perceive risk as a bit higher than WO.
- **Income**: For this attribute, there are no expectations. The levels 'income between €20.000 and €40.000' and 'income between €40.000 and €100.000' and 'I do not want to say are significant. The first level has a value of 0.112, the second level 0.168 and the last 0.259. Having an income between €20.000 and €100.000 contributes to a higher perceived risk.
- **Age**: The age group '20-40' years contributes to a lower perceived risk (in comparison to the base 'younger than 20 years). The other two dummy variables did not turn out to be significant. So the age groups '40-65 years' and '65 years and older' do not contribute to an increase or decrease in perceived risk.
- **Work status**: All of the parameters are highly insignificant. So work status does not influence perceived risk.

Table 4 Left: Significant attributes linear regression, right: insignificant attributes

Model	Main effects			Main & Socio-demographics		
Main attributes	Value	t	p-value	Value	t	p-value
(Constant)	3.039	52.919	0.000	2.815	40.913	0.000
ADVICE3	0.698	15.412	0.000	0.702	15.620	0.000
VACC	-0.006	-12.866	0.000	-0.007	-13.203	0.000
CROW	0.006	11.306	0.000	0.006	11.604	0.000
INFECT1	0.253	5.917	0.000	0.252	5.953	0.000
INFECT2	0.136	2.134	0.033	0.123	1.954	0.051
INFECT3	0.494	12.575	0.000	0.495	12.692	0.000
AIRCO1	-0.388	-5.411	0.000	-0.405	-5.684	0.000
AIRCO2	-0.193	-4.467	0.000	-0.198	-4.617	0.000
AIRCO3	-0.286	-7.419	0.000	-0.289	-7.560	0.000
REQUIRE1	0.138	2.627	0.009	0.144	2.757	0.006
REQUIRE3	-0.267	-7.424	0.000	-0.272	-7.608	0.000
MASK1	-0.261	-5.397	0.000	-0.268	-5.576	0.000
MASK3	-0.137	-3.889	0.000	-0.140	-4.005	0.000
Socio-demographic attributes				Value	t	p-value
GENDER				0.088	3.283	0.001
HBO				0.137	4.191	0.000
WO				0.127	3.783	0.000
INCOME_20_40				0.112	2.634	0.008
INCOME_40_100				0.168	4.005	0.000
AGE_20_40				-0.095	-2.915	0.004
DO_NOT_SAY_INCOME				0.259	5.473	0.000
$R^2 = 0.129$				$R^2 = 0.143$		
Adjusted $R^2 = 0.127$				Adjusted $R^2 = 0.140$		

Model	Main effects			Main & Socio-demographics		
Main attributes	Value	t	p-value	Value	t	p-value
MASK2	-0.036	-1.536	0.125	-0.029	-1.240	0.215
CLEAN1	-0.004	-0.304	0.761	-0.002	-0.132	0.895
CLEAN2	-0.023	-1.074	0.283	-0.020	-0.920	0.358
CLEAN3	0.003	0.216	0.829	0.002	0.139	0.890
REQUIRE2	0.022	0.791	0.429	0.019	0.674	0.500
ADVICE1	-0.003	-0.126	0.900	-0.001	-0.056	0.955
ADVICE2	-0.021	-0.941	0.347	-0.020	-0.886	0.376
Socio-demographic attributes				Value	t	p-value
AGE_40_65				-0.022	-1.581	0.114
AGE_65_AND_OLDER				0.022	1.530	0.126
INCOME_MORE_100				0.025	1.374	0.170
NOT_WORKING				-0.006	-0.382	0.702
STUDENT_SCHOOL				0.008	0.584	0.559
RETIRED				-0.009	-0.709	0.479
OTHER_EDU				0.012	0.801	0.423

4.2 Results MNL model

The results of the different MNL models can be found in table 5. Looking at the results, it can be concluded that all parameters do have the expected sign. As all insignificant parameters have been removed, all remaining parameters are significant at the 5% level. As a consequence of adding the interaction between risk and time, there is no separate parameter for risk. The parameter for time is mode dependent. However, the parameter of time for plane is insignificant, so it is not included in the model.

Table 5 Base model, base model with main attribute interactions & final model

Model	Base model			Main interaction			Final model		
Parameter	Value	t	p-value	Value	t	p-value	Value	t	p-value
ASC_PLANE	-0.359	-5.3	1.16e-07	-1.51	-11.3	0	-2.23	-14.1	0
ASC_TRAIN	0.35	6.4	1.53e-10	-0.197	-2.52	0.0118	-0.414	-4.16	3.15e-05
B_COMFORT	0.281	8.7	0	0.281	8.3	0	0.346	10.4	0
B_COST	-0.00306	-21.2	0	-0.00326	-23	0	-0.00973	-8.02	1.11e-15
B_TIME	-0.167	-26.5	0						
B_RISK	-0.22	-22	0						
B_TIME_C				-0.168	-23.1	0	-0.194	-24.8	0
B_TIME_T				-0.0991	-9.42	0	-0.114	-10.1	0
B_TIME_RISK_P				-0.0437	-9.9	0	-0.011	-3.55	0.000392
B_TIME_RISK_T				-0.039	-17.1	0	-0.0107	-4.32	1.57e-05
B_GENDER_PR							-0.0217	-10.5	0
B_EDU_HBO_PR							-0.0142	-4.26	2.05e-05
B_EDU_WO_PR							-0.0237	-6.63	3.27e-11
B_EDU_OTHER_PR							-0.018	-4.99	6.05e-07
B_AGE_COST_20_40							0.00325	2.6	0.00932
B_AGE_COST_40_65							0.0064	5.18	2.23e-07
B_AGE_COST_65_AND_OLDER							0.00723	5.88	4.17e-09
B_PAYMENT_WORKEDU_COST							0.00223	3.18	0.00146
B_COMPANY_PR_FRIENDS							-0.00746	-2.62	0.00878
B_COMPANY_PR_OTHER							0.0197	4.06	4.86e-05
B_PURPOSE_WORK_TIME							-0.0596	-3.98	6.82e-05
B_PREF_CAR_C							0.496	6.3	2.97e-10
B_PREF_CAR_T							-0.268	-3.33	0.000881
B_PREF_PLANE_P							1.04	14.5	0
B_PREF_TRAIN_C							-0.184	-2.43	0.0152
B_PREF_TRAIN_T							0.404	5.49	3.99e-08
ρ^2	0.126			0.133			0.24		

- ASC plane and ASC train are both negative, so car is more preferred.
- Beta comfort is positive, beta time and beta cost and beta time are both negative. Travel time is more negative for car than for train. For plane, time is insignificant.
- Perceived risk is significant as an interaction with time. This means that for longer travel times, COVID-19 risk is perceived higher. Perceived risk is also negative and different for train and plan; however, there is a very small difference, see table 5.
- Several interactions are significant with perceived, also on the ASCs and travel cost.
- A preference for car has a positive contribution to the ASC of car and a small negative contribution to train. A preference for train has a small negative contribution on the ASC of car and a positive contribution to the ASC of train. A preference for plane has a positive contribution to the ASC of plane.
- For travel time, only purpose work is significant. Respondents that travel for work are more sensitive to travel time.
- For travel cost, both age and payment by work or educational institution are significant. Higher age results in less weight to travel cost, the same counts if the trip is payed by work the educational institution.
- For perceived risk, educational level and gender are significant. Having at least HBO or WO education (in comparison to MBO) results in more weight to perceived risk. Women have a higher weight to perceived risk as well.
- Increasing age results in a lower weight to travel cost.

4.3 Interpretation of parameters

Several interactions contribute to the different main parameters. It is highlighted that the WtP for perceived risk, particularly in terms of travel costs, can vary significantly across people of varied ages, payment, education level and travel companies, and possibly other unobserved background variables not represented in this study. Therefore, the average respondent is used, based on the average (and when not possible) on the most common value, the values are shown in table 6.

Table 6 The average respondent

Age	38
Education	HBO
Gender	Men & women
Trip purpose	Leisure
Travel company	Alone

The 'value of risk' (VoR) is the trade-off between perceived risk and travel cost.Using the average respondent values, the VoR values is calculated in equation 2 & 3.

$$VoR_{in\ travel\ cost} = \frac{\frac{\delta U}{\delta PR}}{\frac{\delta U}{\delta TC}} \tag{1}$$

$$VoR_{men} = \frac{-0.011 - 0.142}{-0.00973 + 0.00325} = 3.89\ euro \tag{2}$$

$$VoR_{women} = \frac{-0.011 - 0.142 - 0.0217}{-0.00973 + 0.00325} = 7.24\ euro \tag{3}$$

The VoR for men is €3.89 per level of perceived risk per hour; for women this is €7.24. For women this is higher due to the interaction between gender and perceived risk. For higher travel times and higher perceived risk levels, this value becomes equivalent 3.89/7.24*TT*PR. VoR is thus linear, there are no quadratic components.

The trade-off between travel comfort and perceived risk does have to following results. The VoR for men in comfort 'points' is 0.072 (per hour); for women, this is 0.134.

$$VoR_{in\ travel\ comfort} = \frac{\frac{\delta U}{\delta PR}}{\frac{\delta U}{\delta CF}} \tag{4}$$

$$VoR_{in\ comfort,men} = \frac{-0.011 - 0.142}{0.346} = 0.072 \tag{5}$$

$$VoR_{in\ comfort,women} = \frac{-0.011 - 0.142 - 0.0217}{0.346} = 0.134 \tag{6}$$

As perceived risk is dependent of time, it is also possible to express what comfort is 'worth' in terms of travel time for different perceived risk levels. The trade-off (TO) for a full reduction/increase in class, this means one full comfort 'point' (so 1^{st} class $\leftrightarrow$ 2^{nd} class or business $\leftrightarrow$ economy) is shown in table 7.

Table 7 Trade off comfort travel time

Men			Women		
PR level	Comfort [points]	TT [hours]	PR level	Comfort [points]	TT [hours]
1	0.072	13.889	1	0.134	7.463
2	0.144	6.944	2	0.268	3.731
3	0.216	4.630	3	0.402	2.488
4	0.288	3.472	4	0.536	1.866
5	0.36	2.778	5	0.670	1.493

The value of time in transportation (economics) is the potential cost of the time that a passenger spends on their route. In essence, this is the amount a passenger is ready to pay to save time, or the amount they would take as compensation for lost time. The amount of time that passengers will save is one of the key justifications for transportation upgrades (Kouwenhoven et al., 2014). The VoT for train has a value of €17.59 for train and €29.94 for car.

$$VoT = \frac{\frac{\delta U}{\delta TT}}{\frac{\delta U}{\delta TC}} \tag{7}$$

$$VoT_{train} = \frac{-0.1140}{-0.0097 + 0.0033} = 17.59 \; euro/hour \tag{8}$$

$$VoT_{car} = \frac{-0.194}{-0.0097 + 0.0033} = 29.94 \; euro/hour \tag{9}$$

The different VoT values found for this thesis in comparison to other studies is shown in table 8. The fact that car has a higher VoT is in line with KiM *Netherlands Institute for Transport Policy Analysis* (KiM, 2020). The parameter of time for plane is not significant; hence, no VoT can be specified for plane. In this study, the VoT for car is just over 2 euros more. For Kouwenhoven et al. (2014) there is a very small difference found between the modes. Shires and De Jong (2009) did not find differences in VoT for train and car. It must be noted that both KiM (2020) and Kouwenhoven et al. (2014) focused on short distance travel, while Shires and De Jong (2009) focused on long-distance. The average VoT found in this paper is almost the same as found in Shires and De Jong (2009).

Table 8 VoT results of different studies

	This paper	KiM (2020)	Van Kouwenhoven (2014)	Shires and De Jong, (2009)
VoT Train	17.59	13.22	9.25	24.00
VoT Car	29.94	15.58	9.00	
Average	23.77	14.40	9.13	24.00

Willingness to pay for comfort can also be interpreted from the result. The equation for WtP for comfort for both train and plane can be found in equation 11.

$$WtP_{comfort} = \frac{\frac{\delta U}{\delta CF}}{\frac{\delta U}{\delta TC}} \tag{10}$$

$$WtP_{comfort} = \frac{0.346}{-0.0097 + 0.0033} = 53.40\ euro/class \tag{11}$$

The result show the willingness to pay for an upgrade in class (so 2^{nd} class $\rightarrow$ 1^{st} class or economy $\rightarrow$ business) is €53.40. No mode-specific beta turned out to be significant, so there is no distinction between classes in train and plane. Balcombe, Fraser, and Harris (2009) found a value of about €120, but they agreed that this value is on the high side. A big amount of the passengers travelling in business or 1^{st} class is flying for business purposes (BusinessAM, 2020). It is therefore of interest to show the willingness to pay for business/educational travellers (respondents who's trip is payed by company of educational institution), the value is shown in equation 12. This shows that the WtP for an upgrade is substantially higher.

$$WtP_{comfort} = \frac{0.346}{-0.0097 + 0.0033 + 0.00222} = 81.22\ euro/class \tag{12}$$

4.4 Combination results both experiment

Both the perceived risk rating experiment and the main (mode) choice experiment are estimated, so now the results of both experiments can be estimated. Within the perceived risk rating experiment, perceived risk was the dependent variable; at the same time, it was an independent variable in the main (mode) choice experiment. For the rating experiment, gender and education level have a positive contribution to perceived risk. Both experiments can be combined by using the (absolute) linear regression coefficients of the perceived risk rating experiment. Then these values are combined with the WtP values that are just calculated. Because of the dummy coding, all different dummy variables have an independent contribution to the value of perceived risk. The results are shown in table 9. Respondents are willing to most for a decrease in level of perceived when there is a risk red travel advice. In comparison to the base (no travel advice), men are willing to pay €32.74 for a decrease in level of risk when travelling for 12 hours; women €60.93. If men are travelling for only 6 hours, this value is half of the 12-hour value, so €16.37; for women, €30.47. Moreover, there is a high WtP when there is a 100% load factor (men €30.17 and women €56.15). At last, 100.000 infections result in a WtP for decrease of one level of perceived risk of €23.09 for men and €42.96 for men. The big difference between men and women is due to the significant interaction between gender and perceived risk.

Table 9 WtP values for the different risk factors

			Men				Women			
		PR level difference	$\Delta_{level} = -1$				$\Delta_{level} = -1$			
		Travel time	3	6	9	12	3	6	9	12
		Value	€11.67	€23.33	€35.00	€46.67	€21.71	€43.43	€65.14	€86.85
Travel advice	Red travel advice	0.702	€8.18	€16.37	€24.55	€32.74	€15.23	€30.47	€45.70	€60.93
Vaccination rate	Parameter	-0.007	€-0.08	€-0.15	€-0.23	€-0.30	€-0.14	€-0.28	€-0.42	€-0.57
	15	-0.098	€-1.14	€-2.28	€-3.42	€-4.55	€-2.12	€-4.24	€-6.36	€-8.48
	30	-0.195	€-2.28	€-4.55	€-6.83	€-9.11	€-4.24	€-8.48	€-12.72	€-16.95
	70	-0.455	€-5.31	€-10.63	€-15.94	€-21.26	€-9.89	€-19.78	€-29.67	€-39.56
	90	-0.586	€-6.83	€-13.66	€-20.50	€-27.33	€-12.72	€-25.43	€-38.15	€-50.86
Load factor	Parameter	0.006	€0.08	€0.15	€0.23	€0.30	€0.14	€0.28	€0.42	€0.56
	25	0.162	€1.89	€3.77	€5.66	€7.54	€3.51	€7.02	€10.53	€14.04
	50	0.323	€3.77	€7.54	€11.31	€15.08	€7.02	€14.04	€21.06	€28.07
	75	0.485	€5.66	€11.31	€16.97	€22.63	€10.53	€21.06	€31.58	€42.11
	100	0.646	€7.54	€15.08	€22.63	€30.17	€14.04	€28.07	€42.11	€56.15
Infection rate	10.000 infections per day	0.252	€2.94	€5.88	€8.82	€11.76	€5.47	€10.95	€16.42	€21.90
	25.000 infections per day	0.123	€1.44	€2.88	€4.32	€5.76	€2.68	€5.36	€8.04	€10.72
	100.000 infections per day	0.495	€5.77	€11.54	€17.31	€23.09	€10.74	€21.48	€32.22	€42.96
Air conditioning	Only ventilation	-0.405	€-4.72	€-9.44	€-14.17	€-18.89	€-8.79	€-17.58	€-26.36	€-35.15
	Airco no HEPA	-0.198	€-2.31	€-4.62	€-6.92	€-9.23	€-4.29	€-8.59	€-12.88	€-17.18
	Airco with HEPA	-0.289	€-3.38	€-6.75	€-10.13	€-13.51	€-6.28	€-12.57	€-18.85	-25.14
Travel requirements	3G-policy	0.144	€1.68	€3.36	€5.03	€6.71	€3.12	€6.25	€9.37	€12.49
	2G+-policy	-0.272	€-3.17	€-6.35	€-9.52	€-12.70	€-5.91	€-11.81	€-17.72	€-23.63
Face mask policy	Any face mask	-0.268	€-3.12	€-6.24	€-9.37	€-12.49	€-5.81	€-11.62	€-17.43	-23.25
	At least FFP2	-0.140	€-1.64	€-3.27	€-4.91	€-6.55	€-3.05	€-6.09	€-9.14	€-12.18
	Socio-demographic									
Gender	Women	0.088	€1.03	€2.06	€3.09	€4.12	€1.92	€3.83	€5.75	€7.66
Education	HBO	0.137	€1.60	€3.21	€4.81	€6.41	€2.98	€5.97	€8.95	€11.94
Income class	Income €20.000 to €40.000	0.112	€1.31	€2.62	€3.92	€5.23	€2.43	€4.87	€7.30	€9.74
	Income €40.000 to €100.000	0.168	€1.96	€3.92	€5.88	€7.85	€3.65	€7.30	€10.95	€14.60
Age	Age 20 to 40 years	-0.095	€-1.11	€-2.22	€-3.32	€-4.43	€-2.06	€-4.12	€-6.18	€-8.25

5 Influence perceived risk on market share

This research takes a broad approach to long-distance travel within Europe. To give a clear example of what the different implications from the results mean for real-life examples, it is chosen to go for three different cases. First, a short route that is popular by train will be used, this is the Amsterdam - London route. At the same time, the plane and car options are viable and popular as well. The second route is Amsterdam - Berlin. This route has all three modes as viable options. This route is popular by train as there is a direct connection. At third, a longer route is chosen to see the difference between a shorter and longer routes. In order to look at the influence of perceived risk on the modal-split, real-world values for travel time, travel cost and comfort are being used for these routes. Then, the perceived risk levels are varied so that the influence can be discussed.

Amsterdam - London: For this route, the maximum market share differences are as follows. When perceived risk in the train increases from 1 to 5 and the plane stays at level 1, train loses 10% point market share for men; for women this is 17%. Plane has 3% point decrease in market share for men; for women this is 5% point. Car increase by 7% for men; for women this is 12% point. This is because train has a longer travel time. Women have a higher weight to perceived risk, therefore market shares differences are higher for women. This is due to the fact that plane has a shorter travel time, for car this is because there is no perceived risk. For smaller differences in perceived risk, these market share changes are even smaller. The perceived risk levels are complex concepts, its score is determined by a lot of factors. Therefore there is no real such thing as risk level 1. So, these extreme changes are not realistic in real life. Moreover, 4 out of the 8 factors are destination specific. In this case, these factors do not change between modes for the same OD pair. To make this clear, the maximum difference in perceived risk points for the same

OD-pair will be (only considering mode-related attributes):

$$PR_{low} = 2.815 - 0.405 - 0.268 = 2.14$$
$$PR_{high} = 2.815 + (0.006 * 100) = 3.46$$
$$\Delta PR = 3.46 - 2.14 = 1.32$$

In this case, when the perceived risk of train is 2 and plane 3 (to approximate risk point 3.46 and 2.14) to train 1 and plane 3, there is only a 1 to 2% point market share difference for men; for women this is 2 to 3% point. Thus, it can be concluded that the influence of perceived risk is very moderate for this route. The change for the different factors that contribute to perceived risk is even smaller.

Amsterdam - Berlin: The maximum market share differences for this route are a bit higher. When perceived risk in train is increasing from 1 to 5 while plane stays at 1, results in a decrease of market share for train of 16% point for men; for women this is 27% point. Plane increases 6% point for men; for women it increase by 10% point. Car increases by 10% point for men; for women this is 17% point. As there is only 1.32 perceived risk level difference possible, the maximum market share difference possible is around 1-4% point for men; for women this is 2-7% point.

Amsterdam - Barcelona: For this route, perceived risk has a bigger influence on the market shares. This could be expected as perceived risk is dependent on time. The maximum market share difference for train is 21% point decrease for men; for women this is 29% point. For men plane increases by 17% point; for women this is 24% point. Car increases by 4% point for men; for women this is 5% point. However, as there is only 1.32 perceived risk level difference, the realistic effect of perceived risk is a maximum of 4-6% point for men; for women 3-16% point.

With the 1.32 perceived risk difference, market share differences are at maximum. For the three routes, the maximum possible market share difference show for London a small effect of perceived risk. For the Amsterdam - Berlin and Amsterdam - Barcelona routes, this effect becomes larger. However, the 1.32 is the maximum possible difference. For real-life trips, not all of the four mode-related attribute will be different. So in this case, the market share difference is smaller.

6 Conclusion

To research the effects of perceived COVID-19 infection risk, this study was conducted. In total, 1147 respondents took part in this study. To study whether perceived COVID-19 risk influences mode choice, the survey consisted of two main parts, a part about socio-demographics and a part about travel behaviour. The first part of the survey was the rating experiment. In this part, respondents had to rate their perceived risk of infection with COVID-19 due to several factors; four of these factors were destination-related, and the other four were mode-related. After this, respondents were faced with the main (mode) choice experiment. In this experiment, respondents had to choose between train, plane and car based on travel time, travel cost, travel comfort (class of travel) and perceived risk (now a given value). To analyse the data, the discrete choice modelling theory was used. At last, an adapted variant of the Hierarchical Information Integration (HII) theory was used to combine both of the results.

In total, eight factors were established: travel advice, vaccination rate, load factor, infections, air conditioning, travel requirements, face mask policy and cleaning policy. Red travel advice is the most important factor in the contribution to perceived risk. Vaccination rate and load factor do have about the same contribution to perceived risk and are the second most important predictors for perceived risk, but the signs are opposite. Vaccination rate has a negative sign, load factor a positive sign. The attribute of 100.000 positive tests has the highest contribution to the number of infections. For the ventilation/air conditioning variables, ventilation has the highest negative contribution. This can be explained as there was a focus on ventilation by the government to reduce the number of infections. Regarding travel requirements, the 3G policy has a positive contribution to perceived risk, which is counter-intuitive. This could be explained, as with the (first) introduction of the 3G-policy, there was a dramatic increase in number of infections. The 2G policy has no effects, but the 2G+ policy does decrease perceived risk. When looking at the face mask policy, having any mask is more important in the decrease of perceived risk than having at least an FFP2 mask. The cleaning policies do not contribute to decrease perceived risk. Regarding socio-demographic values, the following conclusion can be taken. For this question, gender, age, income, education and work status were added to the model. Work status does not have a contribution to perceived risk and thus does not influence perceived risk. Gender has a positive contribution to perceived risk; hence, being woman results in a higher (average) rating for perceived risk. The same counts for education levels' HBO' and 'WO'. Moreover, having an income of €20.000 and €100.000 per year also contributes to perceived risk. When looking at age, most of the age groups did not have a contribution to perceived risk; however, being between 20 and 40 years old leads to a decrease in perceived risk.

Travel cost becomes less important with increasing age; this means that the cost parameter becomes less negative with increasing age. Moreover, if the payment is made by the educational institute or work, people become less sensitive to cost as well. For travel time, only travel purpose has an influence. Travelling for work contributes to a more negative weight on travel time, so travel time becomes more important. For travel comfort, no significant interactions were found. At last, for perceived risk, both gender and travel company turned out to be significant. Being a woman contributes to a higher weight to perceived risk; this was also found in the rating experiment. At last, travelling with friends contributes to perceived risk.

The value of risk for a decrease in the level of risk for a man is €3.89 per decrease in level per hour; for a woman, this is €7.24 per decrease in level per hour. The gender difference is found due to a significant interaction between perceived risk and gender. The levels of risk go from 1-very low to 5-very high. Perceived risk depends on time; therefore, there is no trade-off in travel time. Longer travel time leads to a higher perceived risk. It is noted that this is for the average respondent. Different values are for different ages, travel purposes, who pays for the trips and education levels. In terms of travel comfort, the value of risk for men is 0.072 comfort 'points' for one level decrease of perceived risk per hour; for women, this is 0.134. An upgrade in travel class is worth 13.89 hours for men when perceived risk level is one; for women, this is 7.46 hours. For perceived risk level 5, a full increase in travel class is 'worth' only 2.78 hours of travel for men; for women, 1.49 hours. The value of time found is 17.59 euro/hour for train, and 29.94 euro/hour for car. The willingness to pay for an increase in travel class is 53.40 euros. When work or the educational institution pays for the trip, the WtP becomes 81.22 euros.

For the three routes, the maximum possible market share difference shows a small effect on perceived risk for London. This effect becomes more significant for the Amsterdam - Berlin and Amsterdam - Barcelona routes. However, as there is only a 1.32 perceived risk level difference, the realistic effect of perceived risk is only a maximum of 2-3% market share difference for London; however, for Barcelona this increase to 16% point. However, in this case, all four mode-related attributes have to be at the lowest level for plane and have to be at the highest value for train. Realistically, the differences are smaller, so this market share difference will be smaller as well. **It can thus be concluded that risk perception of infection with COVID-19 has a moderate influence on market share.**

7 Limitations

The rating experiment is based on eight factors that together contribute to perceived risk. However, for respondents, other attributes could be important as well. The usage of focus groups could improve the composition of perceived risk.

Travel cost, travel time (and comfort to a lesser extent) are the most important variables for mode choice. Because of the main choice experimental set-up, with perceived risk next to the other main choice attributes, travel cost, travel time and travel comfort, an overestimation of the importance of perceived risk is possible.

Another limitation of the research is that it did not include an opting-out option. COVID-19 has several influences on travel behaviour.

The biggest limitation is the usage of the MNL model, as the MNL model has some significant shortcomings. A Mixed Logit model is better as it accommodates for the limitations of the MNL model.

The usage of the NS panel to recruit respondents, gives a bias to the sample. The market shares for train are very high in comparison to real world examples.

The rating of perceived risk was an integer, on a 5-points scale, based on the chosen attributes. The experiment is on an ordinal scale, whereas the rating is estimated on a ratio scale. There is some debate about using linear regression to estimate the rating because the respondents gave the answers on an ordinal scale.

References

Aaditya, B., & Rahul, T. (2021). Psychological impacts of covid-19 pandemic on the mode choice behaviour: A hybrid choice modelling approach. *Transport Policy*, *108*, 47-58. Retrieved from https://www.sciencedirect.com/science/article/pii/S0967070X21001359

Aamaas, B., Borken-Kleefeld, J., Peters, G.P. (2013). The climate impact of travel behavior: A german case study with illustrative mitigation options. *Environmental Science & Policy*, *33*, 273-282. Retrieved from https://www.sciencedirect.com/science/article/pii/S1462901113001366

Abdullah, M., Dias, C., Muley, D., Shahin, M. (2020). Exploring the impacts of covid-19 on travel behavior and mode preferences. *Transportation Research Interdisciplinary Perspectives*, *8*, 100255. Retrieved from https://www.sciencedirect.com/science/article/pii/S2590198220301664

Atasoy, B., Glerum, A., Bierlaire, M. (2013). Attitudes towards mode choice in switzerland. *disP - The Planning Review*, *49*(2), 101-117. Retrieved from https://doi.org/10.1080/02513625.2013.827518

Balcombe, K., Fraser, I., Harris, L. (2009). Consumer willingness to pay for in-flight service and comfort levels: A choice experiment. *Journal of Air Transport Management*, *15*(5), 221–226.

