DESIGNING STATED CHOICE EXPERIMENTS: AN ANALYSISOF THE EFFECTS OF AIRPORT LOCATION ON AIR PASSENGER'S TRAVEL CHOICE

Master Thesis
"Transport, Infrastructure & Logistics"

June 2008

Reinier Beelaerts van Blokland

TU Delft
Delft University of Technology

ITLS

DESIGNING STATED CHOICE EXPERIMENTS: AN ANALYSIS OF THE EFFECTS OF AIRPORT LOCATION ON AIR PASSENGER'S TRAVEL CHOICE

Reinier E.F. Beelaerts van Blokland – 1106309

Delft University of Technology | Faculty of Civil Engineering – Faculty of TPM

The University of Sydney | Institute of Transport & Logistics Studies (ITLS)

Email – beelaerts@gmail.com

June 2008

Committee

Chair

Professor Piet H.L Bovy	Delft University of Technology	Faculty of Civil Engineering

Supervisors

Dr. Michiel C.J. Bliemer	Delft University of Technology	Faculty of Civil Engineering
Dr. Eric J.E. Molin	Delft University of Technology	Faculty of TPM
Dr. John M. Rose	The University of Sydney	ITLS
Prof.mr.dr.ir. Sicco C. Santema	Delft University of Technology	Faculty of Aerospace Engineering

Keywords

Air travellers, Discrete choice modelling, Airlines, Airports, Stated choice experiment, Experimental designs

Summary

Since major airports are reaching capacity, some solutions have to be found to solve the actual transportation problem. The government sees the development of secondary airports as a solution to this problem. First, this would result in air traffic diversion and secondly this would permit new economical opportunities to the concerned region. However, airlines and air passengers are the airport users, and in order to make secondary airports economically profitable, the attractiveness of these airports plays an important role. Therefore, getting more insight in the effects of airport location on air passengers' choice would be helpful for the decision process of implementing or not a secondary airport. Additionally, investigating air travellers' preferences may assist airports and airlines to adjust their strategy on air travellers' needs in order to gain a larger market share.

In order to achieve this objective, using discrete choice modelling and more specifically stated choice (SC) experiments, is a proven way for estimating and forecasting travellers' behaviour. Since the reliability of the estimated parameters can be relying on the estimating models used or on the underlying experimental designs, more and more researches realise that so-called efficient design are able to produce more efficient data in the sense that more reliable parameter estimates can be obtained. Because this new knowledge has never been applied on practical cases, the aim of this study is to perform a SC experiment and to reach the following objectives: (i) determining which factors are influencing air passengers' choice in selecting flight itineraries, and (ii) proving that efficient designs result in more reliable parameter estimates than orthogonal designs.

A SC experiment is defined as presenting a sample of respondents with a number of hypothetical scenarios, consisting of universal but finite number of alternatives that differ on a number of attribute dimensions. In this research the alternatives consist of trips from Amsterdam Airport Schiphol (AAS) to Barcelona city. Since the city of Barcelona has two airports, namely Barcelona airport and the further away located Girona airport, travellers may take into account the airport location in their itinerary choice. In each choice situation, respondents will have to choose one alternative between 5 proposed. Each alternative is composed of six attributes: airline, price, transfer time, departure time, egress time, and egress time. Generally, the chosen alternative is the one that maximizes the traveller's utility. In total each respondent will face 6 different choice situations. Obviously, there are many ways of assigning combinations of attribute levels to the attributes in each situation. The purpose of the underlying experimental design is to determine these combinations in the best possible way. The different methods for constructing experimental designs yield the generation of three designs, namely an orthogonal design with 108 choice situations, an efficient design with 18 choice situations, and another efficient design with 108 choice situations. Collecting data can be done by several ways, such as pen & paper surveys, CAPI surveys, or internet surveys. In this research, an internet survey has been performed because of its advantages compared to other methods, looking similarly to the real website of an online travel agent (e.g. ebookers). TeamVier, a Dutch market research company, provided a heterogeneous sample of respondents for this study. In the end, the collected data consist in 3,300 observations collected from 550 respondents.

The collected data has been estimated with two different estimation models. The first model is the well-known multinomial logit (MNL) model and the second model is the more advanced panel mixed logit (ML) model. The results of both models yield better model fit of the panel ML model regarding to the MNL model, which leads in better estimations of the parameter estimates. With the outcomes of this model, statements can be done regrading to travellers' behaviour in itinerary choices. The outcomes of the analyses have shown that price and transfer time are the most important attributes influencing travellers' itinerary choice. The willingness to pay of travellers to avoid one hour of transfer time has been estimated on approximately €37. Additionally, the airport location plays an important role in travellers' itinerary choice. Due to the less egress time, travellers have a preference for arriving at the main airport instead of at the further away located secondary airport. In this study case, the egress time from Girona airport is about 40 to 60 minutes longer than from Barcelona airport, therefore people are willing to pay about €20 to travel to Barcelona instead of Girona.

When comparing the data coming from the different experimental designs, the outcomes of the analyses show that efficient designs result in more reliable parameter estimates than orthogonal designs. Additionally, the results have shown that efficient designs permit the reduction of required sample size to produce a fixed level of reliability in the parameter estimates. This leads to the fact that, designs with a limited number of choice situations can be as efficient as larger designs.

From now on, it is recommended to construct an efficient experimental design instead of an orthogonal design as it leads to more reliable parameter estimates and necessities a lower sample size, which is the most expensive part when performing a SC experiment. The statements on travellers' preferences in itinerary choice can be used by airports and airlines to adapt their strategies on the consumer needs. Although travellers have a preference for arriving at the main airport, this study has shown that the disadvantages of the location of secondary airports can easily be compensate with better price offers, which make them even more attractive than main airports. Developing regional airports could be a good alternative to solve the capacity problems of main airports. However, because this study focussed on the effects of the arriving airport location, further research is required on travellers having business proposes or on the influence of access time and access prices on local travellers' behaviour.

Preface

This report is the result of my master thesis in Transport, Infrastructure and Logistics. The last eight months