Buehler, R. (2011). Determinants of transport mode choice: a comparison of germany and the usa. *Journal of Transport Geography*, *19*(4), 644-657. Retrieved from https://www.sciencedirect.com/science/article/pii/S0966692310001171

Buehler, R., & Nobis, C. (2010). Travel behavior in aging societies: Comparison of germany and the united states. *Transportation research record*, *2182*(1), 62–70.

BusinessAM (2020). *Wie business class vliegt, waant zich nu in economy.* Retrieved from https://businessam.be/wie-business-class-vliegt-waant-zich-nu-in-economy/

Chen, J.-M. (2022). Novel statistics predict the covid-19 pandemic could terminate in 2022. *Journal of Medical Virology.*

de Haas, M., Faber, R., Hamersma, M. (2020). How covid-19 and the dutch 'intelligent lockdown' change activities, work and travel behaviour: Evidence from longitudinal data in the netherlands. *Transportation Research Interdisciplinary Perspectives*, *6*, 100150. Retrieved from https://www.sciencedirect.com/science/article/pii/S2590198220300610

De Vos, J. (2020). The effect of covid-19 and subsequent social distancing on travel behavior. *Transportation Research Interdisciplinary Perspectives*, *5*, 100121. Retrieved from https://www.sciencedirect.com/science/article/pii/S2590198220300324

De Witte, A., Hollevoet, J., Dobruszkes, F., Hubert, M., Macharis, C. (2013). Linking modal choice to motility: A comprehensive review. *Transportation Research Part A: Policy and Practice*, *49*, 329-341. Retrieved from https://www.sciencedirect.com/science/article/pii/S0965856413000165

Dobruszkes, F., Dehon, C., Givoni, M. (2014). Does european high-speed rail affect the current level of air services? an eu-wide analysis. *Transportation Research Part A: Policy and Practice*, *69*, 461-475. Retrieved from https://www.sciencedirect.com/science/article/pii/S0965856414002158

Dryhurst, S., Schneider, C.R., Kerr, J., Freeman, A.L., Recchia, G., Van Der Bles, A.M., ... Van Der Linden, S. (2020). Risk perceptions of covid-19 around the world. *Journal of Risk Research*, *23*(7-8), 994–1006.

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*, *41*(5), 428–443.

Johansson, M.V., Heldt, T., Johansson, P. (2006). The effects of attitudes and personality traits on mode choice. *Transportation Research Part A: Policy and Practice*, *40*(6), 507–525.

Joshi, A., Kale, S., Chandel, S., Pal, D.K. (2015). Likert scale: Explored and explained. *British journal of applied science & technology*, *7*(4), 396.

Karlsson, L.C., Soveri, A., Lewandowsky, S., Karlsson, L., Karlsson, H., Nolvi, S., ... Antfolk, J. (2021). Fearing the disease or the vaccine: The case of covid-19. *Personality and Individual Differences*, *172*, 110590. Retrieved from https://www.sciencedirect.com/science/article/pii/S0191886920307819

KiM (2020). *Kengetallen bereikbaarheid.* Retrieved from https://www.rwseconomie.nl/kengetallen/kengetallen-bereikbaarheid-map

Kouwenhoven, M., de Jong, G.C., Koster, P., van den Berg, V.A., Verhoef, E.T., Bates, J., Warffemius, P.M. (2014). New values of time and reliability in passenger transport in the netherlands. *Research in Transportation Economics*, *47*, 37–49.

Krishnakumari, P., & Cats, O. (2020, 05). *Virus spreading in public transport networks.*

Leppin, A., & Aro, A.R. (2009). Risk perceptions related to sars and avian influenza: theoretical foundations of current empirical research. *International journal of behavioral medicine*, *16*(1), 7–29.

Loewenstein, G.F., Weber, E.U., Hsee, C.K., Welch, N. (2001). Risk as feelings. *Psychological bulletin*, *127*(2), 267.

Louviere, J.J. (1984). Hierarchical information integration: A new method for the design and analysis of complex multlattribute judgment problems. *ACR North American Advances*.

Matiza, T. (2020). Post-covid-19 crisis travel behaviour: towards mitigating the effects of perceived risk. *Journal of Tourism Futures*.

Mertens, G., Gerritsen, L., Duijndam, S., Salemink, E., Engelhard, I.M. (2020). Fear of the coronavirus (covid-19): Predictors in an online study conducted in march 2020. *Journal of anxiety disorders*, *74*, 102258.

Molin, E. (2020). *Complex variables & non-linearity.* Retrieved from https://brightspace.tudelft.nl/d2l/le/content/195054/viewContent/1371927/View

Molin, E., & van Gelder, M. (2008). Freeway access to public transport: A hierarchical multimodal choice model. *Transportation Research Record*, *2076*(1), 106-113. Retrieved from https://doi.org/10.3141/2076-12 https://arxiv.org/abs/https://doi.org/10.3141/2076-12

Morikawa, T., Ben-Akiva, M., McFadden, D. (2002). Discrete choice models incorporating revealed preferences and psychometric data. *Advances in econometrics.* Emerald Group Publishing Limited.

Mouter, N., Hernandez, J.I., Itten, A.V. (2021). Public participation in crisis policymaking. how 30,000 dutch citizens advised their government on relaxing covid-19 lockdown measures. *PloS one*, *16*(5), e0250614.

Nieto-García, M., Muñoz-Gallego, P.A., Gonzalez-Benito, Ó. (2020). The more the merrier? understanding how travel frequency shapes willingness to pay. *Cornell Hospitality Quarterly*, *61*(4), 401–415.

Paulssen, M., Temme, D., Vij, A., Walker, J.L. (2014). Values, attitudes and travel behavior: a hierarchical latent variable mixed logit model of travel mode choice. *Transportation*, *41*(4), 873–888.

Rijksoverheid (2021). *Reisadviezen.* Retrieved from https://www.nederlandwereldwijd.nl/reizen/reisadviezen

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*(1), 84-108. Retrieved from https://www.sciencedirect.com/science/article/pii/S1755534513700307

Shamshiripour, A., Rahimi, E., Shabanpour, R., Mohammadian, A.K. (2020). How is covid-19 reshaping activity-travel behavior? evidence from a comprehensive survey in chicago. *Transportation Research Interdisciplinary Perspectives*, *7*, 100216. Retrieved from https://www.sciencedirect.com/science/article/pii/S2590198220301275

Shires, J.D., & De Jong, G.C. (2009). An international meta-analysis of values of travel time savings. *Evaluation and program planning*, *32*(4), 315–325.

Tirachini, A., & Cats, O. (2020). Covid-19 and public transportation: Current assessment, prospects, and research needs. *Journal of Public Transportation*, *22*(1), 1.

Troko, J., Myles, P., Gibson, J., Hashim, A., Enstone, J., Kingdon, S., … Van-Tam, J.N. (2011). Is public transport a risk factor for acute respiratory infection? *BMC infectious diseases*, *11*(1), 1–6.

van Goeverden, C.D. (2009). Explaining factors for train use in european long-distance travel. *Tourism and Hospitality Planning & Development*, *6*(1), 21–37.

Van Loon, R., & Rouwendal, J. (2013). Vacation behaviour: Frequency, destination choice and expenditures. *Tourism Economics*, *19*(5), 1005–1026.

Velavan, T.P., & Meyer, C.G. (2020). The covid-19 epidemic. *Tropical medicine & international health*, *25*(3), 278.

B

Ngene code & design

This appendix contains the Ngene code that was used to construct the choice for both the perceived risk rating experiment and the two main choice experiments.

Perceived risk rating experiment

```
?Design for rating experiment
design
;alts = alt1, alt2
;rows = 20
;orth = seq
;block = 4
;model:
U(alt1)=
+cr*crow[25,50,75,100]
+ma*mask[0,1,2,3]
+cl*clean[0,1,2,3]
+air*airco[0,1,2,3]
+req*require[0,1,2,3]
+in*infect[0,1,2,3]
+va*vacc[15,30,70,90]
+ad*advice[0,1,2,3]
$
```

Choice set	Crow	Mask	Clean	Airco	Require	Infect	Vacc	Advice	Block
1	25	0	0	0	0	0	15	0	3
2	75	0	2	3	3	2	30	3	4
3	25	2	1	3	2	3	70	2	4
4	75	1	3	0	3	3	70	0	4
5	50	3	0	3	1	2	70	1	1
6	50	0	3	1	2	0	70	3	1
7	25	1	1	2	0	3	30	3	3
8	50	1	0	0	3	1	90	3	1
9	100	0	0	1	1	3	15	2	2
10	75	2	1	1	1	1	90	1	3
11	25	3	3	1	1	0	30	2	4
12	100	1	3	3	0	1	15	1	2
13	100	3	1	2	3	0	15	2	2
14	50	3	2	1	2	3	30	0	1
15	50	0	3	2	1	2	90	0	1
16	75	1	0	3	2	0	90	0	4
17	100	2	1	0	2	2	30	1	3
18	75	3	2	0	0	2	70	3	3
19	100	2	2	2	0	1	90	2	2
20	25	2	2	2	3	1	15	1	2

Main choice 400-600 kilometer

```
? Efficient design 400-600km
design
;alts = train, plane, car
;rows = 12
;block = 3
;eff = (mnl,d,mean)
;model:
U(train)=
+Btt[(n,-0.1,0.05)]*TT_train[3,4.5,6]
+Btc[(n,-0.01,0.005)]*TC_train[30,165,300]
+Bcf[(n,0.5,0.25)]*CF[0,1]
+Bpr[-(n,0.05,0.025)]*PR[1,3,5]/
U(plane)=
+Btt*TT_plane[3,4,5]
+Btc*TC_plane[50,175,300]
+Bcf*CF+Bpr*PR/
U(car)=
+Btt*TT_car[4.5,6.5,8.5]
+Btc*TC_car[80,115,150]
$
```

Choice set	tt_train	tc_train	train,cf	train,pr	tt_plane	tc_plane	plane,cf	plane,pr	tt_car	tc_car	Block
1	6	30	0	5	3	50	1	1	6.5	150	1
2	4.5	300	0	3	3	175	1	1	8.5	80	3
3	4.5	300	0	3	5	175	1	5	4.5	80	1
4	4.5	30	1	3	4	175	0	3	6.5	115	3
5	3	300	1	1	3	175	0	5	8.5	80	1
6	3	165	1	1	5	50	0	5	6.5	150	2
7	6	30	1	1	3	50	0	5	6.5	150	2
8	3	165	0	5	4	300	1	1	8.5	115	2
9	4.5	300	0	3	5	50	1	3	4.5	150	3
10	6	165	1	5	4	300	0	1	4.5	115	1
11	6	165	1	5	4	300	0	3	4.5	80	3
12	3	30	0	1	5	300	1	3	8.5	115	2

Main choice 800-1200 kilometer

```
? Efficient design 800-1200km
design
;alts = train, plane, car
;rows = 12
;block = 3
;eff = (mnl,d,mean)
;model:
U(train)=
+Btt[(n,-0.1,0.05)]*TT_train[6,9,12]
+Btc[(n,-0.01,0.005)]*TC_train[50,200,350]
+Bcf[(n,0.5,0.25)]*CF[0,1]
+Bpr[(n,-0.05,0.025)]*PR[1,3,5]/
U(plane)=
Btt*TT_plane[4,5,6]
+Btc*TC_plane[50,225,400]
+Bcf*CF
+Bpr*PR/
U(car)=
+Btt*TT_car[10,13,16]
+Btc*TC_car[100,150,200]
$
```

Choice set	tt_train	tc_train	train.cf	train.pr	tt_plane	tc_plane	plane.cf	plane.pr	tt_car	tc_car	Block
1	12	200	1	5	5	400	0	3	10	200	2
2	9	350	0	3	4	225	1	1	16	100	3
3	12	50	1	5	4	50	0	1	13	200	2
4	6	350	0	1	5	225	1	5	16	100	1
5	6	200	1	1	6	225	0	5	16	150	3
6	12	200	1	3	6	400	0	3	10	100	1
7	9	50	0	5	5	50	1	1	13	150	1
8	9	50	0	5	5	400	1	1	10	150	2
9	12	50	1	1	4	50	0	5	10	200	3
10	6	350	0	3	6	400	1	3	13	100	3
11	6	200	0	3	6	50	1	3	13	150	2
12	9	350	1	1	4	225	0	5	16	200	1

C

Explanation interactions

This Appendix will explain the different steps to get to the final model. First only interactions between main variables will be tested.

Interactions between main variables

Model 1: Base model only main effects

The base model only contains main parameters, travel time, travel cost, travel comfort and perceived risk. All parameters are significant. The ρ^2 0.126, which is not very high, but a good value to start with. The LRS is 2530.875, so the model is significantly better than a random model. This is a good starting point to move on.

Name	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
ASC_PLANE	-0.359	0.0703	-5.1	3.38e-07	0.0677	-5.3	1.16e-07
ASC_TRAIN	0.35	0.0558	6.28	3.49e-10	0.0547	6.4	1.53e-10
B_COMFORT	0.281	0.0332	8.46	0	0.0323	8.7	0
B_COST	-0.00306	0.000145	-21.1	0	0.000144	-21.2	0
B_RISK	-0.22	0.0103	-21.4	0	0.01	-22	0
B_TIME	-0.167	0.00642	-26	0	0.00629	-26.5	0

Model 2: Estimating mode-specific βtime

From theory, it can be expected that time is perceived different for every different mode (KiM, 2016). Therefore for mode βtime is separately estimated. Also for comfort this is done, but not for car as for car comfort is not varied in the main choice model. As a result all betas are significant, except for βtime for plane, which is highly insignificant. The ρ^2 did increase from 0.126 to 0.134. The LRS = -2*(- 8745.118+8660.3) = 169,636. This is higher than the chi-square value, so the model is significantly better. However, if the correlation between the betas are taken into account, it can be seen that the p-value for βtime_T and βtime_C is 0.44, so there are not significantly different.

Name	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
ASC_PLANE	-0.762	0.136	-5.61	2e-08	0.131	-5.81	6.39e-09
ASC_TRAIN	0.177	0.0864	2.04	0.0409	0.0888	1.99	0.0466
B_COMFORT_P	-0.353	0.062	-5.69	1.26e-08	0.06	-5.88	4.18e-09
B_COMFORT_T	0.796	0.0608	13.1	0	0.0621	12.8	0
B_COST	-0.00318	0.000148	-21.6	0	0.000147	-21.6	0
B_RISK	-0.233	0.0107	-21.7	0	0.0105	-22.2	0
B_TIME_C	-0.176	0.00739	-23.9	0	0.00738	-23.9	0
B_TIME_P	-0.00704	0.0274	-0.257	0.797	0.0275	-0.255	0.798
B_TIME_T	-0.184	0.00916	-20.1	0	0.00905	-20.3	0

Model 3: Estimating mode-specific βrisk

Risk can also be perceived differently per mode. Therefore there will be also two betas specified, for plane and train. For car, risk is not varied. Both betas are highly significant. The LRS =-2*(-8660,3+8658,209) = 4,182, which is just higher than the value at 5% level 3.841, so it is significantly better on the 5% level but not at the 1% level. When looking at the correlations, the p-value of βrisk_P and βrisk_T is 0.0408, so it is significant at the 5% level, but not at the 1% level. The 5% level is used, so therefore both βrisk are specified for each mode.

Name	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
ASC_PLANE	-0.721	0.137	-5.24	1.58e-07	0.133	-5.41	6.45e-08
ASC_TRAIN	0.0881	0.0964	0.914	0.361	0.0989	0.891	0.373
B_COMFORT_P	-0.4	0.0664	-6.03	1.69e-09	0.0655	-6.11	1e-09
B_COMFORT_T	0.82	0.0618	13.3	0	0.0632	13	0
B_COST	-0.00322	0.000149	-21.7	0	0.000148	-21.7	0
B_RISK_P	-0.269	0.0209	-12.9	0	0.0207	-13	0
B_RISK_T	-0.203	0.0182	-11.1	0	0.0181	-11.2	0
B_TIME_C	-0.179	0.00747	-23.9	0	0.00745	-24	0
B_TIME_P	0.00654	0.0282	0.232	0.817	0.0285	0.23	0.818
B_TIME_T	-0.188	0.00943	-20	0	0.00928	-20.3	0

Model 4: Estimating interaction between comfort and time for both plane and train
Both betas are insignificant, and both values are very small. The LRS = -2*(-8658.724+8655.665)=6.1, so this suggest this model is better. At the same time, the ρ^2 value is still 0.135. Because of the reason, the interaction will not be taken into account.

Name	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
ASC_PLANE	-0.788	0.115	-6.87	6.43e-12	0.111	-7.12	1.11e-12
ASC_TRAIN	0.0847	0.0848	0.999	0.318	0.0871	0.972	0.331
B_COMFORT_P	-0.38	0.0658	-5.78	7.48e-09	0.0651	-5.85	5.02e-09
B_COMFORT_T	0.796	0.0623	12.8	0	0.0648	12.3	0
B_COMFORT_TIME_P	8.34e-05	6.78e-05	1.23	0.219	7.07e-05	1.18	0.238
B_COMFORT_TIME_T	-3.02e-05	3.96e-05	-0.764	0.445	3.98e-05	-0.759	0.448
B_COST	-0.00335	0.000288	-11.6	0	0.000298	-11.2	0
B_RISK_P	-0.27	0.0196	-13.7	0	0.0193	-14	0
B_RISK_T	-0.21	0.018	-11.6	0	0.018	-11.6	0
B_TIME	-0.181	0.00674	-26.8	0	0.00666	-27.1	0

Model 5: Combining comfort, mode specific time and risk + addition beta time for plane
Mode-specific comfort does give a counter-intuitive outcome. This can partly explained by the survey design, because in all choice sets the class of plane and train were always the opposite (e.g., when train is 1^{st} class then for plane class is economy and opposite). Therefore, comfort will be combined to a single parameter. Beta time for plane is again added, to test if it is significant. It will be included to test if the parameter becomes significant with other parameters. It is also tested if the interaction between risk and time also is mode-specific. Both parameters for train and plane are highly significant, so will be in the model. At the same time, beta risk for plane and train are highly insignificant.

Name	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
ASC_PLANE	-1.65	0.195	-8.46	0	0.194	-8.52	0
ASC_TRAIN	-0.216	0.096	-2.26	0.0241	0.0974	-2.22	0.0263
B_COMFORT	0.291	0.0362	8.03	1.11e-15	0.0355	8.19	2.22e-16
B_COST	-0.00327	0.000146	-22.4	0	0.000142	-23.1	0
B_RISK_P	0.0829	0.0823	1.01	0.314	0.0802	1.03	0.301
B_RISK_T	0.0119	0.0361	0.33	0.741	0.0362	0.329	0.742
B_TIME_C	-0.168	0.00743	-22.6	0	0.00733	-22.9	0
B_TIME_P	0.0687	0.0462	1.49	0.137	0.0462	1.49	0.137
B_TIME_RISK_P	-0.0625	0.0192	-3.26	0.00113	0.0188	-3.32	0.000906
B_TIME_RISK_T	-0.0405	0.00482	-8.4	0	0.00479	-8.45	0
B_TIME_T	-0.0968	0.0119	-8.14	4.44e-16	0.0118	-8.18	2.22e-16

Model 6: Final base model
This model is the starting point as the basis model for socio-demographic interactions and other travel behaviour interactions. Both betas for risk are removed, so risk is only included as an interaction with time. The final LL=-8682.564, with a ρ^2 of 0.133.

Name	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
ASC_PLANE	-1.51	0.137	-11.1	0	0.134	-11.3	0
ASC_TRAIN	-0.197	0.078	-2.53	0.0114	0.0783	-2.52	0.0118
B_COMFORT	0.281	0.0347	8.11	4.44e-16	0.0339	8.3	0
B_COST	-0.00326	0.000145	-22.4	0	0.000142	-23	0
B_TIME_C	-0.168	0.00734	-22.9	0	0.00728	-23.1	0
B_TIME_P	0.0368	0.0321	1.15	0.252	0.0325	1.14	0.256
B_TIME_RISK_P	-0.0437	0.0046	-9.51	0	0.00442	-9.9	0
B_TIME_RISK_T	-0.039	0.00235	-16.6	0	0.00228	-17.1	0
B_TIME_T	-0.0991	0.0107	-9.22	0	0.0105	-9.42	0

Interactions with socio-demographic variables

Model interaction age and cost

There are three extra parameters estimated, for every dummy level of age. The LRS = -2*(-8682.564+8627.256)=110.616, so this model is significantly better. The ρ^2 0.138.

Name	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
ASC_PLANE	-1.51	0.137	-11	0	0.134	-11.2	0
ASC_TRAIN	-0.217	0.0783	-2.78	0.00551	0.0786	-2.76	0.0057
B_AGE_65_AND_OLDER	0.00612	0.00126	4.88	1.08e-06	0.00123	4.97	6.75e-07
B_AGE_COST_20_40	0.00276	0.00127	2.17	0.0301	0.00125	2.21	0.0268
B_AGE_COST_40_65	0.00554	0.00126	4.41	1.05e-05	0.00123	4.49	7.11e-06
B_COMFORT	0.29	0.0348	8.32	0	0.0341	8.5	0
B_COST	-0.00836	0.00124	-6.75	1.51e-11	0.00121	-6.88	6.03e-12
B_TIME_C	-0.17	0.00738	-23.1	0	0.00733	-23.2	0
B_TIME_P	0.0271	0.0323	0.841	0.4	0.0327	0.83	0.406
B_TIME_RISK_P	-0.043	0.00461	-9.32	0	0.00444	-9.68	0
B_TIME_RISK_T	-0.0394	0.00235	-16.7	0	0.00229	-17.2	0
B_TIME_T	-0.0991	0.0108	-9.2	0	0.0106	-9.39	0

Model interaction age and perceived risk

There are three extra parameters estimated, for every dummy level of age. LRS = -2*(-8682.564+8674.615)= 15,898. This is higher than the chi-square, however, the interactions all do not turn out to be significant. So this interaction will not be taken into account.

Name	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
ASC_PLANE	-1.51	0.137	-11	0	0.134	-11.3	0
ASC_TRAIN	-0.195	0.078	-2.5	0.0123	0.0783	-2.49	0.0126
B_AGE_PR_20_40	-0.0137	0.00857	-1.6	0.11	0.00831	-1.65	0.0993
B_AGE_PR_40_65	-0.015	0.00845	-1.78	0.0755	0.00819	-1.83	0.0667
B_AGE_PR_65_AND_OLDER	-0.00531	0.00838	-0.634	0.526	0.00811	-0.655	0.513
B_COMFORT	0.28	0.0347	8.07	6.66e-16	0.0339	8.26	2.22e-16
B_COST	-0.00326	0.000146	-22.4	0	0.000142	-23	0
B_TIME_C	-0.168	0.00734	-22.9	0	0.00728	-23.1	0
B_TIME_P	0.0366	0.0321	1.14	0.255	0.0324	1.13	0.26
B_TIME_RISK_P	-0.0331	0.00932	-3.55	0.000379	0.00908	-3.65	0.000262
B_TIME_RISK_T	-0.0288	0.00841	-3.42	0.000621	0.00811	-3.55	0.000383
B_TIME_T	-0.099	0.0108	-9.21	0	0.0105	-9.41	0

Model interaction age and comfort

There are three extra parameters estimated, for every dummy level of age. The LRS = -2*(-8682.564+8679.447)= 6,234, which is lower than the chi-square value. None of the parameters are significant on the 1% level, 1 on the 5% level.

Name	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
ASC_PLANE	-1.39	0.0875	-15.9	0	0.0849	-16.4	0
ASC_TRAIN	-0.205	0.0777	-2.64	0.00823	0.078	-2.63	0.00846
B_AGE_CF_20_40	-0.243	0.185	-1.31	0.191	0.184	-1.32	0.187
B_AGE_CF_40_65	-0.251	0.183	-1.37	0.171	0.182	-1.38	0.169
B_AGE_CF_65_AND_OLDER	-0.355	0.183	-1.94	0.0523	0.182	-1.95	0.0513
B_COMFORT	0.582	0.18	3.23	0.00126	0.18	3.24	0.00122
B_COST	-0.00324	0.000144	-22.5	0	0.000141	-23	0
B_TIME_C	-0.169	0.00727	-23.3	0	0.00719	-23.5	0
B_TIME_RISK_P	-0.0408	0.00379	-10.8	0	0.00364	-11.2	0
B_TIME_RISK_T	-0.0394	0.00233	-16.9	0	0.00228	-17.3	0
B_TIME_T	-0.1	0.0107	-9.32	0	0.0105	-9.54	0

Model interaction gender and cost

There is only 1 extra parameter. The LRS = -2*(-8682.564+8682.97) = -0,812 The parameter is not significant.

Name	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
ASC_PLANE	-1.39	0.0876	-15.9	0	0.085	-16.4	0
ASC_TRAIN	-0.206	0.0777	-2.65	0.00811	0.0779	-2.64	0.00829
B_COMFORT	0.296	0.0323	9.17	0	0.031	9.53	0
B_COST	-0.00333	0.000197	-16.9	0	0.000195	-17	0
B_GENDER_COST	0.000191	0.00027	0.71	0.478	0.00027	0.71	0.478
B_TIME_C	-0.169	0.00727	-23.3	0	0.00719	-23.5	0
B_TIME_RISK_P	-0.0408	0.00379	-10.8	0	0.00365	-11.2	0
B_TIME_RISK_T	-0.0394	0.00233	-16.9	0	0.00227	-17.3	0
B_TIME_T	-0.0999	0.0107	-9.32	0	0.0105	-9.55	0

Model interaction gender and time

There is only one extra parameter. LRS = -2*(-8682.564+8683.22) = -1,312

Name	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
ASC_PLANE	-1.39	0.0885	-15.7	0	0.086	-16.2	0
ASC_TRAIN	-0.206	0.0777	-2.66	0.00792	0.0779	-2.65	0.00809
B_COMFORT	0.296	0.0323	9.16	0	0.0311	9.52	0
B_COST	-0.00323	0.000144	-22.5	0	0.000141	-22.9	0
B_GENDER_TIME	-0.000455	0.00803	-0.0566	0.955	0.00839	-0.0542	0.957
B_TIME_C	-0.169	0.00824	-20.5	0	0.00823	-20.5	0
B_TIME_RISK_P	-0.0407	0.0038	-10.7	0	0.00365	-11.1	0
B_TIME_RISK_T	-0.0394	0.00233	-16.9	0	0.00227	-17.3	0
B_TIME_T	-0.0996	0.0114	-8.73	0	0.0112	-8.9	0

Model gender and perceived risk

There is only one extra parameter. LRS = -2*(-8683.222+8683.22) = 0 Parameter is highly significant

Name	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
ASC_PLANE	-1.39	0.0875	-15.9	0	0.0849	-16.4	0
ASC_TRAIN	-0.206	0.0777	-2.66	0.00793	0.0779	-2.65	0.0081
B_COMFORT	0.296	0.0323	9.17	0	0.031	9.53	0
B_COST	-0.00323	0.000144	-22.5	0	0.000141	-23	0
B_GENDER_PR	-0.0267	0.00125	-21.4	0	0.00122	-22	0
B_TIME_C	-0.169	0.00727	-23.3	0	0.00719	-23.5	0
B_TIME_RISK_P	-0.014	0.00287	-4.89	1.02e-06	0.00276	-5.09	3.67e-07
B_TIME_RISK_T	-0.0127	0.0023	-5.52	3.35e-08	0.00222	-5.72	1.08e-08
B_TIME_T	-0.0998	0.0107	-9.31	0	0.0105	-9.55	0

Model interaction income and cost

There are 4 extra parameters. LRS = -2*(-8682.564+8664.316) = 36,496 The parameters turn out to be significant.

Name	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
ASC_PLANE	-1.4	0.0876	-16	0	0.0851	-16.5	0
ASC_TRAIN	-0.213	0.0778	-2.74	0.00613	0.078	-2.73	0.00626
B_COMFORT	0.298	0.0323	9.22	0	0.031	9.6	0
B_COST	-0.00629	0.000569	-11	0	0.000573	-11	0
B_INCOME_COST_100_AND_MORE	0.00322	0.000754	4.27	1.96e-05	0.000741	4.35	1.39e-05
B_INCOME_COST_20_40	0.00327	0.000615	5.31	1.09e-07	0.000618	5.29	1.25e-07
B_INCOME_COST_40_100	0.00297	0.000612	4.85	1.24e-06	0.000616	4.82	1.44e-06
B_INCOME_COST_NOT_SAY	0.00375	0.00064	5.86	4.54e-09	0.000647	5.8	6.83e-09
B_TIME_C	-0.17	0.00728	-23.3	0	0.00721	-23.6	0
B_TIME_RISK_P	-0.0408	0.00379	-10.8	0	0.00365	-11.2	0
B_TIME_RISK_T	-0.0394	0.00233	-16.9	0	0.00227	-17.3	0
B_TIME_T	-0.0999	0.0107	-9.31	0	0.0105	-9.55	0

Model education and PR

There are three extra parameter LRS = -2*(-8683.222+-8664.719) = 37.006 All parameters are significant.

Name	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
ASC_PLANE	-1.39	0.0876	-15.9	0	0.0849	-16.4	0
ASC_TRAIN	-0.199	0.0778	-2.56	0.0106	0.078	-2.55	0.0107
B_COMFORT	0.293	0.0323	9.05	0	0.0311	9.42	0
B_COST	-0.00323	0.000144	-22.4	0	0.000141	-22.9	0
B_EDU_HBO_PR	-0.0116	0.00313	-3.69	0.000228	0.00311	-3.72	0.0002
B_EDU_OTHER_PR	-0.0177	0.00344	-5.14	2.72e-07	0.00336	-5.26	1.42e-07
B_EDU_WO_PR	-0.0165	0.00331	-4.99	6.11e-07	0.00331	-4.98	6.26e-07
B_TIME_C	-0.169	0.00727	-23.2	0	0.00718	-23.5	0
B_TIME_RISK_P	-0.0299	0.00421	-7.1	1.23e-12	0.00409	-7.31	2.63e-13
B_TIME_RISK_T	-0.0291	0.00286	-10.2	0	0.00283	-10.3	0
B_TIME_T	-0.1	0.0107	-9.32	0	0.0105	-9.55	0

Interaction with travel behaviour questions

Model worry and perceived risk

There are two extra parameters. The parameters are not significant. LRS = -2*(-8683.222+8682.886) = 0.672, so not significant.

Name	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
ASC_PLANE	-1.39	0.0875	-15.9	0	0.0849	-16.4	0
ASC_TRAIN	-0.207	0.0777	-2.66	0.00784	0.0779	-2.65	0.00802
B_COMFORT	0.296	0.0323	9.16	0	0.031	9.52	0
B_COST	-0.00323	0.000144	-22.5	0	0.000141	-23	0
B_TIME_C	-0.169	0.00727	-23.3	0	0.00719	-23.5	0
B_TIME_RISK_P	-0.0401	0.00394	-10.2	0	0.00379	-10.6	0
B_TIME_RISK_T	-0.0387	0.00257	-15	0	0.00254	-15.3	0
B_TIME_T	-0.0998	0.0107	-9.31	0	0.0105	-9.55	0
B_WORRY_MORE_RISK	0.000856	0.00498	0.172	0.863	0.00475	0.18	0.857
B_WORRY_SAME_RISK	-0.00186	0.00249	-0.747	0.455	0.00248	-0.752	0.452

Model payment and cost
There are three extra parameters. Higher than chi-square so significantly better. Only payment by work or education is significant. The sign is positive (as expected, so LRS = -2*(-8683.222+8674.014) = 18.416

Name	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
ASC_PLANE	-1.4	0.0877	-15.9	0	0.085	-16.4	0
ASC_TRAIN	-0.204	0.0778	-2.62	0.00869	0.0781	-2.62	0.00891
B_COMFORT	0.296	0.0323	9.15	0	0.0311	9.51	0
B_COMPANY_COST_FAMILY	-0.0022	0.000458	-4.81	1.51e-06	0.000461	-4.78	1.79e-06
B_COMPANY_COST_FRIENDS	-0.00106	0.000354	-2.98	0.0029	0.000356	-2.96	0.00306
B_COMPANY_COST_OTHER	-0.000616	0.000596	-1.03	0.301	0.000607	-1.02	0.31
B_COMPANY_COST_PARTNER	-0.00198	0.000452	-4.37	1.26e-05	0.000454	-4.35	1.35e-05
B_COST	-0.00213	0.00029	-7.33	2.23e-13	0.000292	-7.3	2.92e-13
B_TIME_C	-0.169	0.00728	-23.3	0	0.00719	-23.6	0
B_TIME_RISK_P	-0.0407	0.00379	-10.7	0	0.00365	-11.2	0
B_TIME_RISK_T	-0.0395	0.00233	-16.9	0	0.00228	-17.3	0
B_TIME_T	-0.1	0.0107	-9.33	0	0.0105	-9.57	0

Model travel company and cost
There are four extra parameters. LRS = -2*(-8683.222+8667.473) = 31.498. Only B_COMPANY_COST_OTHER is not significant, but this give no information anyhow, all the other parameters are significant and show face validity.

Name	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
ASC_PLANE	-1.4	0.0877	-15.9	0	0.085	-16.4	0
ASC_TRAIN	-0.204	0.0778	-2.62	0.00869	0.0781	-2.62	0.00891
B_COMFORT	0.296	0.0323	9.15	0	0.0311	9.51	0
B_COMPANY_COST_FAMILY	-0.0022	0.000458	-4.81	1.51e-06	0.000461	-4.78	1.79e-06
B_COMPANY_COST_FRIENDS	-0.00106	0.000354	-2.98	0.0029	0.000356	-2.96	0.00306
B_COMPANY_COST_OTHER	-0.000616	0.000596	-1.03	0.301	0.000607	-1.02	0.31
B_COMPANY_COST_PARTNER	-0.00198	0.000452	-4.37	1.26e-05	0.000454	-4.35	1.35e-05
B_COST	-0.00213	0.00029	-7.33	2.23e-13	0.000292	-7.3	2.92e-13
B_TIME_C	-0.169	0.00728	-23.3	0	0.00719	-23.6	0
B_TIME_RISK_P	-0.0407	0.00379	-10.7	0	0.00365	-11.2	0
B_TIME_RISK_T	-0.0395	0.00233	-16.9	0	0.00228	-17.3	0
B_TIME_T	-0.1	0.0107	-9.33	0	0.0105	-9.57	0

Model travel company and perceived risk
There are four extra parameters. LRS = -2*(-8683.222 +8653.575) = 59.294. Only B_COMPANY_PR_OTHER is not significant, but this give no information anyhow, all the other parameters are significant and show face validity.

Name	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
ASC_PLANE	-1.39	0.0875	-15.8	0	0.0847	-16.4	0
ASC_TRAIN	-0.191	0.0779	-2.45	0.0143	0.078	-2.45	0.0144
B_COMFORT	0.295	0.0323	9.12	0	0.0311	9.46	0
B_COMPANY_PR_FAMILY	-0.00722	0.00381	-1.9	0.058	0.00368	-1.96	0.05
B_COMPANY_PR_FRIENDS	-0.0173	0.00309	-5.59	2.21e-08	0.00301	-5.76	8.45e-09
B_COMPANY_PR_OTHER	0.0078	0.00477	1.64	0.102	0.00473	1.65	0.0993
B_COMPANY_PR_PARTNER	-0.0204	0.00412	-4.94	7.91e-07	0.00414	-4.91	9.02e-07
B_COST	-0.00323	0.000144	-22.4	0	0.000141	-22.9	0
B_TIME_C	-0.169	0.00728	-23.2	0	0.00717	-23.5	0
B_TIME_RISK_P	-0.0302	0.00435	-6.94	3.94e-12	0.00422	-7.16	8.18e-13
B_TIME_RISK_T	-0.0293	0.00309	-9.47	0	0.00296	-9.9	0
B_TIME_T	-0.1	0.0107	-9.35	0	0.0105	-9.59	0

Model work status and cost
There are four extra parameters. LRS = -2*(-8683.222+8641.657) = 83.13, ρ^2 0.136. Only B_WORK_OTHER_COST

and B_WORK_RETIRED is not significant, but this give no information anyhow, all the other parameters are significant and show face validity. Student is not fully significant on the 1%, but it is on the 5% level.

Name	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
ASC_PLANE	-1.41	0.0877	-16.1	0	0.0852	-16.6	0
ASC_TRAIN	-0.222	0.0779	-2.85	0.0044	0.0781	-2.84	0.00449
B_COMFORT	0.299	0.0323	9.24	0	0.031	9.63	0
B_COST	-0.00361	0.000594	-6.07	1.24e-09	0.000606	-5.96	2.6e-09
B_TIME_C	-0.171	0.00729	-23.4	0	0.00722	-23.7	0
B_TIME_RISK_P	-0.0409	0.0038	-10.8	0	0.00365	-11.2	0
B_TIME_RISK_T	-0.0395	0.00233	-16.9	0	0.00228	-17.3	0
B_TIME_T	-0.0998	0.0107	-9.3	0	0.0105	-9.53	0
B_WORK_NOT_COST	-0.00307	0.000762	-4.03	5.65e-05	0.000769	-3.99	6.69e-05
B_WORK_OTHER_COST	0.000493	0.00114	0.434	0.664	0.00116	0.423	0.672
B_WORK_RETIRED_COST	0.000207	0.000623	0.333	0.739	0.000635	0.326	0.744
B_WORK_STUDENT_COST	0.00152	0.000633	2.41	0.0161	0.000645	2.36	0.0181

Model travel purpose and cost
There are four extra parameters. LRS = -2*(-8683.222+8672.635) = 21.174, ρ^2 0.134. Only B_PURPOSE_WORK_COST is significant. The value is positive as expected.

Name	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
ASC_PLANE	-1.39	0.0877	-15.9	0	0.085	-16.4	0
ASC_TRAIN	-0.201	0.0778	-2.58	0.00989	0.078	-2.57	0.0101
B_COMFORT	0.297	0.0323	9.2	0	0.0311	9.55	0
B_COST	-0.00346	0.000171	-20.2	0	0.000169	-20.5	0
B_PURPOSE_OTHER_COST	0.000766	0.000419	1.83	0.0673	0.000426	1.8	0.0724
B_PURPOSE_SCHOOL_COST	0.00101	0.00103	0.981	0.327	0.000979	1.03	0.301
B_PURPOSE_VFR_COST	-9.21e-05	0.00041	-0.225	0.822	0.000408	-0.226	0.822
B_PURPOSE_WORK_COST	0.00264	0.000616	4.28	1.87e-05	0.000618	4.27	1.99e-05
B_TIME_C	-0.169	0.00728	-23.2	0	0.00718	-23.5	0
B_TIME_RISK_P	-0.0409	0.0038	-10.8	0	0.00365	-11.2	0
B_TIME_RISK_T	-0.0393	0.00233	-16.9	0	0.00227	-17.3	0
B_TIME_T	-0.101	0.0107	-9.38	0	0.0105	-9.62	0

Model travel purpose and time
There are four extra parameters. LRS = -2*(-8683.222+8653.351) = 59.742, ρ^2 0.136. Only B_PURPOSE_WORK_TIME is significant. The value is negative as expected.

Name	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
ASC_PLANE	-1.4	0.0878	-16	0	0.085	-16.5	0
ASC_TRAIN	-0.214	0.0778	-2.75	0.00599	0.078	-2.74	0.00611
B_COMFORT	0.297	0.0323	9.2	0	0.0311	9.56	0
B_COST	-0.00325	0.000144	-22.5	0	0.000141	-23	0
B_PURPOSE_OTHER_TIME	0.00212	0.00833	0.255	0.799	0.00895	0.237	0.812
B_PURPOSE_SCHOOL_TIME	0.00357	0.0211	0.169	0.866	0.0217	0.165	0.869
B_PURPOSE_VFR_TIME	0.0122	0.00788	1.55	0.122	0.00792	1.54	0.124
B_PURPOSE_WORK_TIME	-0.097	0.0136	-7.14	9.63e-13	0.0139	-6.98	2.95e-12
B_TIME_C	-0.168	0.0075	-22.4	0	0.00746	-22.5	0
B_TIME_RISK_P	-0.0413	0.00381	-10.8	0	0.00366	-11.3	0
B_TIME_RISK_T	-0.0395	0.00233	-16.9	0	0.00228	-17.4	0
B_TIME_T	-0.0977	0.0109	-8.98	0	0.0107	-9.18	0

Model travel frequency and perceived risk
There are three extra parameters. LRS = -2*(-8683.222+8661.571) = 43.302, ρ^2 0.135. All parameters. Maybe strange outcome, that people that travel 1_3 times per year are more likely to have perceived risk. However, if people are traveling more than 3 time per year, perceived risk is likely to decrease.

Name	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
ASC_PLANE	-1.39	0.0876	-15.9	0	0.0849	-16.4	0
ASC_TRAIN	-0.199	0.0778	-2.56	0.0106	0.078	-2.55	0.0107
B_1_3_PR	-0.00965	0.0027	-3.57	0.000358	0.00266	-3.63	0.000286
B_4_7_PR	0.0154	0.00407	3.78	0.000156	0.00402	3.82	0.000132
B_COMFORT	0.294	0.0323	9.09	0	0.0311	9.45	0
B_COST	-0.00323	0.000144	-22.4	0	0.000141	-22.9	0
B_MORE_8_PR	0.0155	0.00577	2.68	0.00729	0.00578	2.68	0.00735
B_TIME_C	-0.169	0.00728	-23.2	0	0.00718	-23.5	0
B_TIME_RISK_P	-0.0393	0.00394	-9.97	0	0.00378	-10.4	0
B_TIME_RISK_T	-0.0384	0.00258	-14.9	0	0.00254	-15.1	0
B_TIME_T	-0.1	0.0107	-9.32	0	0.0105	-9.54	0

Model preference mode and ASC both distance classes

There are 9 extra parameters. The LRS = -2*(-8683.222+ 7745.085) = 1876,274, ρ^2 0.226. B_PREF_CAR_P and B_ B_PREF_TRAIN_P are not significant.

Name	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
ASC_PLANE	-2.12	0.169	-12.6	0	0.171	-12.4	0
ASC_TRAIN	-0.236	0.146	-1.62	0.106	0.152	-1.55	0.122
B_COMFORT	0.345	0.0344	10	0	0.033	10.5	0
B_COST	-0.00366	0.000155	-23.7	0	0.000152	-24.1	0
B_PREF_CAR_C	0.462	0.0434	10.6	0	0.0453	10.2	0
B_PREF_CAR_P	-0.0804	0.0502	-1.6	0.109	0.0511	-1.57	0.115
B_PREF_CAR_T	-0.382	0.0446	-8.56	0	0.0466	-8.19	2.22e-16
B_PREF_PLANE_C	-0.288	0.0438	-6.57	5.11e-11	0.0458	-6.29	3.18e-10
B_PREF_PLANE_P	0.683	0.0476	14.4	0	0.0488	14	0
B_PREF_PLANE_T	-0.395	0.043	-9.2	0	0.0452	-8.74	0
B_PREF_TRAIN_C	-0.216	0.0422	-5.13	2.93e-07	0.0443	-4.88	1.08e-06
B_PREF_TRAIN_P	-0.0657	0.047	-1.4	0.162	0.0482	-1.36	0.172
B_PREF_TRAIN_T	0.282	0.0404	6.98	2.98e-12	0.043	6.56	5.39e-11
B_TIME_C	-0.193	0.00782	-24.7	0	0.00775	-24.9	0
B_TIME_RISK_P	-0.0453	0.00405	-11.2	0	0.00386	-11.7	0
B_TIME_RISK_T	-0.0441	0.00251	-17.6	0	0.00245	-18	0
B_TIME_T	-0.116	0.0115	-10.1	0	0.0113	-10.3	0

Model with all interactions, then stepwise backwards elimination

Model with all significant interactions included

All of the significant interactions (from the previous separate models) are included with the base model. This gives the following results. Not all parameters are significant. ρ^2 is 0.219. Final LL -7820.572.

Name	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
ASC_PLANE	-1.43	0.0951	-15	0	0.0917	-15.6	0
ASC_TRAIN	-0.344	0.0852	-4.04	5.4e-05	0.0858	-4.01	6.15e-05
B_1_3_PR	-0.00157	0.00306	-0.514	0.607	0.00303	-0.52	0.603
B_4_7_PR	0.0185	0.00457	4.05	5.11e-05	0.00423	4.38	1.19e-05
B_AGE_COST_20_40	0.00228	0.00135	1.68	0.0925	0.0013	1.75	0.0807
B_AGE_COST_40_65	0.00484	0.00142	3.42	0.00063	0.00137	3.54	0.000408
B_AGE_COST_65_AND_OLDER	0.00523	0.0015	3.49	0.000481	0.00145	3.6	0.000322
B_COMFORT	0.336	0.0344	9.75	0	0.0329	10.2	0
B_COMPANY_COST_FAMILY	-0.000772	0.000515	-1.5	0.133	0.000518	-1.49	0.136
B_COMPANY_COST_FRIENDS	-0.000775	0.000396	-1.96	0.0503	0.000398	-1.95	0.0515
B_COMPANY_COST_OTHER	-0.000344	0.000717	-0.48	0.632	0.000712	-0.483	0.629
B_COMPANY_COST_PARTNER	-0.00101	0.000504	-2	0.0455	0.000502	-2.01	0.0442
B_COMPANY_PR_FAMILY	-0.00496	0.0042	-1.18	0.238	0.00401	-1.24	0.216
B_COMPANY_PR_FRIENDS	-0.0116	0.0034	-3.43	0.000607	0.00322	-3.62	0.000294
B_COMPANY_PR_OTHER	0.0163	0.00523	3.11	0.00185	0.00521	3.13	0.00176
B_COMPANY_PR_PARTNER	-0.0102	0.00443	-2.31	0.0208	0.00424	-2.41	0.0158
B_COST	-0.00856	0.00167	-5.12	3.08e-07	0.00163	-5.26	1.45e-07
B_EDU_HBO_PR	-0.0145	0.00339	-4.27	1.95e-05	0.00334	-4.33	1.48e-05
B_EDU_OTHER_PR	-0.0157	0.00396	-3.97	7.12e-05	0.00379	-4.16	3.23e-05
B_EDU_WO_PR	-0.0222	0.00365	-6.1	1.06e-09	0.00358	-6.21	5.13e-10
B_GENDER_PR	-0.0185	0.00249	-7.43	1.11e-13	0.00236	-7.82	5.33e-15
B_INCOME_COST_100_AND_MORE	0.000486	0.000918	0.529	0.596	0.000908	0.536	0.592
B_INCOME_COST_20_40	0.00074	0.000741	0.998	0.318	0.000749	0.988	0.323
B_INCOME_COST_40_100	0.000293	0.000764	0.383	0.701	0.000766	0.382	0.702
B_INCOME_COST_NOT_SAY	0.000924	0.00078	1.18	0.236	0.000789	1.17	0.242
B_MORE_8_PR	0.0161	0.00655	2.45	0.0141	0.00609	2.64	0.00827
B_PAYMENT_ELSE_COST	0.000253	0.00127	0.199	0.842	0.00135	0.188	0.851
B_PAYMENT_OTHER_COST	0.000334	0.000619	0.54	0.589	0.000621	0.539	0.59
B_PAYMENT_WORKEDU_COST	0.00166	0.000745	2.23	0.0257	0.000736	2.26	0.0239
B_PREF_CAR_P	-1.69	0.108	-15.7	0	0.105	-16	0
B_PREF_CAR_T	-1.52	0.0959	-15.9	0	0.0966	-15.7	0
B_PREF_PLANE_P	1.05	0.0982	10.7	0	0.097	10.9	0
B_PREF_PLANE_T	-0.078	0.11	-0.711	0.477	0.112	-0.697	0.486
B_PREF_TRAIN_P	-0.497	0.0894	-5.56	2.65e-08	0.0908	-5.48	4.36e-08
B_PREF_TRAIN_T	1.12	0.0732	15.4	0	0.0755	14.9	0
B_PURPOSE_OTHER_TIME	0.0182	0.00868	2.1	0.0357	0.00896	2.03	0.042
B_PURPOSE_SCHOOL_TIME	0.0135	0.0218	0.618	0.536	0.0221	0.61	0.542
B_PURPOSE_VFR_TIME	0.0118	0.00858	1.38	0.168	0.00885	1.34	0.182
B_PURPOSE_WORK_TIME	-0.0745	0.0143	-5.19	2.09e-07	0.0146	-5.1	3.38e-07
B_TIME_C	-0.19	0.00794	-23.9	0	0.00786	-24.1	0
B_TIME_RISK_P	-0.00863	0.00324	-2.67	0.00765	0.00311	-2.78	0.00548
B_TIME_RISK_T	-0.00983	0.00264	-3.72	0.000201	0.0025	-3.94	8.25e-05
B_TIME_T	-0.112	0.0116	-9.72	0	0.0113	-9.92	0
B_WORK_NOT_COST	-0.00029	0.000948	-0.306	0.759	0.000977	-0.297	0.766
B_WORK_OTHER_COST	0.000227	0.00119	0.19	0.849	0.00119	0.191	0.848
B_WORK_RETIRED_COST	0.000661	0.00071	0.932	0.351	0.000729	0.908	0.364
B_WORK_STUDENT_COST	0.000929	0.000749	1.24	0.214	0.000774	1.2	0.23

Elimination 1

First all parameters with a p-value of 0.3 are deleted. This gives elimination_1. In total 14 parameter have been removed, so there are 33 parameters. This gives the following results. Not all parameters are significant. ρ^2 is 0.218. Final LL -7825.372

Name	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
ASC_PLANE	-1.43	0.0947	-15.1	0	0.0911	-15.7	0
ASC_TRAIN	-0.353	0.0838	-4.22	2.48e-05	0.0843	-4.19	2.79e-05
B_4_7_PR	0.0194	0.00434	4.46	8.08e-06	0.00399	4.86	1.18e-06
B_AGE_COST_20_40	0.00284	0.0013	2.17	0.0297	0.00125	2.26	0.0235
B_AGE_COST_40_65	0.00575	0.0013	4.41	1.04e-05	0.00126	4.58	4.76e-06
B_AGE_COST_65_AND_OLDER	0.00646	0.0013	4.96	7.19e-07	0.00125	5.16	2.5e-07
B_COMFORT	0.335	0.0344	9.73	0	0.0329	10.2	0
B_COMPANY_COST_FAMILY	-0.000884	0.000487	-1.82	0.0691	0.000489	-1.81	0.0706
B_COMPANY_COST_FRIENDS	-0.000759	0.000358	-2.12	0.0339	0.000353	-2.15	0.0318
B_COMPANY_COST_PARTNER	-0.000995	0.000476	-2.09	0.0366	0.00047	-2.12	0.0341
B_COMPANY_PR_FAMILY	-0.00502	0.00419	-1.2	0.232	0.00401	-1.25	0.211
B_COMPANY_PR_FRIENDS	-0.0115	0.00339	-3.4	0.000684	0.00321	-3.59	0.000335
B_COMPANY_PR_OTHER	0.0166	0.00515	3.23	0.00125	0.00514	3.24	0.0012
B_COMPANY_PR_PARTNER	-0.0102	0.00442	-2.3	0.0215	0.00423	-2.4	0.0165
B_COST	-0.00832	0.00132	-6.32	2.58e-10	0.00126	-6.59	4.54e-11
B_EDU_HBO_PR	-0.0147	0.00337	-4.35	1.36e-05	0.00333	-4.41	1.03e-05
B_EDU_OTHER_PR	-0.0165	0.00376	-4.38	1.16e-05	0.00354	-4.65	3.29e-06
B_EDU_WO_PR	-0.0227	0.00357	-6.37	1.93e-10	0.00354	-6.42	1.36e-10
B_GENDER_PR	-0.0187	0.00242	-7.71	1.22e-14	0.00229	-8.18	2.22e-16
B_MORE_8_PR	0.0168	0.00642	2.62	0.00874	0.00599	2.81	0.00496
B_PAYMENT_WORKEDU_COST	0.00172	0.000734	2.35	0.019	0.00072	2.39	0.0168
B_PREF_CAR_P	-1.68	0.107	-15.6	0	0.105	-16	0
B_PREF_CAR_T	-1.51	0.0949	-15.9	0	0.0957	-15.8	0
B_PREF_PLANE_P	1.1	0.0799	13.8	0	0.0783	14.1	0
B_PREF_TRAIN_P	-0.491	0.0891	-5.52	3.45e-08	0.0904	-5.44	5.38e-08
B_PREF_TRAIN_T	1.13	0.072	15.7	0	0.0744	15.2	0
B_PURPOSE_OTHER_TIME	0.0174	0.00858	2.03	0.0427	0.00892	1.95	0.0515
B_PURPOSE_VFR_TIME	0.0117	0.00855	1.36	0.173	0.00883	1.32	0.186
B_PURPOSE_WORK_TIME	-0.0749	0.0143	-5.23	1.73e-07	0.0146	-5.13	2.9e-07
B_TIME_C	-0.189	0.00791	-23.9	0	0.00782	-24.2	0
B_TIME_RISK_P	-0.00874	0.00323	-2.71	0.00671	0.0031	-2.82	0.00474
B_TIME_RISK_T	-0.00996	0.00263	-3.79	0.00015	0.00248	-4.02	5.73e-05
B_TIME_T	-0.112	0.0115	-9.69	0	0.0113	-9.89	0

Elimination 2

First all parameters with a p-value of 0.1 and 0.2 are deleted. This gives elimination_2. Two parameters have been deleted, so there are still 32. The following results were found. ρ^2 is 0.218. Final LL -7826.299.

Name	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
ASC_PLANE	-1.43	0.0947	-15.1	0	0.0911	-15.7	0
ASC_TRAIN	-0.353	0.0838	-4.22	2.48e-05	0.0843	-4.19	2.79e-05
B_4_7_PR	0.0194	0.00434	4.48	7.53e-06	0.00398	4.88	1.09e-06
B_AGE_COST_20_40	0.00282	0.0013	2.16	0.0304	0.00125	2.25	0.0242
B_AGE_COST_40_65	0.00574	0.0013	4.4	1.08e-05	0.00126	4.57	4.93e-06
B_AGE_COST_65_AND_OLDER	0.00644	0.0013	4.95	7.54e-07	0.00125	5.15	2.63e-07
B_COMFORT	0.335	0.0344	9.73	0	0.0329	10.2	0
B_COMPANY_COST_FAMILY	-0.00086	0.000486	-1.77	0.077	0.000489	-1.76	0.0787
B_COMPANY_COST_FRIENDS	-0.000743	0.000358	-2.08	0.0376	0.000353	-2.1	0.0355
B_COMPANY_COST_PARTNER	-0.000979	0.000476	-2.06	0.0398	0.00047	-2.08	0.0372
B_COMPANY_PR_FAMILY	-0.00495	0.00419	-1.18	0.238	0.004	-1.24	0.217
B_COMPANY_PR_FRIENDS	-0.0114	0.00339	-3.37	0.000762	0.00321	-3.55	0.000384
B_COMPANY_PR_OTHER	0.0167	0.00516	3.23	0.00122	0.00514	3.24	0.00119
B_COMPANY_PR_PARTNER	-0.0101	0.00442	-2.28	0.0227	0.00423	-2.38	0.0174
B_COST	-0.00832	0.00131	-6.32	2.56e-10	0.00126	-6.59	4.48e-11
B_EDU_HBO_PR	-0.0147	0.00337	-4.35	1.38e-05	0.00332	-4.41	1.03e-05
B_EDU_OTHER_PR	-0.0164	0.00376	-4.37	1.26e-05	0.00354	-4.63	3.62e-06
B_EDU_WO_PR	-0.0227	0.00357	-6.36	1.96e-10	0.00354	-6.41	1.43e-10
B_GENDER_PR	-0.0188	0.00242	-7.74	9.99e-15	0.00229	-8.21	2.22e-16
B_MORE_8_PR	0.0169	0.00642	2.64	0.00839	0.00601	2.82	0.00484
B_PAYMENT_WORKEDU_COST	0.00173	0.000734	2.35	0.0185	0.000719	2.4	0.0162
B_PREF_CAR_P	-1.67	0.107	-15.6	0	0.105	-16	0
B_PREF_CAR_T	-1.51	0.0949	-15.9	0	0.0957	-15.8	0
B_PREF_PLANE_P	1.1	0.0799	13.8	0	0.0783	14.1	0
B_PREF_TRAIN_P	-0.5	0.0889	-5.63	1.78e-08	0.0902	-5.55	2.89e-08
B_PREF_TRAIN_T	1.13	0.072	15.7	0	0.0745	15.2	0
B_PURPOSE_OTHER_TIME	0.0157	0.00848	1.85	0.0645	0.00883	1.78	0.0758
B_PURPOSE_WORK_TIME	-0.0768	0.0143	-5.38	7.41e-08	0.0146	-5.28	1.3e-07
B_TIME_C	-0.187	0.0078	-24	0	0.00769	-24.3	0
B_TIME_RISK_P	-0.00879	0.00322	-2.73	0.00638	0.00309	-2.84	0.00449
B_TIME_RISK_T	-0.00997	0.00263	-3.79	0.000148	0.00248	-4.02	5.71e-05
B_TIME_T	-0.11	0.0115	-9.6	0	0.0112	-9.81	0

Elimination 3

Then all parameters above 5% significance are removed. This gives elimination 3. Again two parameters have been removed, so there are 30 parameters. ρ^2 is 0.218. Final LL -7828.69.

Name	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
ASC_PLANE	-1.44	0.0947	-15.2	0	0.0912	-15.7	0
ASC_TRAIN	-0.355	0.0837	-4.23	2.29e-05	0.0843	-4.21	2.58e-05
B_4_7_PR	0.0194	0.00433	4.48	7.46e-06	0.00397	4.89	1.01e-06
B_AGE_COST_20_40	0.0028	0.0013	2.15	0.0318	0.00125	2.24	0.0248
B_AGE_COST_40_65	0.00571	0.0013	4.38	1.2e-05	0.00125	4.56	5.1e-06
B_AGE_COST_65_AND_OLDER	0.0064	0.0013	4.91	8.9e-07	0.00125	5.13	2.85e-07
B_COMFORT	0.335	0.0344	9.74	0	0.0329	10.2	0
B_COMPANY_COST_FAMILY	-0.000794	0.000485	-1.64	0.102	0.000488	-1.63	0.104
B_COMPANY_COST_FRIENDS	-0.000703	0.000357	-1.97	0.0488	0.000354	-1.99	0.0466
B_COMPANY_COST_PARTNER	-0.000943	0.000476	-1.98	0.0475	0.00047	-2.01	0.0448
B_COMPANY_PR_FRIENDS	-0.00928	0.0029	-3.2	0.00138	0.00281	-3.3	0.000953
B_COMPANY_PR_OTHER	0.0186	0.00487	3.82	0.000132	0.00496	3.76	0.000173
B_COMPANY_PR_PARTNER	-0.0078	0.00403	-1.93	0.053	0.00391	-2	0.046
B_COST	-0.00831	0.00131	-6.32	2.55e-10	0.00126	-6.61	3.79e-11
B_EDU_HBO_PR	-0.0146	0.00337	-4.34	1.41e-05	0.00333	-4.4	1.07e-05
B_EDU_OTHER_PR	-0.0167	0.00373	-4.48	7.4e-06	0.00351	-4.76	1.93e-06
B_EDU_WO_PR	-0.0226	0.00356	-6.35	2.17e-10	0.00354	-6.4	1.57e-10
B_GENDER_PR	-0.0202	0.00213	-9.47	0	0.00203	-9.91	0
B_MORE_8_PR	0.018	0.00634	2.83	0.00461	0.00597	3.01	0.00263
B_PAYMENT_WORKEDU_COST	0.00174	0.000734	2.38	0.0175	0.00072	2.42	0.0155
B_PREF_CAR_P	-1.67	0.107	-15.6	0	0.105	-16	0
B_PREF_CAR_T	-1.51	0.0949	-15.9	0	0.0957	-15.8	0
B_PREF_PLANE_P	1.1	0.0798	13.8	0	0.0782	14	0
B_PREF_TRAIN_P	-0.491	0.0887	-5.53	3.11e-08	0.0901	-5.45	5.09e-08
B_PREF_TRAIN_T	1.14	0.072	15.8	0	0.0744	15.2	0
B_PURPOSE_WORK_TIME	-0.0792	0.0142	-5.56	2.66e-08	0.0145	-5.46	4.9e-08
B_TIME_C	-0.185	0.00772	-24	0	0.00761	-24.4	0
B_TIME_RISK_P	-0.00954	0.00317	-3.01	0.00263	0.00305	-3.13	0.00174
B_TIME_RISK_T	-0.0106	0.00256	-4.15	3.33e-05	0.00243	-4.37	1.21e-05
B_TIME_T	-0.108	0.0114	-9.47	0	0.0112	-9.68	0

Elimination 4

Some parameters became insignificant on the 5% level, these parameters are removed. This gives elimination_4. One parameter is removed. So in total 29 parameters. ρ^2 is 0.218. Final LL -7830.035.

Name	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
ASC_PLANE	-1.44	0.0947	-15.2	0	0.0911	-15.8	0
ASC_TRAIN	-0.356	0.0837	-4.26	2.07e-05	0.0843	-4.23	2.34e-05
B_4_7_PR	0.0193	0.00433	4.47	7.99e-06	0.00397	4.88	1.08e-06
B_AGE_COST_20_40	0.00307	0.00129	2.38	0.0173	0.00123	2.49	0.0127
B_AGE_COST_40_65	0.0061	0.00128	4.77	1.88e-06	0.00122	4.98	6.37e-07
B_AGE_COST_65_AND_OLDER	0.00679	0.00128	5.31	1.12e-07	0.00122	5.56	2.64e-08
B_COMFORT	0.335	0.0344	9.75	0	0.0329	10.2	0
B_COMPANY_COST_FRIENDS	-0.00044	0.00032	-1.38	0.169	0.000318	-1.38	0.167
B_COMPANY_COST_PARTNER	-0.000659	0.000444	-1.48	0.138	0.000441	-1.49	0.135
B_COMPANY_PR_FRIENDS	-0.0093	0.0029	-3.21	0.00134	0.00281	-3.31	0.000925
B_COMPANY_PR_OTHER	0.0184	0.00486	3.78	0.000154	0.00494	3.72	0.000198
B_COMPANY_PR_PARTNER	-0.00783	0.00403	-1.94	0.0523	0.00391	-2	0.0453
B_COST	-0.00895	0.00125	-7.13	9.71e-13	0.00119	-7.53	5.17e-14
B_EDU_HBO_PR	-0.0146	0.00337	-4.34	1.41e-05	0.00333	-4.4	1.07e-05
B_EDU_OTHER_PR	-0.0167	0.00373	-4.48	7.6e-06	0.00351	-4.76	1.96e-06
B_EDU_WO_PR	-0.0227	0.00357	-6.36	2.08e-10	0.00354	-6.4	1.54e-10
B_GENDER_PR	-0.0202	0.00213	-9.46	0	0.00204	-9.9	0
B_MORE_8_PR	0.0179	0.00634	2.81	0.00488	0.00597	2.99	0.0028
B_PAYMENT_WORKEDU_COST	0.00195	0.000724	2.69	0.00709	0.000709	2.75	0.006
B_PREF_CAR_P	-1.67	0.107	-15.6	0	0.105	-16	0
B_PREF_CAR_T	-1.51	0.0949	-15.9	0	0.0956	-15.7	0
B_PREF_PLANE_P	1.1	0.0798	13.8	0	0.0782	14	0
B_PREF_TRAIN_P	-0.49	0.0886	-5.53	3.21e-08	0.0901	-5.44	5.25e-08
B_PREF_TRAIN_T	1.14	0.072	15.8	0	0.0744	15.3	0
B_PURPOSE_WORK_TIME	-0.0796	0.0142	-5.59	2.31e-08	0.0145	-5.48	4.37e-08
B_TIME_C	-0.185	0.00772	-24	0	0.00761	-24.4	0
B_TIME_RISK_P	-0.00955	0.00317	-3.01	0.0026	0.00305	-3.13	0.00172
B_TIME_RISK_T	-0.0106	0.00256	-4.14	3.51e-05	0.00243	-4.36	1.29e-05
B_TIME_T	-0.108	0.0114	-9.46	0	0.0112	-9.68	0

Elimination 5

Some parameters became insignificant on the 5% level, these parameters are removed. This gives elimination_5. Two parameters are removed. So in total 27 parameters. ρ^2 is 0.218. Final LL -7831.553. All parameters are significant on the 5% level.

Name	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
ASC_PLANE	-1.43	0.0946	-15.2	0	0.0911	-15.8	0
ASC_TRAIN	-0.357	0.0837	-4.26	2.01e-05	0.0842	-4.24	2.25e-05
B_4_7_PR	0.0193	0.00433	4.46	8.22e-06	0.00397	4.87	1.14e-06
B_AGE_COST_20_40	0.00295	0.00129	2.29	0.0221	0.00122	2.41	0.016
B_AGE_COST_40_65	0.0059	0.00127	4.63	3.61e-06	0.00121	4.88	1.09e-06
B_AGE_COST_65_AND_OLDER	0.00665	0.00127	5.24	1.64e-07	0.00121	5.52	3.47e-08
B_COMFORT	0.335	0.0344	9.75	0	0.0329	10.2	0
B_COMPANY_PR_FRIENDS	-0.00911	0.0029	-3.14	0.00166	0.00281	-3.24	0.00118
B_COMPANY_PR_OTHER	0.0183	0.00485	3.78	0.000159	0.00493	3.72	0.000198
B_COMPANY_PR_PARTNER	-0.00759	0.00404	-1.88	0.0602	0.00393	-1.93	0.0532
B_COST	-0.00909	0.00125	-7.26	3.81e-13	0.00119	-7.66	1.87e-14
B_EDU_HBO_PR	-0.0147	0.00337	-4.35	1.34e-05	0.00333	-4.41	1.02e-05
B_EDU_OTHER_PR	-0.0167	0.00373	-4.49	7.25e-06	0.00351	-4.77	1.86e-06
B_EDU_WO_PR	-0.0227	0.00357	-6.37	1.92e-10	0.00354	-6.41	1.43e-10
B_GENDER_PR	-0.0202	0.00213	-9.51	0	0.00203	-9.95	0
B_MORE_8_PR	0.0178	0.00634	2.81	0.00495	0.00596	2.99	0.0028
B_PAYMENT_WORKEDU_COST	0.00222	0.000707	3.14	0.00171	0.000693	3.2	0.00137
B_PREF_CAR_P	-1.67	0.107	-15.6	0	0.105	-16	0
B_PREF_CAR_T	-1.51	0.0949	-15.9	0	0.0957	-15.8	0
B_PREF_PLANE_P	1.1	0.0798	13.8	0	0.0782	14.1	0
B_PREF_TRAIN_P	-0.489	0.0886	-5.52	3.47e-08	0.0901	-5.43	5.72e-08
B_PREF_TRAIN_T	1.14	0.0719	15.8	0	0.0744	15.3	0
B_PURPOSE_WORK_TIME	-0.08	0.0143	-5.61	1.99e-08	0.0146	-5.5	3.86e-08
B_TIME_C	-0.185	0.00772	-24	0	0.0076	-24.4	0
B_TIME_RISK_P	-0.00957	0.00317	-3.02	0.00255	0.00305	-3.14	0.00169
B_TIME_RISK_T	-0.0107	0.00256	-4.16	3.15e-05	0.00243	-4.39	1.13e-05
B_TIME_T	-0.108	0.0114	-9.45	0	0.0112	-9.67	0

Final model 6

The preference for car is also added to the utility function of car. This gives in total three extra parameters. So as a consequence, there are now 30 parameter. ρ^2 increased to 0.24. Final LL -7605.543. However, some parameters turn now out to be insignificant.

Name	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
ASC_PLANE	-2.21	0.169	-13.1	0	0.17	-13	0
ASC_TRAIN	-0.368	0.147	-2.51	0.012	0.149	-2.47	0.0134
B_4_7_PR	0.0222	0.00437	5.07	4.06e-07	0.00398	5.56	2.69e-08
B_AGE_COST_20_40	0.00323	0.00132	2.45	0.0141	0.00125	2.59	0.00963
B_AGE_COST_40_65	0.00638	0.0013	4.9	9.65e-07	0.00124	5.16	2.41e-07
B_AGE_COST_65_AND_OLDER	0.00721	0.0013	5.55	2.81e-08	0.00123	5.86	4.6e-09
B_COMFORT	0.347	0.0348	9.94	0	0.0334	10.4	0
B_COMPANY_PR_FRIENDS	-0.00755	0.00293	-2.58	0.00994	0.00285	-2.65	0.00803
B_COMPANY_PR_OTHER	0.0196	0.00491	3.99	6.57e-05	0.00487	4.02	5.79e-05
B_COMPANY_PR_PARTNER	-0.00617	0.00408	-1.51	0.13	0.00399	-1.55	0.122
B_COST	-0.00971	0.00128	-7.58	3.46e-14	0.00121	-8.01	1.11e-15
B_EDU_HBO_PR	-0.0142	0.00339	-4.17	3.01e-05	0.00333	-4.25	2.1e-05
B_EDU_OTHER_PR	-0.0179	0.00381	-4.69	2.74e-06	0.00363	-4.93	8.27e-07
B_EDU_WO_PR	-0.0237	0.00359	-6.59	4.27e-11	0.00358	-6.61	3.75e-11
B_GENDER_PR	-0.0217	0.00215	-10.1	0	0.00207	-10.5	0
B_MORE_8_PR	0.0197	0.00637	3.1	0.00192	0.00598	3.3	0.000959
B_PAYMENT_WORKEDU_COST	0.00223	0.00072	3.1	0.00195	0.000702	3.18	0.00147
B_PREF_CAR_C	0.431	0.0436	9.87	0	0.0448	9.6	0
B_PREF_CAR_P	-0.0762	0.0501	-1.52	0.129	0.0501	-1.52	0.129
B_PREF_CAR_T	-0.354	0.0445	-7.97	1.55e-15	0.0451	-7.87	3.77e-15
B_PREF_PLANE_C	0.0304	0.0731	0.415	0.678	0.0752	0.404	0.686
B_PREF_PLANE_P	1.05	0.0791	13.3	0	0.0797	13.2	0
B_PREF_TRAIN_C	-0.246	0.0425	-5.8	6.79e-09	0.0439	-5.62	1.96e-08
B_PREF_TRAIN_P	-0.0739	0.0468	-1.58	0.114	0.0472	-1.57	0.117
B_PREF_TRAIN_T	0.32	0.0405	7.92	2.44e-15	0.0413	7.76	8.22e-15
B_PURPOSE_WORK_TIME	-0.0598	0.0145	-4.13	3.64e-05	0.015	-3.99	6.72e-05
B_TIME_C	-0.193	0.00789	-24.5	0	0.00781	-24.8	0
B_TIME_RISK_P	-0.011	0.00323	-3.39	0.000695	0.0031	-3.53	0.000417
B_TIME_RISK_T	-0.0107	0.0026	-4.12	3.73e-05	0.00247	-4.33	1.48e-05
B_TIME_T	-0.114	0.0116	-9.87	0	0.0114	-10	0

Final model

The parameters 4_7 and 8_MORE are not significantly different from each other, as both values are very similar and the standard error captures the value of the other parameter. Therefore these variables will be combined. Some parameters were not significant now, and are removed. There are 27 parameters, and this will be the final model.

Name	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
ASC_PLANE	-2.24	0.158	-14.1	0	0.158	-14.1	0
ASC_TRAIN	-0.417	0.0973	-4.29	1.82e-05	0.0995	-4.19	2.74e-05
B_4_OR_MORE_PR	0.022	0.00372	5.92	3.18e-09	0.00343	6.43	1.26e-10
B_AGE_COST_20_40	0.00328	0.00132	2.49	0.0129	0.00125	2.62	0.00883
B_AGE_COST_40_65	0.00643	0.0013	4.93	8.16e-07	0.00124	5.19	2.1e-07
B_AGE_COST_65_AND_OLDER	0.00726	0.0013	5.59	2.25e-08	0.00123	5.89	3.87e-09
B_COMFORT	0.347	0.0348	9.95	0	0.0333	10.4	0
B_COMPANY_PR_FRIENDS	-0.00586	0.00273	-2.15	0.0317	0.00267	-2.19	0.0285
B_COMPANY_PR_OTHER	0.0214	0.00477	4.49	7.05e-06	0.00476	4.5	6.92e-06
B_COST	-0.00976	0.00128	-7.62	2.64e-14	0.00122	-8.03	8.88e-16
B_EDU_HBO_PR	-0.0139	0.00338	-4.11	4e-05	0.00333	-4.18	2.96e-05
B_EDU_OTHER_PR	-0.0179	0.00379	-4.73	2.25e-06	0.00359	-5	5.75e-07
B_EDU_WO_PR	-0.0238	0.00358	-6.65	2.93e-11	0.00356	-6.69	2.26e-11
B_GENDER_PR	-0.0229	0.002	-11.4	0	0.00194	-11.8	0
B_PAYMENT_WORKEDU_COST	0.00222	0.000721	3.08	0.00204	0.000702	3.16	0.00155
B_PREF_CAR_C	0.498	0.0783	6.37	1.93e-10	0.0786	6.34	2.35e-10
B_PREF_CAR_T	-0.267	0.0805	-3.32	0.000913	0.0806	-3.31	0.000917
B_PREF_PLANE_P	1.04	0.0713	14.6	0	0.0719	14.5	0
B_PREF_TRAIN_C	-0.185	0.0749	-2.47	0.0137	0.0758	-2.44	0.0148
B_PREF_TRAIN_T	0.405	0.0729	5.56	2.74e-08	0.0736	5.51	3.58e-08
B_PURPOSE_WORK_TIME	-0.0599	0.0145	-4.14	3.55e-05	0.015	-3.99	6.54e-05
B_TIME_C	-0.194	0.00789	-24.6	0	0.00781	-24.8	0
B_TIME_RISK_P	-0.0116	0.0032	-3.63	0.00028	0.00307	-3.79	0.000153
B_TIME_RISK_T	-0.0112	0.00257	-4.37	1.27e-05	0.00246	-4.57	4.99e-06
B_TIME_T	-0.114	0.0116	-9.87	0	0.0114	-10	0

Now all parameters are significant on the 5% level (most of them even at the 1% level). ρ^2 is 0.24. Final LL -7606.821. **This will be the final model.**

D

PandasBiogeme code for Python & estimation results

Base MNL model only main variables

```
import pandas as pd
import biogeme.database as db
import biogeme.biogeme as bio
from biogeme import models
from biogeme.expressions import Beta

# Read the data
df = pd.read_csv(’C:/Users/mauri/OneDrive/Documenten/TIL5060 TIL thesis/Data
survey/Volledige_dataset_download/Main_choice_gecombineerd/Volledige_data
_gecombineerd_omgezet_alleen_main.dat’, sep=’\t’)
database = db.Database(’main_choice_volledig’, df)

# The following statement allows you to use the names of the
# variable as Python variable.
globals().update(database.variables)

# Parameters to be estimated
ASC_CAR = Beta(’ASC_CAR’,0,-1000,1000,1)
ASC_TRAIN = Beta(’ASC_TRAIN’,0,-1000,1000,0)
ASC_PLANE = Beta(’ASC_PLANE’,0,-1000,1000,0)
B_TIME = Beta(’B_TIME’,0,-1000,1000,0)
B_COMFORT = Beta(’B_COMFORT’,0,-1000,1000,0)
B_RISK = Beta(’B_RISK’,0,-1000,1000,0)
B_COST = Beta(’B_COST’,0,-1000,1000,0)

# Definition of the utility functions
V1 = ASC_TRAIN \
+ B_TIME * TT_T\
+ B_COST * TC_T\
+ B_COMFORT * CF_T\
+ B_RISK * PR_T
V2 = ASC_PLANE\
+ B_TIME * TT_P\
+ B_COST * TC_P\
+ B_COMFORT * CF_P\
+ B_RISK * PR_P
```

```
V3 = ASC_CAR\
+ B_TIME * TT_C + B_COST * TC_C

# Associate utility functions with the numbering of alternatives
V = {1: V1, 2: V2, 3: V3}

AV1 = 1
AV2 = 1
AV3 = 1

Associate the availability conditions with the alternatives
av =
{1: AV1,
2: AV2,
3: AV3}

# Definition of the model. This is the contribution of each
# observation to the log likelihood function.
logprob = models.loglogit(V, av, CHOICE)

# Create the Biogeme object
biogeme = bio.BIOGEME(database, logprob)
biogeme.modelName = 'Basis_model'

# Calculate the null log likelihood for reporting.
biogeme.calculateNullLoglikelihood(av)

# Estimate the parameters
results = biogeme.estimate()

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

Estimation report of base model with only main variables

Estimation report

Number of estimated parameters: 6
Sample size: 9112
Excluded observations: 0
Null log likelihood: -10010.56
Init log likelihood: -10010.56
Final log likelihood: -8745.118
Likelihood ratio test for the null model: 2530.875
Rho-square for the null model: 0.126
Rho-square-bar for the null model: 0.126
Likelihood ratio test for the init. model: 2530.875
Rho-square for the init. model: 0.126
Rho-square-bar for the init. model: 0.126
Akaike Information Criterion: 17502.24
Bayesian Information Criterion: 17544.94
Final gradient norm: 1.5920E-03
Nbr of threads: 12
Algorithm: Newton with trust region for simple bound constraints
Proportion analytical hessian: 100.0%
Relative projected gradient: 1.380441e-07
Relative change: 2.0528238262908227e-09
Number of iterations: 4
Number of function evaluations: 13
Number of gradient evaluations: 5
Number of hessian evaluations: 5
Cause of termination: Relative change = 2.05e-09 <= 1e-05
Optimization time: 0:00:00.331703

Estimated parameters

Name	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
ASC_PLANE	-0.359	0.0703	-5.1	3.38e-07	0.0677	-5.3	1.16e-07
ASC_TRAIN	0.35	0.0558	6.28	3.49e-10	0.0547	6.4	1.53e-10
B_COMFORT	0.281	0.0332	8.46	0	0.0323	8.7	0
B_COST	-0.00306	0.000145	-21.1	0	0.000144	-21.2	0
B_RISK	-0.22	0.0103	-21.4	0	0.01	-22	0
B_TIME	-0.167	0.00642	-26	0	0.00629	-26.5	0

Correlation of coefficients

Coefficient1	Coefficient2	Covariance	Correlation	t-test	p-value	Rob. cov.	Rob. corr.	Rob. t-test	Rob. p-value
ASC_TRAIN	ASC_PLANE	0.00349	0.889	21.5	0	0.00327	0.884	22.1	0
B_COMFORT	ASC_PLANE	-0.00159	-0.681	6.66	2.79e-11	-0.00148	-0.68	6.9	5.25e-12
B_COMFORT	ASC_TRAIN	-0.00122	-0.662	-0.85	0.395	-0.00114	-0.647	-0.873	0.383
B_COST	ASC_PLANE	-2.27e-06	-0.223	5.05	4.31e-07	-1.97e-06	-0.201	5.25	1.51e-07
B_COST	ASC_TRAIN	-2.18e-06	-0.27	-6.33	2.52e-10	-1.93e-06	-0.244	-6.45	1.09e-10
B_COST	B_COMFORT	8.45e-07	0.176	-8.56	0	6.59e-07	0.142	-8.8	0
B_RISK	ASC_PLANE	-0.000505	-0.697	1.78	0.0745	-0.000466	-0.689	1.85	0.0639
B_RISK	ASC_TRAIN	-0.000425	-0.74	-8.94	0	-0.000404	-0.739	-9.13	0
B_RISK	B_COMFORT	0.000158	0.463	-16.8	0	0.000134	0.416	-17	0
B_RISK	B_COST	3.46e-07	0.232	-21.1	0	3.23e-07	0.224	-21.7	0
B_TIME	ASC_PLANE	0.000326	0.723	2.91	0.00357	0.000299	0.702	3.02	0.0025
B_TIME	ASC_TRAIN	0.000218	0.609	-9.92	0	0.000201	0.584	-10.1	0
B_TIME	B_COMFORT	-8.26e-05	-0.388	-12.4	0	-7.52e-05	-0.371	-12.8	0
B_TIME	B_COST	-3.95e-09	-0.00425	-25.5	0	-3.84e-09	-0.00423	-26	0
B_TIME	B_RISK	-2.31e-05	-0.35	3.81	0.00014	-2.22e-05	-0.353	3.9	9.51e-05

Base MNL model with main interactions

import pandas as pd
import biogeme.database as db
import biogeme.biogeme as bio
from biogeme **import** models
from biogeme.expressions **import** Beta

# Read the data
df = pd.read_csv('C:/Users/mauri/OneDrive/Documenten/TIL5060 TIL thesis/Data survey/Volledige_dataset
_download/Main_choice_gecombineerd/Modellen_biogeme/Volledige_data_gecombineerd_omgezet_met_socio
_new.dat', sep='\ t')
$database = db.Database('main\_choice\_volledig', df)$

# The following statement allows you to use the names of the
# variable as Python variable.
globals().update(database.variables)

# Parameters to be estimated
ASC_CAR = Beta('ASC_CAR',0,-1000,1000,1)
ASC_TRAIN = Beta('ASC_TRAIN',0,-1000,1000,0)
ASC_PLANE = Beta('ASC_PLANE',0,-1000,1000,0)
B_TIME_T = Beta('B_TIME_T',0,-1000,1000,0)
B_TIME_C = Beta('B_TIME_C',0,-1000,1000,0)
B_TIME_P = Beta('B_TIME_P',0,-1000,1000,0)
B_COMFORT = Beta('B_COMFORT',0,-1000,1000,0)
B_COST = Beta('B_COST',0,-1000,1000,0)

# Interaction of main variables
B_TIME_RISK_T = Beta('B_TIME_RISK_T',0,-1000,1000,0)
B_TIME_RISK_P = Beta('B_TIME_RISK_P',0,-1000,1000,0)

Definition of the utility functions
V1 = ASC_TRAIN \
$+ B\_COMFORT * CF\_T\backslash$
$+ (B\_TIME\_T + B\_TIME\_RISK\_T * (PR\_T - 1)) * TT\_T\backslash$
$+ B\_COST * TC\_T$
$V2 = ASC\_PLANE\backslash$
$+ B\_COMFORT * CF\_P\backslash$
$+ (B\_TIME\_RISK\_P * (PR\_P - 1)) * TT\_P\backslash$
$+ B\_COST * TC\_P$
$V3 = ASC\_CAR\backslash$
$+ B\_TIME\_C * TT\_C\backslash$
$+ B\_COST * TC\_C$

# Associate utility functions with the numbering of alternatives
V = 1: V1, 2: V2, 3: V3

AV1 = 1
AV2 = 1
AV3 = 1

Associate the availability conditions with the alternatives
av =
{1: AV1,
2: AV2,
3: AV3}

```
# Definition of the model. This is the contribution of each
# observation to the log likelihood function.
logprob = models.loglogit(V, av, CHOICE)

# Create the Biogeme object
biogeme = bio.BIOGEME(database, logprob)
biogeme.modelName = 'Final_basis_model'

# Calculate the null log likelihood for reporting.
biogeme.calculateNullLoglikelihood(av)

# Estimate the parameters
results = biogeme.estimate()

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

Estimation report of base model with main interactions

Estimation report

Number of estimated parameters:	8
Sample size:	9112
Excluded observations:	0
Null log likelihood:	-10010.56
Init log likelihood:	-10010.56
Final log likelihood:	-8683.222
Likelihood ratio test for the null model:	2654.667
Rho-square for the null model:	0.133
Rho-square-bar for the null model:	0.132
Likelihood ratio test for the init. model:	2654.667
Rho-square for the init. model:	0.133
Rho-square-bar for the init. model:	0.132
Akaike Information Criterion:	17382.44
Bayesian Information Criterion:	17439.38
Final gradient norm:	3.7728E-02
Nbr of threads:	12
Algorithm:	Newton with trust region for simple bound constraints
Proportion analytical hessian:	100.0%
Relative projected gradient:	2.927481e-06
Relative change:	2.116480836943102e-09
Number of iterations:	4
Number of function evaluations:	13
Number of gradient evaluations:	5
Number of hessian evaluations:	5
Cause of termination:	Relative change = 2.12e-09 <= 1e-05
Optimization time:	0:00:00.251012

Estimated parameters

Name	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
ASC_PLANE	-1.39	0.0875	-15.9	0	0.0849	-16.4	0
ASC_TRAIN	-0.206	0.0777	-2.66	0.00793	0.0779	-2.65	0.0081
B_COMFORT	0.296	0.0323	9.17	0	0.031	9.53	0
B_COST	-0.00323	0.000144	-22.5	0	0.000141	-23	0
B_TIME_C	-0.169	0.00727	-23.3	0	0.00719	-23.5	0
B_TIME_RISK_P	-0.0408	0.00379	-10.8	0	0.00365	-11.2	0
B_TIME_RISK_T	-0.0394	0.00233	-16.9	0	0.00227	-17.3	0
B_TIME_T	-0.0998	0.0107	-9.31	0	0.0105	-9.55	0

Correlation of coefficients

Coefficient1	Coefficient2	Covariance	Correlation	t-test	p-value	Rob. cov.	Rob. corr.	Rob. t-test	Rob. p-value
ASC_TRAIN	ASC_PLANE	0.00406	0.597	15.9	0	0.00405	0.612	16.5	0
B_COMFORT	ASC_PLANE	-0.00123	-0.435	16	0	-0.0011	-0.417	16.6	0
B_COMFORT	ASC_TRAIN	-0.000165	-0.0657	5.83	5.47e-09	-0.000223	-0.0921	5.8	6.46e-09
B_COST	ASC_PLANE	-1.49e-06	-0.118	15.9	0	-1.24e-06	-0.104	16.4	0
B_COST	ASC_TRAIN	-8.57e-07	-0.0766	2.61	0.00897	-6.64e-07	-0.0604	2.61	0.00916
B_COST	B_COMFORT	7.5e-07	0.161	-9.27	0	5.5e-07	0.126	-9.64	0
B_TIME_C	ASC_PLANE	0.000519	0.815	15	0	0.000502	0.822	15.5	0
B_TIME_C	ASC_TRAIN	0.000385	0.681	0.51	0.61	0.000371	0.663	0.507	0.612
B_TIME_C	B_COMFORT	-3.06e-05	-0.131	-13.7	0	-2.86e-05	-0.128	-14.2	0
B_TIME_C	B_COST	4.27e-08	0.0408	-22.8	0	3.88e-08	0.0384	-23.1	0
B_TIME_RISK_P	ASC_PLANE	-0.000132	-0.396	15.2	0	-0.000126	-0.406	15.6	0
B_TIME_RISK_P	ASC_TRAIN	1.32e-06	0.00447	2.13	0.0333	-3.55e-06	-0.0125	2.12	0.0339
B_TIME_RISK_P	B_COMFORT	2.57e-05	0.21	-10.6	0	1.64e-05	0.145	-11	0
B_TIME_RISK_P	B_COST	7.28e-08	0.133	-9.94	0	6.8e-08	0.132	-10.3	0
B_TIME_RISK_P	B_TIME_C	3.48e-08	0.00126	15.7	0	-1.15e-06	-0.0438	15.7	0
B_TIME_RISK_T	ASC_PLANE	-1.89e-05	-0.0925	15.4	0	-1.39e-05	-0.0722	15.9	0
B_TIME_RISK_T	ASC_TRAIN	-5.47e-06	-0.0302	2.15	0.0319	-2.36e-06	-0.0133	2.14	0.0323
B_TIME_RISK_T	B_COMFORT	1.76e-05	0.235	-10.5	0	1.67e-05	0.237	-11	0
B_TIME_RISK_T	B_COST	4.27e-08	0.128	-15.6	0	2.94e-08	0.0918	-16	0
B_TIME_RISK_T	B_TIME_C	-2.86e-06	-0.169	16.2	0	-2.28e-06	-0.139	16.6	0
B_TIME_RISK_T	B_TIME_RISK_P	-2.85e-06	-0.323	0.271	0.787	-2.57e-06	-0.31	0.281	0.778
B_TIME_T	ASC_PLANE	0.000298	0.318	15.3	0	0.000245	0.276	15.7	0
B_TIME_T	ASC_TRAIN	-0.00026	-0.312	1.3	0.192	-0.000281	-0.345	1.3	0.195
B_TIME_T	B_COMFORT	-0.000135	-0.389	-10.5	0	-0.000114	-0.353	-11	0
B_TIME_T	B_COST	-7.78e-08	-0.0504	-9	0	-5.14e-08	-0.0349	-9.23	0
B_TIME_T	B_TIME_C	2.32e-05	0.297	6.29	3.15e-10	2.12e-05	0.282	6.37	1.9e-10
B_TIME_T	B_TIME_RISK_P	4.28e-06	0.105	-5.37	7.66e-08	3.55e-06	0.0932	-5.5	3.89e-08
B_TIME_T	B_TIME_RISK_T	-1.51e-05	-0.607	-4.92	8.51e-07	-1.41e-05	-0.593	-5.06	4.2e-07

Final model

import pandas as pd
import biogeme.database as db
import biogeme.biogeme as bio
from biogeme import models
from biogeme.expressions import Beta

Read the data
df = pd.read_csv('C:/Users/mauri/OneDrive/Documenten/TIL5060 TIL thesis/Data survey/Volledige_dataset_download/Main_choice_gecombineerd/Modellen_biogeme/Volledige_data_gecombineerd_omgezet_met_socio_new.dat',
sep='\ $t')database = db.Database('main\_choice\_volledig', df)$
The following statement allows you to use the names of the
variable as Python variable.
globals().update(database.variables)

Parameters to be estimated
ASC_CAR = Beta('ASC_CAR',0,-1000,1000,1)
ASC_TRAIN = Beta('ASC_TRAIN',0,-1000,1000,0)
ASC_PLANE = Beta('ASC_PLANE',0,-1000,1000,0)
B_TIME_T = Beta('B_TIME_T',0,-1000,1000,0)
B_TIME_C = Beta('B_TIME_C',0,-1000,1000,0)
B_TIME_P = Beta('B_TIME_P',0,-1000,1000,0)
B_COMFORT = Beta('B_COMFORT',0,-1000,1000,0)
B_COST = Beta('B_COST',0,-1000,1000,0)

Interactions
B_TIME_RISK_T = Beta('B_TIME_RISK_T',0,-1000,1000,0)
B_TIME_RISK_P = Beta('B_TIME_RISK_P',0,-1000,1000,0)

Socio-demographic interactions
B_AGE_COST_20_40 = Beta('B_AGE_COST_20_40',0,-1000,1000,0)
B_AGE_COST_40_65 = Beta('B_AGE_COST_40_65',0,-1000,1000,0)
B_AGE_COST_65_AND_OLDER = Beta('B_AGE_COST_65_AND_OLDER',0,-1000,1000,0)

B_GENDER_PR = Beta('B_GENDER_PR',0,-1000,1000,0)

B_EDU_HBO_PR = Beta('B_EDU_HBO_PR',0,-1000,1000,0)
B_EDU_WO_PR = Beta('B_EDU_WO_PR',0,-1000,1000,0)
B_EDU_OTHER_PR = Beta('B_EDU_OTHER_PR',0,-1000,1000,0)

Travel behaviour interactions
B_PAYMENT_WORKEDU_COST = Beta('B_PAYMENT_WORKEDU_COST',0,-1000,1000,0)

B_COMPANY_PR_FRIENDS = Beta('B_COMPANY_PR_FRIENDS',0,-1000,1000,0)
B_COMPANY_PR_PARTNER = Beta('B_COMPANY_PR_PARTNER',0,-1000,1000,0)
B_COMPANY_PR_OTHER = Beta('B_COMPANY_PR_OTHER',0,-1000,1000,0)

B_PURPOSE_WORK_TIME = Beta('B_PURPOSE_WORK_TIME',0,-1000,1000,0)
B_PURPOSE_OTHER_TIME = Beta('B_PURPOSE_OTHER_TIME',0,-1000,1000,0)

B_4_OR_MORE_PR = Beta('B_4_OR_MORE_PR',0,-1000,1000,0)

B_PREF_TRAIN_T = Beta('B_PREF_TRAIN_T',0,-1000,1000,0)
B_PREF_CAR_T = Beta('B_PREF_CAR_T',0,-1000,1000,0)

```
B_PREF_TRAIN_P = Beta('B_PREF_TRAIN_P',0,-1000,1000,1)
B_PREF_PLANE_P = Beta('B_PREF_PLANE_P',0,-1000,1000,0)
B_PREF_CAR_P = Beta('B_PREF_CAR_P',0,-1000,1000,1)

B_PREF_TRAIN_C = Beta('B_PREF_TRAIN_C',0,-1000,1000,0)
B_PREF_PLANE_C = Beta('B_PREF_PLANE_C',0,-1000,1000,1)
B_PREF_CAR_C = Beta('B_PREF_CAR_C',0,-1000,1000,0)

Definition of the utility functions
V1 = ASC_TRAIN + (B_PREF_TRAIN_T * (TRAIN_400_600 + TRAIN_800_1200) + B_PREF_CAR_T * (CAR_400_600
+ CAR_800_1200))
+ B_COMFORT * CF_T
+ ((B_TIME_T + (B_PURPOSE_WORK_TIME * WORK)) + (B_TIME_RISK_T + B_GENDER_PR + (B_EDU_HBO_PR
* HBO + B_EDU_WO_PR * WO + B_EDU_OTHER_PR * EDU_OTHER) + (B_COMPANY_PR_FRIENDS * WITH_FRIENDS
+ B_COMPANY_PR_PARTNER * WITH_PARTNER + B_COMPANY_PR_OTHER * WITH_OTHER) + (B_4_OR_MORE_PR
* (TRAVEL_4_7 + M8_OR_MORE))) * (PR_T - 1)) * TT_T
+ (B_COST + (B_AGE_COST_20_40 * AGE_20_40 + B_AGE_COST_40_65 * AGE_40_65 + B_AGE_COST_65_AND_OLDER
* AGE_65_AND_OLDER) + (B_PAYMENT_WORKEDU_COST * WORK_EDU)) * TC_T

V2 = ASC_PLANE + (B_PREF_TRAIN_P * (TRAIN_400_600 + TRAIN_800_1200) + B_PREF_PLANE_P * (PLANE_400_600
+ PLANE_800_1200) + B_PREF_CAR_P * (CAR_400_600 + CAR_800_1200)) + B_COMFORT * CF_P + ((B_TIME_RISK_P
+ B_GENDER_PR + (B_EDU_HBO_PR * HBO + B_EDU_WO_PR * WO + B_EDU_OTHER_PR * EDU_OTHER) +
(B_COMPANY_PR_FRIENDS * WITH_FRIENDS + B_COMPANY_PR_PARTNER * WITH_PARTNER + B_COMPANY_PR_OTHER
* WITH_OTHER) + (B_4_OR_MORE_PR * (TRAVEL_4_7 + M8_OR_MORE))) * (PR_P - 1)) * TT_P
+ (B_COST + (B_AGE_COST_20_40 * AGE_20_40 + B_AGE_COST_40_65 * AGE_40_65 + B_AGE_COST_65_AND_OLDER
* AGE_65_AND_OLDER) + (B_PAYMENT_WORKEDU_COST * WORK_EDU)) * TC_P

V3 = ASC_CAR + (B_PREF_TRAIN_C * (TRAIN_400_600 + TRAIN_800_1200) + B_PREF_PLANE_C * (PLANE_400_600
+ PLANE_800_1200) + B_PREF_CAR_C * (CAR_400_600 + CAR_800_1200)) + (B_TIME_C + (B_PURPOSE_WORK_TIME
* WORK)) * TT_C
+ (B_COST + (B_AGE_COST_20_40 * AGE_20_40 + B_AGE_COST_40_65 * AGE_40_65 + B_AGE_COST_65_AND_OLDER
* AGE_65_AND_OLDER) + (B_PAYMENT_WORKEDU_COST * WORK_EDU)) * TC_C
Associate utility functions with the numbering of alternatives
V = {1: V1, 2: V2, 3: V3}

AV1 = 1
AV2 = 1
AV3 = 1

# Associate the availability conditions with the alternatives
av =
{1: AV1,
2: AV2,
3: AV3}

# Definition of the model. This is the contribution of each
# observation to the log likelihood function.
logprob = models.loglogit(V, av, CHOICE)

# Create the Biogeme object
biogeme = bio.BIOGEME(database, logprob)
biogeme.modelName = 'Final_basis_model'

# Calculate the null log likelihood for reporting.
biogeme.calculateNullLoglikelihood(av)
```

```
# Estimate the parameters
results = biogeme.estimate()

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

Estimation report of final model

Report file: Final_model~00.html
Database name: main_choice_volledig

Estimation report

Number of estimated parameters: 25
Sample size: 9112
Excluded observations: 0
Null log likelihood: -10010.56
Init log likelihood: -10010.56
Final log likelihood: -7606.821
Likelihood ratio test for the null model: 4807.468
Rho-square for the null model: 0.24
Rho-square-bar for the null model: 0.238
Likelihood ratio test for the init. model: 4807.468
Rho-square for the init. model: 0.24
Rho-square-bar for the init. model: 0.238
Akaike Information Criterion: 15263.64
Bayesian Information Criterion: 15441.58
Final gradient norm: 1.7496E-02
Nbr of threads: 12
Algorithm: Newton with trust region for simple bound constraints
Proportion analytical hessian: 100.0%
Relative projected gradient: 9.821563e-07
Relative change: 0.0010595381964233441
Number of iterations: 4
Number of function evaluations: 13
Number of gradient evaluations: 5
Number of hessian evaluations: 5
Cause of termination: Relative gradient = 9.8e-07 <= 6.1e-06
Optimization time: 0:00:07.825193

Estimated parameters

Name	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
ASC_PLANE	-2.24	0.158	-14.1	0	0.158	-14.1	0
ASC_TRAIN	-0.417	0.0973	-4.29	1.82e-05	0.0995	-4.19	2.74e-05
B_4_OR_MORE_PR	0.022	0.00372	5.92	3.18e-09	0.00343	6.43	1.26e-10
B_AGE_COST_20_40	0.00328	0.00132	2.49	0.0129	0.00125	2.62	0.00883
B_AGE_COST_40_65	0.00643	0.0013	4.93	8.16e-07	0.00124	5.19	2.1e-07
B_AGE_COST_65_AND_OLDER	0.00726	0.0013	5.59	2.25e-08	0.00123	5.89	3.87e-09
B_COMFORT	0.347	0.0348	9.95	0	0.0333	10.4	0
B_COMPANY_PR_FRIENDS	-0.00586	0.00273	-2.15	0.0317	0.00267	-2.19	0.0285
B_COMPANY_PR_OTHER	0.0214	0.00477	4.49	7.05e-06	0.00476	4.5	6.92e-06
B_COST	-0.00976	0.00128	-7.62	2.64e-14	0.00122	-8.03	8.88e-16
B_EDU_HBO_PR	-0.0139	0.00338	-4.11	4e-05	0.00333	-4.18	2.96e-05
B_EDU_OTHER_PR	-0.0179	0.00379	-4.73	2.25e-06	0.00359	-5	5.75e-07
B_EDU_WO_PR	-0.0238	0.00358	-6.65	2.93e-11	0.00356	-6.69	2.26e-11
B_GENDER_PR	-0.0229	0.002	-11.4	0	0.00194	-11.8	0
B_PAYMENT_WORKEDU_COST	0.00222	0.000721	3.08	0.00204	0.000702	3.16	0.00155
B_PREF_CAR_C	0.498	0.0783	6.37	1.93e-10	0.0786	6.34	2.35e-10
B_PREF_CAR_T	-0.267	0.0805	-3.32	0.000913	0.0806	-3.31	0.000917
B_PREF_PLANE_P	1.04	0.0713	14.6	0	0.0719	14.5	0
B_PREF_TRAIN_C	-0.185	0.0749	-2.47	0.0137	0.0758	-2.44	0.0148
B_PREF_TRAIN_T	0.405	0.0729	5.56	2.74e-08	0.0736	5.51	3.58e-08
B_PURPOSE_WORK_TIME	-0.0599	0.0145	-4.14	3.55e-05	0.015	-3.99	6.54e-05
B_TIME_C	-0.194	0.00789	-24.6	0	0.00781	-24.8	0
B_TIME_RISK_P	-0.0116	0.0032	-3.63	0.00028	0.00307	-3.79	0.000153
B_TIME_RISK_T	-0.0112	0.00257	-4.37	1.27e-05	0.00246	-4.57	4.99e-06
B_TIME_T	-0.114	0.0116	-9.87	0	0.0114	-10	0

E

Survey

This appendix shows the survey as made in the MWM2 (Crowdtech) software. Note that for the rating, main choice 400-600km and main choice 800-1200km only *block 1* is used. The full design can be found in the Ngene design in appendix B.

1	**De invloed van COVID-19 risico op de keuze van vervoerswijze voor langeafstandsvervoer binnen Europa**	*Tussenpagina*

TUDelft NS International Royal HaskoningDHV Enhancing Society Together

2	**Om deel te nemen aan deze enquête moet u akkoord gaan met de voorwaarden.** Als u niet akkoord gaat, dan wordt u direct naar het einde van de enquête geleid.	*Tabelvraag (single response)*

	Akkoord	Niet akkoord
Door akkoord te gaan neem ik vrijwillig deel aan deze enquête.	○	○
Ik ben me ervan bewust dat de informatie die ik zal verstrekken door het invullen van deze enquête alleen voor onderzoeksdoeleinden zal worden gebruikt. Mogelijk worden bevindingen van dit onderzoek verspreid door middel van wetenschappelijke publicaties.	○	○
Ik ben me ervan bewust dat deelname te allen tijde anoniem zal zijn.	○	○

3	**Deel 1: Reisgedrag** U krijgt eerst een paar vragen over uw reisgedrag voorgelegd.	*Tussenpagina*

4	**Hoeveel keer per jaar maakte u een reis vanuit Nederland naar Europese bestemmingen in 2021?** Dit kan met alle vervoerswijzen zijn (bijvoorbeeld auto, vliegtuig, trein, bus et cetera).	*Single-responsevraag*

- ◯ Ik heb in het afgelopen jaar niet gereisd naar Europese bestemmingen
- ◯ 1-3 keer per jaar
- ◯ 4-7 keer per jaar
- ◯ 8-10 keer per jaar
- ◯ Meer dan 10 keer per jaar

5	**Wat was uw voornaamste reden om te reizen vanuit Nederland naar Europese bestemmingen in 2021?** Als u niet heeft gereisd in 2021, kies dan het antwoord dat u zou verwachten als u wel zou gaan reizen.	*Single-responsevraag*

- ◯ Werk
- ◯ School, studie of opleiding
- ◯ Vrije tijd (vakantie/voor plezier)
- ◯ Bezoek van familie of vrienden
- ◯ Anders, namelijk

6	**U heeft gereisd met de reden zoals ingevuld door u in de vorige vraag. Wie heeft betaald voor de reis?** Als u niet heeft gereisd in 2021, kies dan het antwoord dat u zou verwachten als u wel zou gaan reizen.	*Single-responsevraag*

- ◯ Ik betaal zelf voor de reis
- ◯ Een ander persoon betaalt
- ◯ Werkgever, onderwijsinstelling of opdrachtgever betaalt
- ◯ Anders, namelijk

7	**Met welk reisgezelschap reist u meestal vanuit Nederland naar Europese bestemmingen?** Als u niet heeft gereisd in 2021, kies dan het antwoord dat u zou verwachten als u wel zou gaan reizen.	*Single-responsevraag*

- ◯ Alleen
- ◯ Met partner
- ◯ Met gezin
- ◯ Met vriend/vrienden
- ◯ Met familie
- ◯ Anders, namelijk

8	**Welke vervoerswijze heeft uw voorkeur als u reist vanuit Nederland naar Europese bestemmingen in de afstandsklasse 400-600 kilometer?** Dit zijn bestemmingen zoals Parijs, Londen, Zürich, Berlijn en Kopenhagen.	*Single-responsevraag*

- ◯ Trein
- ◯ Auto
- ◯ Vliegtuig
- ◯ Bus
- ◯ Anders, namelijk

9	**Welke vervoerswijze heeft uw voorkeur als u reist vanuit Nederland naar Europese bestemmingen in de afstandsklasse 800-1200 kilometer?** Dit zijn bestemmingen zoals Bordeaux, Milaan, Barcelona, Warschau en Stockholm.	*Single-responsevraag*

- ◯ Trein
- ◯ Auto
- ◯ Vliegtuig
- ◯ Bus
- ◯ Anders, namelijk

10	**Heeft u altijd beschikking tot een auto?**	*Single-responsevraag*

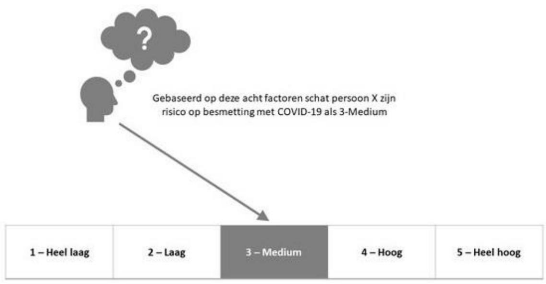
- ◯ Ja, ik heb een eigen auto die ik niet deel met anderen
- ◯ Nee, ik gebruik de auto in overleg met mensen in mijn huishouden
- ◯ Nee, ik gebruik de auto in overleg met mensen buiten mijn huishouden
- ◯ Nee, ik maak wel eens gebruik van een deel-/huurauto
- ◯ Nee, ik heb helemaal geen beschikking tot een auto

11	**Maakt u zich meer of minder zorgen om de omikron-variant van het COVID-19 virus in vergelijking met de delta-variant?**	*Single-responsevraag*

- ◯ Ik maak me er minder zorgen om
- ◯ Ik maak me er meer zorgen om
- ◯ Het is hetzelfde gebleven

12	**Deel 2: Inschatting kans op besmetting met COVID-19** We zijn benieuwd hoe groot u de kans op een besmetting met COVID-19 inschat als u met het vliegtuig of de trein reist naar een Europese bestemming. Dit is natuurlijk afhankelijk van verschillende factoren. Denk aan de bezetting aan boord, het mondkapjesbeleid, et cetera. U krijgt hierna verschillende scenario's te zien, waarbij er op 8 verschillende factoren wordt aangegeven wat de situatie is, zoals in het voorbeeld hieronder: Geef op een vijf-punts schaal aan hoe u de kans op een besmetting met COVID-19 inschat tijdens uw vlieg- of treinreis, gebaseerd op de acht onderstaande factoren.	*Tussenpagina*

	Bezetting aan boord		75%
	Mondkapjesbeleid	Chirurgisch [type 2]	Chirurgisch [type 2] mondkapje verplicht
	Schoonmaakbeleid		Verbeterd [focus op contactpunten]
	Airconditioning		Alleen ventilatie [geen airconditioning]
	Aantal besmettingen [op bestemming]	Medium	10.000 positieve testen per dag [november 2020/juli 2021]
	Inreisvereisten	3G	Ofwel test-, vaccinatie- of herstelbewijs vereist [3g-regel]
	Vaccinatiegraad		90%
	Reisadvies		Geel



13	**Uitleg factoren inschatting kans op besmetting met COVID-19** Als u deze informatie niet wenst, dan kunt u doorgaan naar de volgende pagina. HEPA (high-efficiency particulate air) filters zijn luchtfilters die 99,9% van de lucht- en virusdeeltjes kunnen filteren uit de lucht.	*Tussenpagina*

Bezetting aan boord	De bezetting is het percentage stoelen dat bezet is van het totaal aantal stoelen aan boord van de trein of het vliegtuig. Een percentage van 50% betekent dat de helft van alle stoelen bezet is. Er zijn vier verschillende niveaus: 25%, 50%, 75% en 100%.
Mondkapjesbeleid	Dit is het beleid aan boord van het vliegtuig of de trein of mondkapjes verplicht zijn en zo ja welk soort mondkapje. Er zijn vier verschillende niveaus: geen mondkapje verplicht, mondkapje verplicht [kan elk type zijn], chirurgisch masker verplicht [type 2], FFP2 masker verplicht. Elk opeenvolgend soort mondkapje biedt een hogere bescherming (dus FFP2 beschermt het beste).
Schoonmaakbeleid	Dit is het beleid van de vervoerder of maatschappij hoe en hoe vaak er wordt schoongemaakt. Er zijn vier verschillende niveaus: hetzelfde als voor COVID-19 [dus hetzelfde schoonmaakbeleid], verbeterd [focus op contactpunten], *wekelijkse* desinfectie van het gehele voertuig en *dagelijkse* desinfectie van het gehele voertuig
Airconditioning	Dit geeft aan of het voertuig of vliegtuig airconditioning of ventilatie heeft en wat voor soort. Er zijn vier verschillende niveaus: geen airconditioning of ventilatie, alleen ventilatie, airconditioning *zonder* HEPA-filters, airconditioning *met* HEPA-filters.
Aantal besmettingen [op bestemming]	Dit is het aantal positieve testen per dag in het land waar u naartoe reist. Er zijn vier verschillende niveaus: 100 positieve testen per dag [situatie zoals in de zomer 2020 of juni 2021], 10.000 positieve testen per dag [situatie zoals in november 2020 en juli 2021], 25.000 positieve testen per dag [situatie zoals in november/december 2021] of 100.000 positieve testen per dag.
Inreisvereisten	Dit is het beleid van het land waar u heen reist omrent test-, herstel-, of vaccinatiebewijzen. Er zijn vier verschillende niveaus: geen vereisten, ofwel test-, vaccinatie- of herstelbewijs vereist [3g-regel], vaccinatie- of herstelbewijs vereist (alleen een test is dan niet meer voldoende) [2g-regel], of testen vereist *EN* vaccinatie- of herstelbewijs vereist [2g plus-regel].
Vaccinatiegraad	Dit is het percentage volledig gevaccineerde inwoners van het land waar u naartoe reist. Er zijn vier verschillende niveaus: 15%, 30%, 70% en 90%.
Reisadvies	Dit is het reisadvies van de overheid voor het land waar u naartoe reist. Er zijn vier verschillende niveaus: groen [u kunt reizen], geel [let op, er zijn risico's], oranje [alleen noodzakelijke reizen] en rood [niet reizen].

14	Rating blok 1	*Tussenpagina*

VRAAG 14 ALLEEN TONEN ALS AAN DE ONDERSTAANDE VOORWAARDEN WORDT VOLDAAN, INDIEN NIET VOLDAAN SPRING NAAR: **>> VOLGENDE VRAAG**
Variable column **Rating blok** Bevat **5**

15	**Geef op een vijfpuntsschaal aan hoe u de kans op een besmetting met COVID-19 inschat tijdens uw vlieg- of treinreis, gebaseerd op de acht onderstaande factoren.**	*Single-responsevraag*

VRAAG 15 ALLEEN TONEN ALS AAN DE ONDERSTAANDE VOORWAARDEN WORDT VOLDAAN, INDIEN NIET VOLDAAN SPRING NAAR: **>> VOLGENDE VRAAG**
Variable column **Rating blok** Bevat **1**

	Bezetting aan boord		50%
	Mondkapjesbeleid	FFP2	FFP2 gezichtsmasker verplicht
	Schoonmaakbeleid		Hetzelfde als voor COVID-19
	Airconditioning	HEPA	Airconditioning **met** HEPA-filters
	Aantal besmettingen [op bestemming]	Hoog	25.000 positieve testen per dag [november 2021]
	Inreisvereisten	3G	Ofwel test-, vaccinatie- of herstelbewijs vereist [3g-regel]
	Vaccinatiegraad		70%
	Reisadvies		Geel

- ◯ 1 - Heel laag
- ◯ 2 - Laag
- ◯ 3 - Medium
- ◯ 4 - Hoog
- ◯ 5 - Heel hoog

16	**Geef op een vijfpuntsschaal aan hoe u de kans op een besmetting met COVID-19 inschat tijdens uw vlieg- of treinreis, gebaseerd op de acht onderstaande factoren.**	*Single-responsevraag*

VRAAG 16 ALLEEN TONEN ALS AAN DE ONDERSTAANDE VOORWAARDEN WORDT VOLDAAN, INDIEN NIET VOLDAAN SPRING NAAR: **>> VOLGENDE VRAAG**
Variable column **Rating blok** Bevat **1**

	Bezetting aan boord		50%
	Mondkapjesbeleid		Geen mondkapje verplicht
	Schoonmaakbeleid	**Dagelijks**	**Dagelijkse** desinfectie gehele voertuig
	Airconditioning		Alleen ventilatie [geen airconditioning]
	Aantal besmettingen [op bestemming]	**Laag**	100 positieve testen per dag [zomer 2020/juni 2021]
	Inreisvereisten	**2G**	Vaccinatie- of herstelbewijs vereist [2g-regel]
	Vaccinatiegraad		70%
	Reisadvies		Rood

- ◯ 1 - Heel laag
- ◯ 2 - Laag
- ◯ 3 - Medium
- ◯ 4 - Hoog
- ◯ 5 - Heel hoog

17	**Geef op een vijfpuntsschaal schaal aan hoe u de kans op een besmetting met COVID-19 inschat tijdens uw vlieg- of treinreis, gebaseerd op de acht onderstaande factoren.**	*Single-responsevraag*

VRAAG 17 ALLEEN TONEN ALS AAN DE ONDERSTAANDE VOORWAARDEN WORDT VOLDAAN, INDIEN NIET VOLDAAN SPRING NAAR: **>> VOLGENDE VRAAG**
Variable column **Rating blok** Bevat **1**

	Bezetting aan boord		50%
	Mondkapjesbeleid	Elk type	Mondkapje verplicht [kan elk type zijn]
	Schoonmaakbeleid		Hetzelfde als voor COVID-19
	Airconditioning		Geen ventilatie of airconditioning
	Aantal besmettingen [op bestemming]	Medium	10.000 positieve testen per dag [november 2020/juli 2021]
	Inreisvereisten	2G+	Testen vereist **EN** vaccinatie- of herstelbewijs vereist [2g plus regel]
	Vaccinatiegraad		90%
	Reisadvies		Rood

- ◯ 1 - Heel laag
- ◯ 2 - Laag
- ◯ 3 - Medium
- ◯ 4 - Hoog
- ◯ 5 - Heel hoog

18	**Geef op een vijfpuntsschaal aan hoe u de kans op een besmetting met COVID-19 inschat tijdens uw vlieg- of treinreis, gebaseerd op de acht onderstaande factoren.**	*Single-responsevraag*

VRAAG 18 ALLEEN TONEN ALS AAN DE ONDERSTAANDE VOORWAARDEN WORDT VOLDAAN, INDIEN NIET VOLDAAN SPRING NAAR: **>> VOLGENDE VRAAG**
Variable column **Rating blok** Bevat **1**

	Bezetting aan boord		50%
	Mondkapjesbeleid	FFP2	FFP2 gezichtsmasker verplicht
	Schoonmaakbeleid	Wekelijks	**Wekelijkse** desinfectie gehele voertuig
	Airconditioning		Alleen ventilatie [geen airconditioning]
	Aantal besmettingen [op bestemming]	Heel hoog	100.000 positieve testen per dag
	Inreisvereisten	2G	Vaccinatie- of herstelbewijs vereist [2g-regel]
	Vaccinatiegraad		30%
	Reisadvies		Groen

- 1 - Heel laag
- 2 - Laag
- 3 - Medium
- 4 - Hoog
- 5 - Heel hoog

19	**Geef op een vijfpuntsschaal aan hoe u de kans op een besmetting met COVID-19 inschat tijdens uw vlieg- of treinreis, gebaseerd op de acht onderstaande factoren.**	*Single-responsevraag*

VRAAG 19 ALLEEN TONEN ALS AAN DE ONDERSTAANDE VOORWAARDEN WORDT VOLDAAN, INDIEN NIET VOLDAAN SPRING NAAR: **>> VOLGENDE VRAAG**
Variable column **Rating blok** Bevat **1**

Bezetting aan boord		50%
Mondkapjesbeleid		Geen mondkapje verplicht
Schoonmaakbeleid	Dagelijks	**Dagelijkse** desinfectie gehele voertuig
Airconditioning	Geen HEPA	Airconditioning **zonder** HEPA-filters
Aantal besmettingen [op bestemming]	Hoog	25.000 positieve testen per dag [november 2021]
Inreisvereisten	3G	Ofwel test-, vaccinatie- of herstelbewijs vereist [3g-regel]
Vaccinatiegraad		90%
Reisadvies		Groen

- ◯ 1 - Heel laag
- ◯ 2 - Laag
- ◯ 3 - Medium
- ◯ 4 - Hoog
- ◯ 5 - Heel hoog

38	**Deel 3: Keuze-experiment voor uw reis vanuit Nederland naar Europese bestemmingen.** Wij willen u nu een aantal scenario's voorleggen voor een reis naar een Europese bestemming, waarbij u telkens kunt kiezen tussen een reis met de auto, het vliegtuig of de trein. Per scenario verschillen de situaties van een reis, als het gaat om reistijd, reiskosten, reiscomfort en de ingeschatte besmettingskans. Kunt u hierna voor elk scenario aangeven voor welk vervoermiddel u zou kiezen?	*Tussenpagina*

39	**Uitleg factoren keuze-experiment.** Als u deze informatie niet wenst, dan kunt u doorgaan naar de volgende pagina. In dit deel is de kans op besmetting een gegeven, en is het dus niet de bedoeling dat u zelf een inschatting maakt van de kans op besmetting zoals u in deel 2 gedaan heeft.	*Tussenpagina*

Reistijd	Dit is de reistijd vanuit Amsterdam (centrum) tot aan uw bestemming in het centrum. Voor alle vervoerswijzen (trein, vliegtuig en auto) geldt dat de reistijd dus van deur tot deur is (dus inclusief eventueel voor- en natransport van en naar bijvoorbeeld de luchthaven of het station).
Reiskosten	Dit zijn de totale reiskosten voor de gehele reis, dit is ook inclusief eventueel voor- en na transport. Voor de auto geldt dat dit de benzinekosten zijn en alle andere overige kosten van een auto.
Comfort	Dit is de reisklasse waarin u zult reizen. Voor de trein kan dit de 2^e klasse of 1^e klasse zijn, voor het vliegtuig is dit economy class of business class. In de auto varieert het comfort niet. Dit betekent dus dat u altijd met hetzelfde type auto rijdt.
Inschatting kans op besmetting met COVID-19	Dit is de inschatting van de kans op besmetting met COVID-19 aan boord van het vliegtuig of de trein. Dit komt voort uit uw inschattingen zoals gedaan in deel 2. In dit geval betekent het dus dat het besmettingsrisico een gegeven is, en niet dat u zelf een inschatting maakt van het besmettingsrisico zoals u in deel 2 gedaan heeft. De niveaus zijn 1-Heel laag, 3-Medium of 5-Heel hoog. In de auto zal de inschatting van het op een risico besmetting met COVID-19 niet gevarieerd worden.

40	**U krijgt nu 4 vragen over een reis naar een Europese bestemming, met een reisafstand van 400 tot 600 kilometer.** Denk bijvoorbeeld aan bestemmingen zoals Londen, Parijs, Zürich, Berlijn en Kopenhagen	*Tussenpagina*

41	**400-600km blok 1**	*Tussenpagina*

VRAAG 41 ALLEEN TONEN ALS AAN DE ONDERSTAANDE VOORWAARDEN WORDT VOLDAAN, INDIEN NIET VOLDAAN SPRING NAAR: **>> VOLGENDE VRAAG**
Variable column **400-600km blok** Bevat **5**

42	**Welk vervoermiddel zou u kiezen, als u kunt kiezen uit de onderstaande drie opties?** U maakt een reis naar een Europese bestemming, met een reisafstand van 400 tot 600 kilometer. Denk bijvoorbeeld aan bestemmingen zoals Londen, Parijs, Zürich, Berlijn of Kopenhagen.	*Single-responsevraag*

VRAAG 42 ALLEEN TONEN ALS AAN DE ONDERSTAANDE VOORWAARDEN WORDT VOLDAAN, INDIEN NIET VOLDAAN SPRING NAAR: **>> VOLGENDE VRAAG**
Variable column **400-600km blok** Bevat **1**

	Modaliteit	Trein	Vliegtuig	Auto
	Reistijd	6 uur	3 uur	6.5 uur
	Reiskosten	€ 30	€ 50	€ 150
	Comfort	2e klasse	Business	n.v.t
	Inschatting kans op besmetting met COVID-19	5-Heel hoog	1-Heel laag	1-Heel laag

- ◯ Trein
- ◯ Vliegtuig
- ◯ Auto

43	**Welk vervoermiddel zou u kiezen, als u kunt kiezen uit de onderstaande drie opties?** U maakt een reis naar een Europese bestemming, met een reisafstand van 400 tot 600 kilometer. Denk bijvoorbeeld aan bestemmingen zoals Londen, Parijs, Zürich, Berlijn of Kopenhagen.	*Single-responsevraag*

VRAAG 43 ALLEEN TONEN ALS AAN DE ONDERSTAANDE VOORWAARDEN WORDT VOLDAAN, INDIEN NIET VOLDAAN SPRING NAAR: **>> VOLGENDE VRAAG**
Variable column **400-600km blok** Bevat **1**

	Modaliteit			
	Reistijd	4.5 uur	5 uur	4.5 uur
	Reiskosten	€ 300	€ 175	€ 80
	Comfort	2e klasse	Business	n.v.t
	Inschatting kans op besmetting met COVID-19	3-Medium	5-Heel hoog	1-Heel laag

- ◯ Trein
- ◯ Vliegtuig
- ◯ Auto

44	**Welk vervoermiddel zou u kiezen, als u kunt kiezen uit de onderstaande drie opties?** U maakt een reis naar een Europese bestemming, met een reisafstand van 400 tot 600 kilometer. Denk bijvoorbeeld aan bestemmingen zoals Londen, Parijs, Zürich, Berlijn of Kopenhagen.	*Single-responsevraag*

VRAAG 44 ALLEEN TONEN ALS AAN DE ONDERSTAANDE VOORWAARDEN WORDT VOLDAAN, INDIEN NIET VOLDAAN SPRING NAAR: **>> VOLGENDE VRAAG**
Variable column **400-600km blok** Bevat **1**

	Modaliteit			
	Reistijd	3 uur	3 uur	8.5 uur
	Reiskosten	€ 300	€ 175	€ 80
	Comfort	1^e klasse	Economy	n.v.t
	Inschatting kans op besmetting met COVID-19	1-Heel laag	5-Heel hoog	1-Heel laag

- ○ Trein
- ○ Vliegtuig
- ○ Auto

45	**Welk vervoermiddel zou u kiezen, als u kunt kiezen uit de onderstaande drie opties?** U maakt een reis naar een Europese bestemming, met een reisafstand van 400 tot 600 kilometer. Denk bijvoorbeeld aan bestemmingen zoals Londen, Parijs, Zürich, Berlijn of Kopenhagen.	*Single-responsevraag*

VRAAG 45 ALLEEN TONEN ALS AAN DE ONDERSTAANDE VOORWAARDEN WORDT VOLDAAN, INDIEN NIET VOLDAAN SPRING NAAR: **>> VOLGENDE VRAAG**
Variable column **400-600km blok** Bevat **1**

	Modaliteit	Trein	Vliegtuig	Auto
	Reistijd	6 uur	4 uur	4.5 uur
	Reiskosten	€ 165	€ 300	€ 115
	Comfort	1e klasse	Economy	n.v.t
	Inschatting kans op besmetting met COVID-19	5-Heel hoog	1-Heel laag	1-Heel laag

- ◯ Trein
- ◯ Vliegtuig
- ◯ Auto

56	**U krijgt nu 4 vragen over een reis naar een Europese bestemming, met een reisafstand van 800 tot 1200 kilometer.** Denk bijvoorbeeld aan bestemmingen zoals Bordeaux, Milaan, Barcelona, Warschau en Stockholm.	*Tussenpagina*

57	**800-1200km blok 1**	*Tussenpagina*

VRAAG 57 ALLEEN TONEN ALS AAN DE ONDERSTAANDE VOORWAARDEN WORDT VOLDAAN, INDIEN NIET VOLDAAN SPRING NAAR: **>> VOLGENDE VRAAG**
Variable column **800-1200km blok** Bevat **5**

58	**Welk vervoermiddel zou u kiezen, als u kunt kiezen uit de onderstaande drie opties?** U maakt een reis naar een Europese bestemming, met een reisafstand van 800 tot 1200 kilometer. Denk bijvoorbeeld aan bestemmingen zoals Bordeaux, Milaan, Barcelona, Warschau en Stockholm.	*Single-responsevraag*

VRAAG 58 ALLEEN TONEN ALS AAN DE ONDERSTAANDE VOORWAARDEN WORDT VOLDAAN, INDIEN NIET VOLDAAN SPRING NAAR: **>> VOLGENDE VRAAG**
Variable column **800-1200km blok** Bevat **1**

	Modaliteit	Trein	Vliegtuig	Auto
	Reistijd	6 uur	5 uur	16 uur
	Reiskosten	€ 350	€ 225	€ 100
	Comfort	2e klasse	Business	n.v.t
	Inschatting kans op besmetting met COVID-19	1-Heel laag	5-Heel hoog	1-Heel laag

- ◯ Trein
- ◯ Vliegtuig
- ◯ Auto

59	**Welk vervoermiddel zou u kiezen, als u kunt kiezen uit de onderstaande drie opties?** U maakt een reis naar een Europese bestemming, met een reisafstand van 800 tot 1200 kilometer. Denk bijvoorbeeld aan bestemmingen zoals Bordeaux, Milaan, Barcelona, Warschau en Stockholm.	*Single-responsevraag*

VRAAG 59 ALLEEN TONEN ALS AAN DE ONDERSTAANDE VOORWAARDEN WORDT VOLDAAN, INDIEN NIET VOLDAAN SPRING NAAR: **>> VOLGENDE VRAAG**
Variable column **800-1200km blok** Bevat **1**

	Modaliteit			
	Reistijd	12 uur	6 uur	10 uur
	Reiskosten	€ 200	€ 400	€ 100
	Comfort	1e klasse	Economy	n.v.t
	Inschatting kans op besmetting met COVID-19	3-Medium	3-Medium	1-Heel laag

- ◯ Trein
- ◯ Vliegtuig
- ◯ Auto

60	**Welk vervoermiddel zou u kiezen, als u kunt kiezen uit de onderstaande drie opties?** U maakt een reis naar een Europese bestemming, met een reisafstand van 800 tot 1200 kilometer. Denk bijvoorbeeld aan bestemmingen zoals Bordeaux, Milaan, Barcelona, Warschau en Stockholm.	*Single-responsevraag*

VRAAG 60 ALLEEN TONEN ALS AAN DE ONDERSTAANDE VOORWAARDEN WORDT VOLDAAN, INDIEN NIET VOLDAAN SPRING NAAR: **>> VOLGENDE VRAAG**
Variable column **800-1200km blok** Bevat **1**

	Modaliteit	Trein	Vliegtuig	Auto
	Reistijd	9 uur	5 uur	13 uur
	Reiskosten	€ 50	€ 50	€ 150
	Comfort	2^e klasse	Business	n.v.t
	Inschatting kans op besmetting met COVID-19	5-Heel hoog	1-Heel laag	1-Heel laag

- ◯ Trein
- ◯ Vliegtuig
- ◯ Auto

61	**Welk vervoermiddel zou u kiezen, als u kunt kiezen uit de onderstaande drie opties?** U maakt een reis naar een Europese bestemming, met een reisafstand van 800 tot 1200 kilometer. Denk bijvoorbeeld aan bestemmingen zoals Bordeaux, Milaan, Barcelona, Warschau en Stockholm.	*Single-responsevraag*

VRAAG 61 ALLEEN TONEN ALS AAN DE ONDERSTAANDE VOORWAARDEN WORDT VOLDAAN, INDIEN NIET VOLDAAN SPRING NAAR: **>> VOLGENDE VRAAG**
Variable column **800-1200km blok** Bevat **1**

	Modaliteit	Trein	Vliegtuig	Auto
	Reistijd	9 uur	4 uur	16 uur
	Reiskosten	€ 350	€ 225	€ 200
	Comfort	1e klasse	Economy	n.v.t
	Inschatting kans op besmetting met COVID-19	1-Heel laag	5-Heel hoog	1-Heel laag

- ◯ Trein
- ◯ Vliegtuig
- ◯ Auto

72	**Deel 4: Achtergrondkenmerken** In dit deel worden enkele achtergrondkenmerken gevraagd.	*Tussenpagina*

73	**Wat is uw geslacht?**	*Single-responsevraag*

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

74	**In welke leeftijdscategorie valt u?**	*Single-responsevraag*

- Jonger dan 20
- 20 tot 40 jaar
- 40 tot 65 jaar
- 65 tot 80 jaar
- 80 jaar of ouder
- Zeg ik liever niet

75	**Hoe is op dit moment de samenstelling van uw huishouden (inclusief uzelf)?**	*Single-responsevraag*

- Alleen ikzelf
- Ikzelf met partner
- Ikzelf met partner en kind(eren) (ook deeltijdzorg)
- Ikzelf met kind(eren) (ook deeltijdzorg)
- Ikzelf met ouder(s) of verzorger(s)
- Ikzelf met ouder(s) of verzorger(s), broertje(s) en/of zusje(s)
- Ikzelf met meerdere volwassenen (16 jaar of ouder) (bijv. studentenhuis, verzorgingstehuis)
- Anders, namelijk
- Zeg ik liever niet

76	Wat was het totale inkomen van u afgelopen jaar (2021)?	*Single-responsevraag*

- ◯ Minder dan €10.000
- ◯ € 10.000 tot € 20.000
- ◯ € 20.000 tot € 30.000
- ◯ € 30.000 tot € 40.000
- ◯ € 40.000 tot € 50.000
- ◯ € 50.000 tot € 100.000
- ◯ € 100.000 tot € 200.000
- ◯ € 200.000 of meer
- ◯ Zeg ik liever niet

77	Welke van onderstaande situaties is het meest op u van toepassing?	*Single-responsevraag*

- ◯ Schoolgaand / scholier
- ◯ Studerend / student
- ◯ Werkzaam in loondienst
- ◯ Werkzaam bij de overheid
- ◯ Zelfstandig ondernemer
- ◯ Freelancer of ZZP'er
- ◯ Vrijwilliger
- ◯ Werkloos / werkzoekend / bijstand
- ◯ Huisvrouw / huisman
- ◯ Gepensioneerd / VUT
- ◯ Arbeidsongeschikt
- ◯ Anders, namelijk ________
- ◯ Zeg ik liever niet

78	Weet u zeker dat u de enquête wilt verlaten? Zo nee, klik dan op terug.	*Single-responsevraag*

VRAAG 78 ALLEEN TONEN ALS AAN DE ONDERSTAANDE VOORWAARDEN WORDT VOLDAAN, INDIEN NIET VOLDAAN SPRING NAAR: **>> VOLGENDE VRAAG**
Vraag 2. **Om deel te nemen aan deze enquête moet u akkoord gaan met de voorwaarden.** Is gegeven of Vraag 2. **Om deel te nemen aan deze enquête moet u akkoord gaan met de voorwaarden.** Is gegeven of Vraag 2. **Om deel te nemen aan deze enquête moet u akkoord gaan met de voorwaarden.** Is gegeven

○ Ik weet het zeker

79 Niet akkoord respondenten

VRAAG 79 ALLEEN TONEN ALS AAN DE ONDERSTAANDE VOORWAARDEN WORDT VOLDAAN, INDIEN NIET VOLDAAN SPRING NAAR: **>> EINDE ONDERZOEK**
Vraag 2. **Om deel te nemen aan deze enquête moet u akkoord gaan met de voorwaarden.** Is gegeven of Vraag 2. **Om deel te nemen aan deze enquête moet u akkoord gaan met de voorwaarden.** Is gegeven of Vraag 2. **Om deel te nemen aan deze enquête moet u akkoord gaan met de voorwaarden.** Is gegeven

80 Vul na deze pagina uw in als u wilt deelnemen

The influence of perceived COVID-19 risk on the modal-split for long-distance travel in Europe: a Hierarchical Information Integration and Stated-Preference study approach

M.M. van Dalen

Faculty of Civil Engineering & Geosciences, Delft University of Technology, Stevinweg 1, Delft, 2628 CN, The Netherlands.

Contributing authors: M.vanDalen

Abstract

COVID-19 had (and still has) a huge impact on travel behaviour. Most research on the effects of COVID-19 on travel behaviour is based on daily travel behaviour. Long-distance travel is most of the time not included. Therefore, this paper studies the impact of perceived COVID-19 risk on mode-choice for long-distance travel in Europe. This paper uses a Stated-Preference approach which consists of two parts, the perceived risk rating experiment and the main mode choice experiment. In total 1147 responses were collected. In the rating experiment, respondents rated their perceived risk based on eight attributes. In the main choice experiment, respondents chose between train, plane and care based on travel time, travel cost, travel comfort and perceived risk. With this, the Value of Risk (VoR) for a decrease in perceived risk is derived, both expressed in travel cost and travel comfort. Moreover, the Willingness to Pay (WtP) for comfort and Value of Time (VoT) are derived. With the combination of both models, the WtP for risk attributes given a perceived risk level are derived. The combination of both models shows the Willingness to Pay values for the different risk attributes, given a certain perceived risk level. To test the influence of perceived risk on modal split, three routes with real world values are used. The results implicate a maximum of 5% market share difference possible. This shows that perceived COVID-19 has a moderate effect on modal split.

Keywords: Stated-Preference, Perceived COVID-19 risk, MNL model, Discrete Choice Modelling, Hierarchical Information Integration (HII), long-distance travel, risk

1 Introduction

The new coronavirus (SARS-CoV-2) (from now: COVID-19) is an ongoing global pandemic that is considered one of the worst post-World War II pandemics that affects the world surpassing both MERS and SARS outbreaks (Matiza, 2020). Based on rising infection rates in China and then to over the world, the WHO Emergency Committee declared a global health emergency on January 30, 2020 (Velavan and Meyer, 2020). Countries have taken several measures to reduce the number of infections; some of these measures were restrictions from entering other countries, the closing of restaurants/bars and measures during the trip (like a face mask or negative test). Because of these measures, COVID-19 has a huge impact on travel behaviour (De Vos, 2020).

One of the elements of travel behaviour is mode choice. It is determined by a lot of factors that are often interrelated to each other to a smaller or larger extent (De Witte, Hollevoet, Dobruszkes, Hubert, and Macharis, 2013). Quite some research is done on the effects of mode-choice, mainly focusing on daily (short-distance) travel behaviour (Buehler, 2011; Atasoy, Glerum, and Bierlaire, 2013; De Witte et al., 2013). However, long-distance travel behaviour is often excluded from the analysis. At the same time, over 50% of the passenger-kilometers travelled come from long-distance travel (Aamaas, Borken-Kleefeld, and Peters, 2013). Several studies researched the main drive for passengers to choose a certain mode for long-distance travel in Europe. The main finding of these studies was that travel time seems to be the most important factor (van Goeverden, 2009). Important other factors are travel cost, comfort, reliability, access & egress time and number of interchanges (Román, Espino, and Martín, 2010).

Several papers researched the effects of COVID-19 infection on mode-choice, but the focus was on short-distance travel. Abdullah, Dias, Muley, and Shahin (2020) found that people tend to use less public transport services and more private cars during the pandemic. There was a shift found to active modes as well. Moreover, sanitisation measures and social distancing characterise the perception of safety and, therefore, the willingness to use public transport services (Aaditya and Rahul, 2021). According to de Haas, Faber, and Hamersma (2020) in the Netherlands, around 80% of the people reduced their activities, with an increase from 6% of the people working at home to 39%. For public transport, there was a drop in usage of 90%. Shamshiripour, Rahimi, Shabanpour, and Mohammadian (2020) noticed that people tend to shift more to individual and active modes (e.g., walking or cycling) of travel or not travelling at all. Another possible reason for the reduction of travel by public transport is that people believe that public transport is an unhygienic place, with a high chance of infection with viruses (Troko et al., 2011).

At this stage, most of the COVID-19 measures are gone. Partly this is a consequence of the omicron-variant of the COVID-19 virus (Chen, 2022). Thus, lower hospitalisation rates can be found. The Dutch society accepted to 'live' with COVID-19 because COVID-19 will not fully be gone. By definition, a lot of different people are transported at the same time in PT systems; therefore, the virus can relatively easy be transmitted among travellers. To illustrate this, Krishnakumari and Cats (2020) found that, on average, a person interacts with 1200 other people on a single trip in the metro system network of Washington D.C. Now, it seems evident that measures were (and still are) taken in order to reduce the number of infections and mitigate the public health crisis.

To study the (risk) effects of COVID-19 on mode-choice, a new variable is introduced: perceived risk. Perceived risk is the risk people perceive during their trip of getting infected with COVID-19. It is often defined as the perceived likelihood of getting the disease times the perceived severity of the symptoms (Karlsson et al., 2021). For this study, however, perceived risk consists on mode-related and destination-related attributes. Perceived risk does also have an emotional dimension like worry and fear (Loewenstein, Weber, Hsee, and Welch, 2001). Potentially, people will change their travel behaviour (and thus mode-choice) due to these factors as well as other factors of their trip. The outcome of this study could be beneficial in transport planning and policy-making during health crises. Providers of PT services (like airlines or rail operators) can use the information the optimise their services and operations.

This work contributes to the scientific literature because it is one of the first to examine the effects of COVID-19 on modal-split for long-distance travel in Europe. This is studied using a mode choice experiment with Stated-Preference data. In addition, a perceived risk rating experiment is done to determine the factors that influence the perception of risk posed by a COVID-19 infection. Using a modified variant of the HII methodology, combining both models gives additional insights, such as the Value of Risk (Willingness to Pay for reduction in risk). On data acquired from a sample of 1147 (predominantly regular) train passengers recruited in the Netherlands, this method is used and model results are provided.

2 Methodology

Perceived risk due to COVID-19 is complicated to measure, with several elements/attributes that could contribute to this perceived risk. These attributes possibly weigh different for each individual. Therefore, perceived risk is a complex variable, its score is determined by other variables (Molin, 2020).

2.1 Hierarchical Information Integration (HII) theory

This research uses the Hierarchical Information Integration (HII) theory which was introduced by Louviere (1984). This theory is used when decision-makers are confronted with many attributes. Decision-makers categorise these attributes into 'decision constructs'. Decision-makers (respondents) trade-off attributes that belong to such a 'decision-construct' in the first (sub) experiment, the 'rating' experiment. Then in the 'bridging' experiment, decision-makers make a trade-off between the construct evaluations that are done in the 'rating' experiment (Molin, 2020). For this study, an adapted version of the classical HII experiment is used, with only one 'decision construct', to determine perceived risk. Estimating the 'rating' model allows for predicting the perceived risk. Then in the main choice experiment, the perceived risk attribute is shown among the other attributes that are defined. This is not a true 'bridging' experiment as in a conventional 'bridging' experiment more sub experiments for the decision constructs are used; in this research; however, only one. Molin and van Gelder, (2008) showed that this approach is successful.

2.2 Perceived risk rating experiment

The objective of the perceived risk rating experiment is to determine to what degree COVID-19 infection risk variables impact the perceived risk rating of a train or plane trip. To do this, respondents will be asked to score their perceived risk for various trip combinations (by either train or plane). This risk rating is affected by several variables. The determinants/attributes are based mainly on the papers of Dryhurst et al. (2020), Mertens, Gerritsen, Duijndam, Salemink, and Engelhard (2020), Tirachini and Cats, (2020) and Leppin and Aro (2009). All attributes do have four levels. The first four attributes are mode-related attributes. The last four attributes are destination-related attributes.

1. **On-board crowding**: This is specified as the percentage of seats occupied on-board of the vehicle/plane. The levels are: 25% of the seats occupied, 50% of the seats occupied, 75% of the seats occupied and 100% of the seats occupied.
2. **Face mask policy**: This is the policy on-board the vehicle whether or not or which mask is mandatory. The attribute levels are a representation of the masks available. The four levels are: no mask mandatory, any face mask mandatory, at least a surgical (type II) mask mandatory or at least an FFP2 mask mandatory. Every increase in the level of the type of mask gives better protection.
3. **Cleaning policy**: This is policy the train company or airlines has regarding cleaning. The four levels are: same cleaning policy as before COVID-19, increased cleaning policy (focus on touching points), weekly disinfection of the whole vehicle and daily disinfection of the whole vehicle. The levels are based on several policies that airlines and rail companies implemented during the pandemic.

4. **Air conditioning/ventilation**: This attributes is about the type of ventilation or air conditioning is on-board the vehicle or plane. The levels are no ventilation and air conditioning, only ventilation, air conditioning without HEPA filters and air conditioning with HEPA filters.
5. **Travel requirements**: Several policies are implemented within Europe, like 3G or 2G. The following levels were chosen, they reflect different policies within Europe: no mandatory requirements, either testing, vaccination or recovery proof required (3G-rule), only vaccination or recovery proof required (2G-rule), and vaccination or recovery + testing required or booster required (2G+-rule).
6. **Infection rate**: The infection rate levels do reflect different time moments during the pandemic. The levels are: 100 positive tests per day (reflects the situation of the summer of 2020 or in June 2021), 10.000 positive tests per day (reflects the situation of November 2020 and July 2021), 25.000 tests per day (reflects the situation of November 2021) and 100.000 positive tests per day (fictitious situation).
7. **Vaccination rate**: The levels reflect the vaccination rates in different European countries. The levels are from December 2021: 15% vaccination rate (level as in Bulgaria), 30% vaccination rate (level as in Romania), 70% vaccination rate (level as in the Netherlands and EU average), 90% vaccination rate (level as in Portugal).
8. **Travel advice**: These are the travel advice from the Dutch government (Rijksoverheid, 2021). The levels are: green, yellow, orange and red travel advice.

Ngene is used as the software to generate choice sets (the experimental design). There is no prior information available for the rating experiment. This is required in order to create efficient designs. An orthogonal design is chosen because it seeks an attribute level balanced design, this means that the attribute levels occur the same number of times in the option sets. It is decided to do an unlabeled experiment. The rating experiment will only be conducted for train and plane modes. To simplify, there is no differentiation inside the rating experiment. This suggests that people are either taking the train or flying. The rating experiment makes no difference between these modes. As a consequence, the choice sets may be constructed in a sequential order. Ngene generates an orthogonal design with 20 rows. Four blocks are used within the design and each respondent is given five questions to assess their level of perceived COVID-19 risk. An example of the presented rating experiment can be found in figure 1.

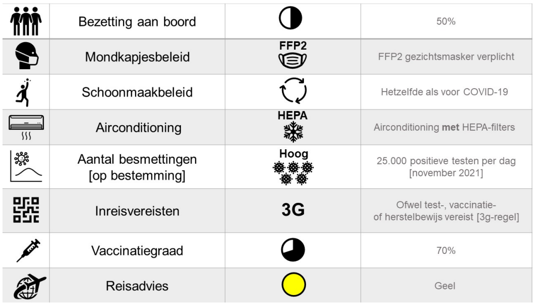


Fig. 1 Example of rating experiment

2.3 Main choice experiment

The main goal of the main (mode) choice experiment is to study how perceived risk is weighted against other factors, such as for example travel cost, while making mode choice decisions. In this case, perceived risk is an independent variable amongst the other main variables. The determinants of the main choice experiment are travel time, travel cost, travel comfort and perceived risk. All attributes do have three levels. The attributes of travel time and travel cost are varied for all modes. As there are two distance classes, in total there are six attribute levels per mode. Travel comfort and perceived risk are only varied for plane and train. Perceived risk and travel comfort are not varied for car. This is because it is assumed that respondents are not sharing their car with strangers. As a consequence, perceived risk in the car is always very low. For travel comfort, it is assumed that people 'own' the same car within the experiment. Therefore, the comfort of the car does not change; thus, the levels of comfort are not varied. The attribute levels can be found in table 1 & 2. For every attribute it is shortly explained why this is included in the main choice experiment.

Travel cost: One of the most important variables in travel behaviour research, and (almost) always included within stated (mode) choice experiments. Trip cost does refer to the cost of making a trip. This can be either the ticket price or the total price for driving the car (fuel + any additional cost). Travel cost is often used in choice experiments to retrieve the willingness to pay for improvements in one of the other attributes of interest. In the case of this research, trade-offs regarding COVID-19 risk are of interest regarding mode-choice for long-distance travel within Europe. From literature, it can be concluded that travel cost is one of the most important variables for mode-choice on long-distance travel (van Goeverden, 2009; Román et al., 2010; Dobruszkes, Dehon, and Givoni, 2014).

Travel time: This variable/attribute is also (almost) always taken into account in studies regarding mode-choice (Morikawa, Ben-Akiva, and McFadden, 2002; van Goeverden, 2009; Román et al., 2010). Some studies refer to the total travel time (thus including access, egress, transfer and waiting time). Other studies only refer to the in-vehicle time. In this study, the total travel time will be used. This includes in-vehicle time and transfer time (if this applies). **Access and egress** are incorporated in the total travel time in this study. This is done in order to keep alternatives simple and understandable for the respondents.

Travel comfort: Travel comfort is also an important factor regarding mode choice for long-distance travel. Román et al. (2010) included comfort as an attribute in their research. The willingness to pay increased if the level of comfort was lower. Furthermore, they found that increased levels of comfort in the aircraft did decrease the perception of time. Train companies and airlines do offer different levels of comfort by distinguishing travel classes. For plane, often, there is a choice between economy class and business class. For the train, this is mostly 2^{nd} class and 1^{st} class.

Perceived risk: Perceived risk is the last attribute that will be included in the main choice experiment. This attribute is directly connected with the rating experiment. In the rating experiment, respondents did rate the risk of getting infected with COVID-19 on their train or plane journey based on eight factors. Respondents rated their journey on a Likert scale with 1-very low, 2-low, 3-medium, 4-high and 5-very high. In the main choice experiment, perceived risk is an attribute amongst the other main attributes. As all levels in the main choice experiment do have three levels, perceived risk does have three levels as well. Therefore, the levels are 1-low, 3-medium and 5-high.

Table 1 Travel time & travel cost values

Travel time							
400-600 km				800-1200km			
# levels	Mode	Value attribute levels		# levels	Mode	Value attribute levels	
		180 min	3h			360 min	6h
	Train	270 min	4.5h		Train	540 min	9h
		360 min	6h			720 min	12h
		180 min	3h			240 min	4h
3	Airplane	240 min	4h	3	Airplane	300 min	5h
		300 min	5h			360 min	6h
		270 min	4.5h			600 min	10h
	Car	390 min	6.5h		Car	780 min	13h
		510 min	8.5h			960 min	16h

Travel cost					
400-600 km			800-1200km		
# levels	Mode	Value attribute levels	# levels	Mode	Value attribute levels
		30 euro			50 euro
	Train	165 euro		Train	200 euro
		300 euro			350 euro
		50 euro			50 euro
3	Airplane	175 euro	3	Airplane	225 euro
		300 euro			400 euro
		80 euro			100 euro
	Car	115 euro		Car	150 euro
		150 euro			200 euro

Table 2 Travel comfort & perceived risk levels

Travel comfort		
# levels	Mode	Value attribute levels
2	Train	2nd class
		1st class
2	Airplane	Economy
		Business
1	Car	Same level

Perceived risk		
# levels	Mode	Value attribute levels
		1-Very low
3	Train	2-Medium
		3-Very High
		1-Very low
3	Airplane	2-Medium
		3-Very High
1	Car	1-Very low

In this part, the included socio-demographic attributes and the travel behaviour questions are discussed.

Age: Is one of the most common used socio-demographic variables. Several papers and research did look include this into their research: Buehler and Nobis (2010), Hensher and Rose (2007), Paulssen, Temme, Vij, and Walker (2014), Román et al. (2010). Often different age groups do have different preferences for certain modes.

Gender: This socio-demographic variable is also a very common variable to include in stated-choice experiments. Almost all studies include this socio-demographic variable (Buehler and Nobis 2010; Hensher and Rose 2007; Paulssen et al. 2014; Román et al. 2010; Johansson, Heldt, and Johansson 2006). With this variable, it can be analysed if women and men do have do have different preferences regarding the variables in the main choice experiment, this could be for example in the preference for a certain mode.

Income: Also an important socio-demographic variable to take into account. It is not always clear how this income is asked in the survey. Some papers ask about dispensable income, while at the same time other papers ask for gross income (Buehler and Nobis 2010; Hensher and Rose 2007; Paulssen et al. 2014; Román et al. 2010; Johansson et al. 2006). For this study, the gross-income is used. As it is expected that higher income influences mode-choice, this variable is included in the model. It is expected that higher income will increase the willingness to pay for the attributes, like time and comfort for example.

Work status: Hensher and Rose (2007) included this variable in their research in order to check if the sample was representative, but they did not include this in the model specification. However, for this study it is included in the model specification as well.

Education level: Socio-demographic variable that is often included in models as well. Johansson et al. (2006) stated that they previously did not find any literature on including education level for long-distance travel; however, in his research, it turned out to be significant. Education level is also expected to have an influence on mode-choice and, therefore, is included in the model.

Travel frequency: This variable is sometimes included in research. Several research included this such as Román et al. (2010), Van Loon and Rouwendal (2013) & Nieto-García, Muñoz-Gallego, and Gonzalez-Benito (2020). For this research, this attribute will be included to test whether travel frequency of respondents influences perceived risk.

Preferred travel mode: Hensher and Rose (2007) included the preferred travel mode in their study. For this mode-choice research, it is interesting to see if the preferred travel mode influences the mode choice. It is expected that people will stick to their main preferred mode when making a choice. In this research, it will be questioned for both the 400-600km and 800-1200km distance classes.

Payment: This attribute is included in order test whether the value of time values changes when payment is done by the respondents or by someone else, or education/work. Kouwenhoven et al. (2014) found significant different in the VoT values for different purpose of work.

Trip purpose: Both Buehler and Nobis (2010) and Román et al. (2010) included trip purpose in their research, and for both studies this turned out to be an important factor. Often willingness to pay for business trips is higher (as the respondent does not pay by him or herself) than for leisure trips. Therefore this will also be taken into account in this research. However, this is not done in the same manner as in these studies. In this research, people are asked if they pay for themselves or if someone else is paying for their trip. In this case, it can still be analysed if the willingness to pay is higher if respondents do not have to pay for themselves.

Travel company: A study done by Mertens et al. (2020) showed that risk for family and loved ones, was one of the most important predictors for COVID-19 fear. Therefore it is included also in this research, to test whether travel company contributes to perceived risk.

For the main choice experiment, an efficient design is chosen. An efficient design results in fewer choice sets for the survey than an orthogonal design. In this case, there is prior information available; however, not for all four variables. Therefore a Bayesian D-efficient design is chosen, so the prior values can differ around the mean. Ngene finds an efficient design with only ten rows. It is chosen to go for a design with 12 rows as this number can be divided by three. The reason for this is the fact that 12 choice sets give more information. As the main choice is divided into two distance classes, this gives in total eight main choice questions for the respondents. An example of the main choice experiment is shown in figure 2.

	Modaliteit			
	Reistijd	4.5 uur	5 uur	4.5 uur
	Reiskosten	€ 300	€ 175	€ 80
	Comfort	2e klasse	Business	n.v.t
	Inschatting kans op besmetting met COVID-19	3-Medium	5-Heel hoog	1-Heel laag

Fig. 2 Example of main choice experiment question (in Dutch)

3 Characteristics of the sample

The intended target group is a representative sample of the Dutch population. Data was collected in several ways: by sharing the link of the survey on social media platforms, by sharing the link to colleagues and by distributing the survey to the customer panel of NS, the Dutch Railways. The survey was open for a response from the 8^{th} of February to the 8^{th} of March. Cooperation with NS was done. The survey was also distributed to their panel. In total, the survey was sent to 5000 people. In total, 938 respondents took part in the survey and fully completed the survey. This is a response rate of 18.7%. Moreover, the link to the survey was distributed among friends, family, and colleagues and through social media platforms like LinkedIn, Instagram, and the Royal HaskoningDHV C-Infra department). This resulted in a total of 209 completed responses. All in all, the total number of completed responses of this survey was 1147, which is way above the minimum needed respondents. The sample consists mainly of highly educated people and people with high income. For age and gender, the sample approximates the Dutch population. The consequence of the higher incomes and higher education is that presumably the Value of Time and Willingness to Pay values are overestimated. The frequencies, percentages and percentages from CBS can be found in table 3.

Table 3 Characteristics of the sample

Socio-demographic variable	Category	Frequency	Percentage sample	Percentage CBS
Age	0-19	27	2.4%	21
	20-40	261	23.1%	25% / 34%[1]
	40-65	418	36.9%	34% / 43%[1]
	65 to 80	371	32.8%	15% / 19%[1]
	80+	55	4.9%	5% / 6%[1]
	Total	1132		
Gender	Female	547	49.0%	50.3%
	Male	569	51.0%	49.7%
	Total	1116		
Income	€10.000	77	8.5%	13.6%
	€10.000-€20.000	85	9.4%	23.3%
	€20.000-€30.000	117	12.9%	18%
	€30.000-€40.000	156	17.2%	14.7%
	€40.000-€50.000	141	15.5%	10.9%
	€50.000-€100.000	247	27.2%	16.5%
	€100.000-€200.000	76	8.4%	2.6%
	€200.000 or more	10	1.1%	0.4%
	Total	909		
Education	Basisonderwijs	7	0.8%	8.3%
	Vmbo-b/k, mbo1	22	2.4%	10.7%
	Vmbo-g/t, vwo-onderbouw	65	7.0%	8.4%
	Mbo2, mbo3 en mbo4	119	12.9%	26.6%
	Havo, vwo	93	10.0%	9.5%
	Hbo-,wo-bachelor	258	27.9%	21.9%
	Hbo-,wo-master, doctor	323	34.9%	13%
	Do not know	39	4.2%	1.7%
	Total	926		

4 Model estimation

For the rating experiment, respondents rated their perceived risk of COVID-19 based on a Likert scale. A Likert scale is a method for interrogating data that is difficult to quantify and providing it with an ordinal level of measurement (Joshi, Kale, Chandel, and Pal, 2015). Therefore it is widely used in questionnaires and surveys. It was chosen to go for a five-point scale, as this is easy for respondents to rate. The rating is analysed using a linear regression model. The data is analysed with IBM SPSS statistics 26.0. The following regression formula is for the perceived risk rating.

$$
\begin{gathered}
PR_{COVID-19} = C + \beta_{RA} * RA + \beta_{VR} * VR + \beta_{LF} * LF \\
+\beta_{INFECT1} * INFECT1 + \beta_{INFECT2} * INFECT2 \\
+\beta_{INFECT3} * INFECT3 + \beta_{VL} * VL + \beta_{NH} * NH \\
+\beta_{HP} * HP + \beta_{3G} * 3G + \beta_{2G+} * 2G + +\beta_{AM} * AM \\
+\beta_{FFP2} * FFP2 + \beta_{GEN} * GEN + \beta_{HBO} * HBO \\
+\beta_{WO} * WO + \beta_{INC_{20-40}} * income_{20-40}
\end{gathered}
$$

C = constant, RA = Red travel advice, VR = vaccination rate, LF = load factor, INFECT1= 10.000 infections. INFECT2 = 20.000 infections, INFECT3

= 100.00 infections, VL = ventilation only, NH = air conditioning no HEPA filter, HP = air conditioning with HEPA filter, 3G = 3G policy, 2G+ = 2G+ policy, AM = any mask, FFP2 = FFP2 mask mandatory, GEN = gender, HBO = HBO education level, WO = university education level, INC_{20-40} = income between €20.000 and €40.000.

4.1 Results regression

The results of linear regression give several implications. In total, eight main attributes are included in the regression model, two of them are ratio scales, and six of them are ordinal scales. All the ordinal scale variables are dummy coded. In total, there are 20 main parameters estimated. Also, five socio-demographic attributes are included that are also dummy coded, ensuring a total of 14 parameters. For the main attributes, all parameters do have the expected sign, except 3G policy (i.e., either testing, recovery or vaccination proof needed to travel). In this case, a negative sign would be expected as this policy will decrease the probability of someone infected when travelling (in comparison to the base level, with no travel requirements).

All other main parameters do have the expected sign, and there are also insignificant main parameters. The constant is 2.8; this is the value if all parameters are set to be zero. In this case all levels are set to the base; respondents rate perceived risk at 2.8 (so that is around the mean value of 3). This means that respondents, on average, rated their perceived risk with all levels set to the base, as a little under 3-medium perceived risk. All significant parameters are highly significant, except for the level '25.000 infections per day'.

Main parameters

- **Travel advice**: Both the yellow and orange advice parameters are not significant, so these levels do turn not out to be different from the base level green advice. However, red travel advice has the largest positive effect on perceived risk with a value of 0.698.
- **Vaccination rate**: This is a ratio variable with a contribution of -0.007 for every percentage point increase in vaccination rate in the country of destination. For example a vaccination rate of 50% gives the following parameter: $50 \times -0.007 = -0.35$. When travelling to a country with a vaccination rate of 90% (Portugal), perceived risk is decreased with -0.63 rating points.
- **Load factor**: This is also a ratio variable, with (almost) the same but opposite contribution of vaccination rate. The value of this parameter is 0.006. A load factor of 75% would lead to $70 \times 0.006 = 0.42$ increase in rating points. A load factor of 100% would lead to an increase of 0.6 points on perceived risk.
- **Infections**: There is some counter-intuitive outcome, as 10.000 tests per day contribute more (with a value of 0.252) to perceived risk than 25.000 tests per day (with a value of 0.123). A reason could be that respondents

find it hard to imagine what the difference in levels means. INFECT2 is also just significant (or just insignificant) on the 5% level (p-value of 0.051). The highest level, i.e., 100.000 tests per day, has the highest contribution of the dummy variables. This is in line with expectations. It also has the second-highest contribution of the attributes, with a value of 0.495.

- **Ventilation/air conditioning**: All dummy variables turn out to be significant. Only ventilation has the highest contribution to the decrease in perceived risk with a value of -0.405. This is in line with expectations as there was a focus from society on ventilation. Therefore it could be expected that people do think this is important. Air conditioning without HEPA filter has the lowest contribution of the dummies (with a value of -0.198); again, this could be expected, as air conditioning without HEPA filters has a lower level of protection against viruses than air conditioning with HEPA filters. Air conditioning with HEPA filter has a higher contribution than the previous level. The value of air conditioning with HEPA filters is -0.289.
- **Travel requirements**: As explained earlier, the first dummy variable 3G policy, has a positive sign with the value of 0.144, which is not in line with expectations. A possible explanation for this could be the fact that when first introducing the 3G policy last summer 2021, there was an exponential increase in infections. 2G policy is not significant. This can be explained as the 2G policy was never introduced, and there was a lot of resistance. Also, there effectiveness of both the 3G and 2G has been questioned and is reduced (Mouter, Hernandez, and Itten, 2021). 2G+ is an extra level of security in comparison to 2G, with people also needing to test even with a vaccination or recovery proof, it turns out to be significant. This level shows a reduction of perceived risk, with a value of -0.272.
- **Face mask policy**: Any face mask required and at least FFP2 face mask required are significant. The level 'any face mask required' has a higher contribution (value of -0.268) than 'at least FFP2' (value of -0.140). So the need to put on any face mask is more important to reduce perceived risk than having at least an FFP2 mask, according to the respondents.
- **Cleaning policy**: None of the dummy variables turned out to be significant. This shows that there is no difference from the base level 'same cleaning policy as before COVID-19' and therefore does not reduce perceived risk.

Socio-demographic attributes

- **Gender**: As in line with the expectations, gender turns out to be significant with a value of 0.088. This means that being women & 'other' increases perceived risk with 0.088 rating points. This is not a very high value in comparison to other attributes, but the value is significant, so there is a difference between men and women.
- **Education**: The level HBO and WO are both significant and positive. This means in comparison to the base (MBO or lower), people with education HBO and WO perceive the risk of COVID-19 as higher than people with

MBO or lower education level. The value for HBO is 0.137 and for WO 0.127, so people with HBO perceive risk as a bit higher than WO.

- **Income**: For this attribute, there are no expectations. The levels 'income between €20.000 and €40.000' and 'income between €40.000 and €100.000' and 'I do not want to say are significant. The first level has a value of 0.112, the second level 0.168 and the last 0.259. Having an income between €20.000 and €100.000 contributes to a higher perceived risk.
- **Age**: The age group '20-40' years contributes to a lower perceived risk (in comparison to the base 'younger than 20 years). The other two dummy variables did not turn out to be significant. So the age groups '40-65 years' and '65 years and older' do not contribute to an increase or decrease in perceived risk.
- **Work status**: All of the parameters are highly insignificant. So work status does not influence perceived risk.

Table 4 Left: Significant attributes linear regression, right: insignificant attributes

Model	Main effects			Main & Socio-demographics		
Main attributes	Value	t	p-value	Value	t	p-value
(Constant)	3.039	52.919	0.000	2.815	40.913	0.000
ADVICE3	0.698	15.412	0.000	0.702	15.620	0.000
VACC	-0.006	-12.866	0.000	-0.007	-13.203	0.000
CROW	0.006	11.306	0.000	0.006	11.604	0.000
INFECT1	0.253	5.917	0.000	0.252	5.953	0.000
INFECT2	0.136	2.134	0.033	0.123	1.954	0.051
INFECT3	0.494	12.575	0.000	0.495	12.692	0.000
AIRCO1	-0.388	-5.411	0.000	-0.405	-5.684	0.000
AIRCO2	-0.193	-4.467	0.000	-0.198	-4.617	0.000
AIRCO3	-0.286	-7.419	0.000	-0.289	-7.560	0.000
REQUIRE1	0.138	2.627	0.009	0.144	2.757	0.006
REQUIRE3	-0.267	-7.424	0.000	-0.272	-7.608	0.000
MASK1	-0.261	-5.397	0.000	-0.268	-5.576	0.000
MASK3	-0.137	-3.889	0.000	-0.140	-4.005	0.000
Socio-demographic attributes				Value	t	p-value
GENDER				0.088	3.283	0.001
HBO				0.137	4.191	0.000
WO				0.127	3.783	0.000
INCOME_20_40				0.112	2.634	0.008
INCOME_40_100				0.168	4.005	0.000
AGE_20_40				-0.095	-2.915	0.004
DO_NOT_SAY_INCOME				0.259	5.473	0.000
$R^2 = 0.129$				$R^2 = 0.143$		
Adjusted $R^2 = 0.127$				Adjusted $R^2 = 0.140$		

Model	Main effects			Main & Socio-demographics		
Main attributes	Value	t	p-value	Value	t	p-value
MASK2	-0.036	-1.536	0.125	-0.029	-1.240	0.215
CLEAN1	-0.004	-0.304	0.761	-0.002	-0.132	0.895
CLEAN2	-0.023	-1.074	0.283	-0.020	-0.920	0.358
CLEAN3	0.003	0.216	0.829	0.002	0.139	0.890
REQUIRE2	0.022	0.791	0.429	0.019	0.674	0.500
ADVICE1	-0.003	-0.126	0.900	-0.001	-0.056	0.955
ADVICE2	-0.021	-0.941	0.347	-0.020	-0.886	0.376
Socio-demographic attributes				Value	t	p-value
AGE_40_65				-0.022	-1.581	0.114
AGE_65_AND_OLDER				0.022	1.530	0.126
INCOME_MORE_100				0.025	1.374	0.170
NOT_WORKING				-0.006	-0.382	0.702
STUDENT_SCHOOL				0.008	0.584	0.559
RETIRED				-0.009	-0.709	0.479
OTHER_EDU				0.012	0.801	0.423

4.2 Results MNL model

The results of the different MNL models can be found in table 5. Looking at the results, it can be concluded that all parameters do have the expected sign. As all insignificant parameters have been removed, all remaining parameters are significant at the 5% level. As a consequence of adding the interaction between risk and time, there is no separate parameter for risk. The parameter for time is mode dependent. However, the parameter of time for plane is insignificant, so it is not included in the model.

Table 5 Base model, base model with main attribute interactions & final model

Model	Base model			Main interaction			Final model		
Parameter	Value	t	p-value	Value	t	p-value	Value	t	p-value
ASC_PLANE	-0.359	-5.3	1.16e-07	-1.51	-11.3	0	-2.23	-14.1	0
ASC_TRAIN	0.35	6.4	1.53e-10	-0.197	-2.52	0.0118	-0.414	-4.16	3.15e-05
B_COMFORT	0.281	8.7	0	0.281	8.3	0	0.346	10.4	0
B_COST	-0.00306	-21.2	0	-0.00326	-23	0	-0.00973	-8.02	1.11e-15
B_TIME	-0.167	-26.5	0						
B_RISK	-0.22	-22	0						
B_TIME_C				-0.168	-23.1	0	-0.194	-24.8	0
B_TIME_T				-0.0991	-9.42	0	-0.114	-10.1	0
B_TIME_RISK_P				-0.0437	-9.9	0	-0.011	-3.55	0.000392
B_TIME_RISK_T				-0.039	-17.1	0	-0.0107	-4.32	1.57e-05
B_GENDER_PR							-0.0217	-10.5	0
B_EDU_HBO_PR							-0.0142	-4.26	2.05e-05
B_EDU_WO_PR							-0.0237	-6.63	3.27e-11
B_EDU_OTHER_PR							-0.018	-4.99	6.05e-07
B_AGE_COST_20_40							0.00325	2.6	0.00932
B_AGE_COST_40_65							0.0064	5.18	2.23e-07
B_AGE_COST_65_AND_OLDER							0.00723	5.88	4.17e-09
B_PAYMENT_WORKEDU_COST							0.00223	3.18	0.00146
B_COMPANY_PR_FRIENDS							-0.00746	-2.62	0.00878
B_COMPANY_PR_OTHER							0.0197	4.06	4.86e-05
B_PURPOSE_WORK_TIME							-0.0596	-3.98	6.82e-05
B_PREF_CAR_C							0.496	6.3	2.97e-10
B_PREF_CAR_T							-0.268	-3.33	0.000881
B_PREF_PLANE_P							1.04	14.5	0
B_PREF_TRAIN_C							-0.184	-2.43	0.0152
B_PREF_TRAIN_T							0.404	5.49	3.99e-08
ρ^2	0.126			0.133			0.24		

- ASC plane and ASC train are both negative, so car is more preferred.
- Beta comfort is positive, beta time and beta cost and beta time are both negative. Travel time is more negative for car than for train. For plane, time is insignificant.
- Perceived risk is significant as an interaction with time. This means that for longer travel times, COVID-19 risk is perceived higher. Perceived risk is also negative and different for train and plan; however, there is a very small difference, see table 5.
- Several interactions are significant with perceived, also on the ASCs and travel cost.
- A preference for car has a positive contribution to the ASC of car and a small negative contribution to train. A preference for train has a small negative contribution on the ASC of car and a positive contribution to the ASC of train. A preference for plane has a positive contribution to the ASC of plane.
- For travel time, only purpose work is significant. Respondents that travel for work are more sensitive to travel time.
- For travel cost, both age and payment by work or educational institution are significant. Higher age results in less weight to travel cost, the same counts if the trip is payed by work the educational institution.
- For perceived risk, educational level and gender are significant. Having at least HBO or WO education (in comparison to MBO) results in more weight to perceived risk. Women have a higher weight to perceived risk as well.
- Increasing age results in a lower weight to travel cost.

4.3 Interpretation of parameters

Several interactions contribute to the different main parameters. It is highlighted that the WtP for perceived risk, particularly in terms of travel costs, can vary significantly across people of varied ages, payment, education level and travel companies, and possibly other unobserved background variables not represented in this study. Therefore, the average respondent is used, based on the average (and when not possible) on the most common value, the values are shown in table 6.

Table 6 The average respondent

Age	38
Education	HBO
Gender	Men & women
Trip purpose	Leisure
Travel company	Alone

The 'value of risk' (VoR) is the trade-off between perceived risk and travel cost.Using the average respondent values, the VoR values is calculated in equation 2 & 3.

$$VoR_{in\ travel\ cost} = \frac{\frac{\delta U}{\delta PR}}{\frac{\delta U}{\delta TC}} \tag{1}$$

$$VoR_{men} = \frac{-0.011 - 0.142}{-0.00973 + 0.00325} = 3.89\ euro \tag{2}$$

$$VoR_{women} = \frac{-0.011 - 0.142 - 0.0217}{-0.00973 + 0.00325} = 7.24\ euro \tag{3}$$

The VoR for men is €3.89 per level of perceived risk per hour; for women this is €7.24. For women this is higher due to the interaction between gender and perceived risk. For higher travel times and higher perceived risk levels, this value becomes equivalent 3.89/7.24*TT*PR. VoR is thus linear, there are no quadratic components.

The trade-off between travel comfort and perceived risk does have to following results. The VoR for men in comfort 'points' is 0.072 (per hour); for women, this is 0.134.

$$VoR_{in\ travel\ comfort} = \frac{\frac{\delta U}{\delta PR}}{\frac{\delta U}{\delta CF}} \tag{4}$$

$$VoR_{in\ comfort,men} = \frac{-0.011 - 0.142}{0.346} = 0.072 \tag{5}$$

$$VoR_{in\ comfort,women} = \frac{-0.011 - 0.142 - 0.0217}{0.346} = 0.134 \tag{6}$$

As perceived risk is dependent of time, it is also possible to express what comfort is 'worth' in terms of travel time for different perceived risk levels. The trade-off (TO) for a full reduction/increase in class, this means one full comfort 'point' (so 1^{st} class ↔ 2^{nd} class or business ↔ economy) is shown in table 7.

Table 7 Trade off comfort travel time

Men			Women		
PR level	Comfort [points]	TT [hours]	PR level	Comfort [points]	TT [hours]
1	0.072	13.889	1	0.134	7.463
2	0.144	6.944	2	0.268	3.731
3	0.216	4.630	3	0.402	2.488
4	0.288	3.472	4	0.536	1.866
5	0.36	2.778	5	0.670	1.493

The value of time in transportation (economics) is the potential cost of the time that a passenger spends on their route. In essence, this is the amount a passenger is ready to pay to save time, or the amount they would take as compensation for lost time. The amount of time that passengers will save is one of the key justifications for transportation upgrades (Kouwenhoven et al., 2014). The VoT for train has a value of €17.59 for train and €29.94 for car.

$$VoT = \frac{\frac{\delta U}{\delta TT}}{\frac{\delta U}{\delta TC}} \tag{7}$$

$$VoT_{train} = \frac{-0.1140}{-0.0097 + 0.0033} = 17.59\ euro/hour \tag{8}$$

$$VoT_{car} = \frac{-0.194}{-0.0097 + 0.0033} = 29.94\ euro/hour \tag{9}$$

The different VoT values found for this thesis in comparison to other studies is shown in table 8. The fact that car has a higher VoT is in line with KiM *Netherlands Institute for Transport Policy Analysis* (KiM, 2020). The parameter of time for plane is not significant; hence, no VoT can be specified for plane. In this study, the VoT for car is just over 2 euros more. For Kouwenhoven et al. (2014) there is a very small difference found between the modes. Shires and De Jong (2009) did not find differences in VoT for train and car. It must be noted that both KiM (2020) and Kouwenhoven et al. (2014) focused on short distance travel, while Shires and De Jong (2009) focused on long-distance. The average VoT found in this paper is almost the same as found in Shires and De Jong (2009).

Table 8 VoT results of different studies

	This paper	KiM (2020)	Van Kouwenhoven (2014)	Shires and De Jong, (2009)
VoT Train	17.59	13.22	9.25	24.00
VoT Car	29.94	15.58	9.00	
Average	23.77	14.40	9.13	24.00

Willingness to pay for comfort can also be interpreted from the result. The equation for WtP for comfort for both train and plane can be found in equation 11.

$$WtP_{comfort} = \frac{\frac{\delta U}{\delta CF}}{\frac{\delta U}{\delta TC}} \tag{10}$$

$$WtP_{comfort} = \frac{0.346}{-0.0097 + 0.0033} = 53.40\ euro/class \tag{11}$$

The result show the willingness to pay for an upgrade in class (so 2^{nd} class $\rightarrow$ 1^{st} class or economy $\rightarrow$ business) is €53.40. No mode-specific beta turned out to be significant, so there is no distinction between classes in train and plane. Balcombe, Fraser, and Harris (2009) found a value of about €120, but they agreed that this value is on the high side. A big amount of the passengers travelling in business or 1^{st} class is flying for business purposes (BusinessAM, 2020). It is therefore of interest to show the willingness to pay for business/educational travellers (respondents who's trip is payed by company of educational institution), the value is shown in equation 12. This shows that the WtP for an upgrade is substantially higher.

$$WtP_{comfort} = \frac{0.346}{-0.0097 + 0.0033 + 0.00222} = 81.22\ euro/class \tag{12}$$

4.4 Combination results both experiment

Both the perceived risk rating experiment and the main (mode) choice experiment are estimated, so now the results of both experiments can be estimated. Within the perceived risk rating experiment, perceived risk was the dependent variable; at the same time, it was an independent variable in the main (mode) choice experiment. For the rating experiment, gender and education level have a positive contribution to perceived risk. Both experiments can be combined by using the (absolute) linear regression coefficients of the perceived risk rating experiment. Then these values are combined with the WtP values that are just calculated. Because of the dummy coding, all different dummy variables have an independent contribution to the value of perceived risk. The results are shown in table 9. Respondents are willing to most for a decrease in level of perceived when there is a risk red travel advice. In comparison to the base (no travel advice), men are willing to pay €32.74 for a decrease in level of risk when travelling for 12 hours; women €60.93. If men are travelling for only 6 hours, this value is half of the 12-hour value, so €16.37; for women, €30.47. Moreover, there is a high WtP when there is a 100% load factor (men €30.17 and women €56.15). At last, 100.000 infections result in a WtP for decrease of one level of perceived risk of €23.09 for men and €42.96 for men. The big difference between men and women is due to the significant interaction between gender and perceived risk.

Table 9 WtP values for the different risk factors

			Men				Women			
		PR level difference	$\Delta_{level} = -1$				$\Delta_{level} = -1$			
		Travel time	3	6	9	12	3	6	9	12
		Value	€11.67	€23.33	€35.00	€46.67	€21.71	€43.43	€65.14	€86.85
Travel advice	Red travel advice	0.702	€8.18	€16.37	€24.55	€32.74	€15.23	€30.47	€45.70	€60.93
Vaccination rate	Parameter	-0.007	€-0.08	€-0.15	€-0.23	€-0.30	€-0.14	€-0.28	€-0.42	€-0.57
	15	-0.098	€-1.14	€-2.28	€-3.42	€-4.55	€-2.12	€-4.24	€-6.36	€-8.48
	30	-0.195	€-2.28	€-4.55	€-6.83	€-9.11	€-4.24	€-8.48	€-12.72	€-16.95
	70	-0.455	€-5.31	€-10.63	€-15.94	€-21.26	€-9.89	€-19.78	€-29.67	€-39.56
	90	-0.586	€-6.83	€-13.66	€-20.50	€-27.33	€-12.72	€-25.43	€-38.15	€-50.86
Load factor	Parameter	0.006	€0.08	€0.15	€0.23	€0.30	€0.14	€0.28	€0.42	€0.56
	25	0.162	€1.89	€3.77	€5.66	€7.54	€3.51	€7.02	€10.53	€14.04
	50	0.323	€3.77	€7.54	€11.31	€15.08	€7.02	€14.04	€21.06	€28.07
	75	0.485	€5.66	€11.31	€16.97	€22.63	€10.53	€21.06	€31.58	€42.11
	100	0.646	€7.54	€15.08	€22.63	€30.17	€14.04	€28.07	€42.11	€56.15
Infection rate	10.000 infections per day	0.252	€2.94	€5.88	€8.82	€11.76	€5.47	€10.95	€16.42	€21.90
	25.000 infections per day	0.123	€1.44	€2.88	€4.32	€5.76	€2.68	€5.36	€8.04	€10.72
	100.000 infections per day	0.495	€5.77	€11.54	€17.31	€23.09	€10.74	€21.48	€32.22	€42.96
Air conditioning	Only ventilation	-0.405	€-4.72	€-9.44	€-14.17	€-18.89	€-8.79	€-17.58	€-26.36	€-35.15
	Airco no HEPA	-0.198	€-2.31	€-4.62	€-6.92	€-9.23	€-4.29	€-8.59	€-12.88	€-17.18
	Airco with HEPA	-0.289	€-3.38	€-6.75	€-10.13	€-13.51	€-6.28	€-12.57	€-18.85	-25.14
Travel requirements	3G-policy	0.144	€1.68	€3.36	€5.03	€6.71	€3.12	€6.25	€9.37	€12.49
	2G+-policy	-0.272	€-3.17	€-6.35	€-9.52	€-12.70	€-5.91	€-11.81	€-17.72	€-23.63
Face mask policy	Any face mask	-0.268	€-3.12	€-6.24	€-9.37	€-12.49	€-5.81	€-11.62	€-17.43	-23.25
	At least FFP2	-0.140	€-1.64	€-3.27	€-4.91	€-6.55	€-3.05	€-6.09	€-9.14	€-12.18
	Socio-demographic									
Gender	Women	0.088	€1.03	€2.06	€3.09	€4.12	€1.92	€3.83	€5.75	€7.66
Education	HBO	0.137	€1.60	€3.21	€4.81	€6.41	€2.98	€5.97	€8.95	€11.94
Income class	Income €20.000 to €40.000	0.112	€1.31	€2.62	€3.92	€5.23	€2.43	€4.87	€7.30	€9.74
	Income €40.000 to €100.000	0.168	€1.96	€3.92	€5.88	€7.85	€3.65	€7.30	€10.95	€14.60
Age	Age 20 to 40 years	-0.095	€-1.11	€-2.22	€-3.32	€-4.43	€-2.06	€-4.12	€-6.18	€-8.25

5 Influence perceived risk on market share

This research takes a broad approach to long-distance travel within Europe. To give a clear example of what the different implications from the results mean for real-life examples, it is chosen to go for three different cases. First, a short route that is popular by train will be used, this is the Amsterdam - London route. At the same time, the plane and car options are viable and popular as well. The second route is Amsterdam - Berlin. This route has all three modes as viable options. This route is popular by train as there is a direct connection. At third, a longer route is chosen to see the difference between a shorter and longer routes. In order to look at the influence of perceived risk on the modal-split, real-world values for travel time, travel cost and comfort are being used for these routes. Then, the perceived risk levels are varied so that the influence can be discussed.

Amsterdam - London: For this route, the maximum market share differences are as follows. When perceived risk in the train increases from 1 to 5 and the plane stays at level 1, train loses 10% point market share for men; for women this is 17%. Plane has 3% point decrease in market share for men; for women this is 5% point. Car increase by 7% for men; for women this is 12% point. This is because train has a longer travel time. Women have a higher weight to perceived risk, therefore market shares differences are higher for women. This is due to the fact that plane has a shorter travel time, for car this is because there is no perceived risk. For smaller differences in perceived risk, these market share changes are even smaller. The perceived risk levels are complex concepts, its score is determined by a lot of factors. Therefore there is no real such thing as risk level 1. So, these extreme changes are not realistic in real life. Moreover, 4 out of the 8 factors are destination specific. In this case, these factors do not change between modes for the same OD pair. To make this clear, the maximum difference in perceived risk points for the same

OD-pair will be (only considering mode-related attributes):

$$PR_{low} = 2.815 - 0.405 - 0.268 = 2.14$$
$$PR_{high} = 2.815 + (0.006 * 100) = 3.46$$
$$\Delta PR = 3.46 - 2.14 = 1.32$$

In this case, when the perceived risk of train is 2 and plane 3 (to approximate risk point 3.46 and 2.14) to train 1 and plane 3, there is only a 1 to 2% point market share difference for men; for women this is 2 to 3% point. Thus, it can be concluded that the influence of perceived risk is very moderate for this route. The change for the different factors that contribute to perceived risk is even smaller.

Amsterdam - Berlin: The maximum market share differences for this route are a bit higher. When perceived risk in train is increasing from 1 to 5 while plane stays at 1, results in a decrease of market share for train of 16% point for men; for women this is 27% point. Plane increases 6% point for men; for women it increase by 10% point. Car increases by 10% point for men; for women this is 17% point. As there is only 1.32 perceived risk level difference possible, the maximum market share difference possible is around 1-4% point for men; for women this is 2-7% point.

Amsterdam - Barcelona: For this route, perceived risk has a bigger influence on the market shares. This could be expected as perceived risk is dependent on time. The maximum market share difference for train is 21% point decrease for men; for women this is 29% point. For men plane increases by 17% point; for women this is 24% point. Car increases by 4% point for men; for women this is 5% point. However, as there is only 1.32 perceived risk level difference, the realistic effect of perceived risk is a maximum of 4-6% point for men; for women 3-16% point.

With the 1.32 perceived risk difference, market share differences are at maximum. For the three routes, the maximum possible market share difference show for London a small effect of perceived risk. For the Amsterdam - Berlin and Amsterdam - Barcelona routes, this effect becomes larger. However, the 1.32 is the maximum possible difference. For real-life trips, not all of the four mode-related attribute will be different. So in this case, the market share difference is smaller.

6 Conclusion

To research the effects of perceived COVID-19 infection risk, this study was conducted. In total, 1147 respondents took part in this study. To study whether perceived COVID-19 risk influences mode choice, the survey consisted of two main parts, a part about socio-demographics and a part about travel behaviour. The first part of the survey was the rating experiment. In this part, respondents had to rate their perceived risk of infection with COVID-19 due to several factors; four of these factors were destination-related, and the other four were mode-related. After this, respondents were faced with the main (mode) choice experiment. In this experiment, respondents had to choose between train, plane and car based on travel time, travel cost, travel comfort (class of travel) and perceived risk (now a given value). To analyse the data, the discrete choice modelling theory was used. At last, an adapted variant of the Hierarchical Information Integration (HII) theory was used to combine both of the results.

In total, eight factors were established: travel advice, vaccination rate, load factor, infections, air conditioning, travel requirements, face mask policy and cleaning policy. Red travel advice is the most important factor in the contribution to perceived risk. Vaccination rate and load factor do have about the same contribution to perceived risk and are the second most important predictors for perceived risk, but the signs are opposite. Vaccination rate has a negative sign, load factor a positive sign. The attribute of 100.000 positive tests has the highest contribution to the number of infections. For the ventilation/air conditioning variables, ventilation has the highest negative contribution. This can be explained as there was a focus on ventilation by the government to reduce the number of infections. Regarding travel requirements, the 3G policy has a positive contribution to perceived risk, which is counter-intuitive. This could be explained, as with the (first) introduction of the 3G-policy, there was a dramatic increase in number of infections. The 2G policy has no effects, but the 2G+ policy does decrease perceived risk. When looking at the face mask policy, having any mask is more important in the decrease of perceived risk than having at least an FFP2 mask. The cleaning policies do not contribute to decrease perceived risk. Regarding socio-demographic values, the following conclusion can be taken. For this question, gender, age, income, education and work status were added to the model. Work status does not have a contribution to perceived risk and thus does not influence perceived risk. Gender has a positive contribution to perceived risk; hence, being woman results in a higher (average) rating for perceived risk. The same counts for education levels' HBO' and 'WO'. Moreover, having an income of €20.000 and €100.000 per year also contributes to perceived risk. When looking at age, most of the age groups did not have a contribution to perceived risk; however, being between 20 and 40 years old leads to a decrease in perceived risk.

Travel cost becomes less important with increasing age; this means that the cost parameter becomes less negative with increasing age. Moreover, if the payment is made by the educational institute or work, people become less sensitive to cost as well. For travel time, only travel purpose has an influence. Travelling for work contributes to a more negative weight on travel time, so travel time becomes more important. For travel comfort, no significant interactions were found. At last, for perceived risk, both gender and travel company turned out to be significant. Being a woman contributes to a higher weight to perceived risk; this was also found in the rating experiment. At last, travelling with friends contributes to perceived risk.

The value of risk for a decrease in the level of risk for a man is €3.89 per decrease in level per hour; for a woman, this is €7.24 per decrease in level per hour. The gender difference is found due to a significant interaction between perceived risk and gender. The levels of risk go from 1-very low to 5-very high. Perceived risk depends on time; therefore, there is no trade-off in travel time. Longer travel time leads to a higher perceived risk. It is noted that this is for the average respondent. Different values are for different ages, travel purposes, who pays for the trips and education levels. In terms of travel comfort, the value of risk for men is 0.072 comfort 'points' for one level decrease of perceived risk per hour; for women, this is 0.134. An upgrade in travel class is worth 13.89 hours for men when perceived risk level is one; for women, this is 7.46 hours. For perceived risk level 5, a full increase in travel class is 'worth' only 2.78 hours of travel for men; for women, 1.49 hours. The value of time found is 17.59 euro/hour for train, and 29.94 euro/hour for car. The willingness to pay for an increase in travel class is 53.40 euros. When work or the educational institution pays for the trip, the WtP becomes 81.22 euros.

For the three routes, the maximum possible market share difference shows a small effect on perceived risk for London. This effect becomes more significant for the Amsterdam - Berlin and Amsterdam - Barcelona routes. However, as there is only a 1.32 perceived risk level difference, the realistic effect of perceived risk is only a maximum of 2-3% market share difference for London; however, for Barcelona this increase to 16% point. However, in this case, all four mode-related attributes have to be at the lowest level for plane and have to be at the highest value for train. Realistically, the differences are smaller, so this market share difference will be smaller as well. **It can thus be concluded that risk perception of infection with COVID-19 has a moderate influence on market share.**

7 Limitations

The rating experiment is based on eight factors that together contribute to perceived risk. However, for respondents, other attributes could be important as well. The usage of focus groups could improve the composition of perceived risk.

Travel cost, travel time (and comfort to a lesser extent) are the most important variables for mode choice. Because of the main choice experimental set-up, with perceived risk next to the other main choice attributes, travel cost, travel time and travel comfort, an overestimation of the importance of perceived risk is possible.

Another limitation of the research is that it did not include an opting-out option. COVID-19 has several influences on travel behaviour.

The biggest limitation is the usage of the MNL model, as the MNL model has some significant shortcomings. A Mixed Logit model is better as it accommodates for the limitations of the MNL model.

The usage of the NS panel to recruit respondents, gives a bias to the sample. The market shares for train are very high in comparison to real world examples.

The rating of perceived risk was an integer, on a 5-points scale, based on the chosen attributes. The experiment is on an ordinal scale, whereas the rating is estimated on a ratio scale. There is some debate about using linear regression to estimate the rating because the respondents gave the answers on an ordinal scale.

References

Aaditya, B., & Rahul, T. (2021). Psychological impacts of covid-19 pandemic on the mode choice behaviour: A hybrid choice modelling approach. *Transport Policy*, *108*, 47-58. Retrieved from https://www.sciencedirect.com/science/article/pii/S0967070X21001359

Aamaas, B., Borken-Kleefeld, J., Peters, G.P. (2013). The climate impact of travel behavior: A german case study with illustrative mitigation options. *Environmental Science & Policy*, *33*, 273-282. Retrieved from https://www.sciencedirect.com/science/article/pii/S1462901113001366

Abdullah, M., Dias, C., Muley, D., Shahin, M. (2020). Exploring the impacts of covid-19 on travel behavior and mode preferences. *Transportation Research Interdisciplinary Perspectives*, *8*, 100255. Retrieved from https://www.sciencedirect.com/science/article/pii/S2590198220301664

Atasoy, B., Glerum, A., Bierlaire, M. (2013). Attitudes towards mode choice in switzerland. *disP - The Planning Review*, *49*(2), 101-117. Retrieved from https://doi.org/10.1080/02513625.2013.827518

Balcombe, K., Fraser, I., Harris, L. (2009). Consumer willingness to pay for in-flight service and comfort levels: A choice experiment. *Journal of Air Transport Management*, *15*(5), 221–226.

Buehler, R. (2011). Determinants of transport mode choice: a comparison of germany and the usa. *Journal of Transport Geography*, *19*(4), 644-657. Retrieved from https://www.sciencedirect.com/science/article/pii/S0966692310001171

Buehler, R., & Nobis, C. (2010). Travel behavior in aging societies: Comparison of germany and the united states. *Transportation research record*, *2182*(1), 62–70.

BusinessAM (2020). *Wie business class vliegt, waant zich nu in economy.* Retrieved from https://businessam.be/wie-business-class-vliegt-waant-zich-nu-in-economy/

Chen, J.-M. (2022). Novel statistics predict the covid-19 pandemic could terminate in 2022. *Journal of Medical Virology*.

de Haas, M., Faber, R., Hamersma, M. (2020). How covid-19 and the dutch 'intelligent lockdown' change activities, work and travel behaviour: Evidence from longitudinal data in the netherlands. *Transportation Research Interdisciplinary Perspectives*, *6*, 100150. Retrieved from https://www.sciencedirect.com/science/article/pii/S2590198220300610

De Vos, J. (2020). The effect of covid-19 and subsequent social distancing on travel behavior. *Transportation Research Interdisciplinary Perspectives*, *5*, 100121. Retrieved from https://www.sciencedirect.com/science/article/pii/S2590198220300324

De Witte, A., Hollevoet, J., Dobruszkes, F., Hubert, M., Macharis, C. (2013). Linking modal choice to motility: A comprehensive review. *Transportation Research Part A: Policy and Practice*, *49*, 329-341. Retrieved from https://www.sciencedirect.com/science/article/pii/S0965856413000165

Dobruszkes, F., Dehon, C., Givoni, M. (2014). Does european high-speed rail affect the current level of air services? an eu-wide analysis. *Transportation Research Part A: Policy and Practice*, *69*, 461-475. Retrieved from https://www.sciencedirect.com/science/article/pii/S0965856414002158

Dryhurst, S., Schneider, C.R., Kerr, J., Freeman, A.L., Recchia, G., Van Der Bles, A.M., ... Van Der Linden, S. (2020). Risk perceptions of covid-19 around the world. *Journal of Risk Research*, *23*(7-8), 994–1006.

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*, *41*(5), 428–443.

Johansson, M.V., Heldt, T., Johansson, P. (2006). The effects of attitudes and personality traits on mode choice. *Transportation Research Part A: Policy and Practice*, *40*(6), 507–525.

Joshi, A., Kale, S., Chandel, S., Pal, D.K. (2015). Likert scale: Explored and explained. *British journal of applied science & technology*, *7*(4), 396.

Karlsson, L.C., Soveri, A., Lewandowsky, S., Karlsson, L., Karlsson, H., Nolvi, S., ... Antfolk, J. (2021). Fearing the disease or the vaccine: The case of covid-19. *Personality and Individual Differences*, *172*, 110590. Retrieved from https://www.sciencedirect.com/science/article/pii/S0191886920307819

KiM (2020). *Kengetallen bereikbaarheid.* Retrieved from https://www.rwseconomie.nl/kengetallen/kengetallen-bereikbaarheid-map

Kouwenhoven, M., de Jong, G.C., Koster, P., van den Berg, V.A., Verhoef, E.T., Bates, J., Warffemius, P.M. (2014). New values of time and reliability in passenger transport in the netherlands. *Research in Transportation Economics*, *47*, 37–49.

Krishnakumari, P., & Cats, O. (2020, 05). *Virus spreading in public transport networks.*

Leppin, A., & Aro, A.R. (2009). Risk perceptions related to sars and avian influenza: theoretical foundations of current empirical research. *International journal of behavioral medicine*, *16*(1), 7–29.

Loewenstein, G.F., Weber, E.U., Hsee, C.K., Welch, N. (2001). Risk as feelings. *Psychological bulletin*, *127*(2), 267.

Louviere, J.J. (1984). Hierarchical information integration: A new method for the design and analysis of complex multlattribute judgment problems. *ACR North American Advances*.

Matiza, T. (2020). Post-covid-19 crisis travel behaviour: towards mitigating the effects of perceived risk. *Journal of Tourism Futures*.

Mertens, G., Gerritsen, L., Duijndam, S., Salemink, E., Engelhard, I.M. (2020). Fear of the coronavirus (covid-19): Predictors in an online study conducted in march 2020. *Journal of anxiety disorders*, *74*, 102258.

Molin, E. (2020). *Complex variables & non-linearity.* Retrieved from https://brightspace.tudelft.nl/d2l/le/content/195054/viewContent/1371927/View

Molin, E., & van Gelder, M. (2008). Freeway access to public transport: A hierarchical multimodal choice model. *Transportation Research Record*, *2076*(1), 106-113. Retrieved from https://doi.org/10.3141/2076-12 https://arxiv.org/abs/https://doi.org/10.3141/2076-12

Morikawa, T., Ben-Akiva, M., McFadden, D. (2002). Discrete choice models incorporating revealed preferences and psychometric data. *Advances in econometrics.* Emerald Group Publishing Limited.

Mouter, N., Hernandez, J.I., Itten, A.V. (2021). Public participation in crisis policymaking. how 30,000 dutch citizens advised their government on relaxing covid-19 lockdown measures. *PloS one*, *16*(5), e0250614.

Nieto-García, M., Muñoz-Gallego, P.A., Gonzalez-Benito, Ó. (2020). The more the merrier? understanding how travel frequency shapes willingness to pay. *Cornell Hospitality Quarterly*, *61*(4), 401–415.

Paulssen, M., Temme, D., Vij, A., Walker, J.L. (2014). Values, attitudes and travel behavior: a hierarchical latent variable mixed logit model of travel mode choice. *Transportation*, *41*(4), 873–888.

Rijksoverheid (2021). *Reisadviezen.* Retrieved from https://www.nederlandwereldwijd.nl/reizen/reisadviezen

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*(1), 84-108. Retrieved from https://www.sciencedirect.com/science/article/pii/S1755534513700307

Shamshiripour, A., Rahimi, E., Shabanpour, R., Mohammadian, A.K. (2020). How is covid-19 reshaping activity-travel behavior? evidence from a comprehensive survey in chicago. *Transportation Research Interdisciplinary Perspectives*, *7*, 100216. Retrieved from https://www.sciencedirect.com/science/article/pii/S2590198220301275

Shires, J.D., & De Jong, G.C. (2009). An international meta-analysis of values of travel time savings. *Evaluation and program planning*, *32*(4), 315–325.

Tirachini, A., & Cats, O. (2020). Covid-19 and public transportation: Current assessment, prospects, and research needs. *Journal of Public Transportation*, *22*(1), 1.

Troko, J., Myles, P., Gibson, J., Hashim, A., Enstone, J., Kingdon, S., . . . Van-Tam, J.N. (2011). Is public transport a risk factor for acute respiratory infection? *BMC infectious diseases*, *11*(1), 1–6.

van Goeverden, C.D. (2009). Explaining factors for train use in european long-distance travel. *Tourism and Hospitality Planning & Development*, *6*(1), 21–37.

Van Loon, R., & Rouwendal, J. (2013). Vacation behaviour: Frequency, destination choice and expenditures. *Tourism Economics*, *19*(5), 1005–1026.

Velavan, T.P., & Meyer, C.G. (2020). The covid-19 epidemic. *Tropical medicine & international health*, *25*(3), 278.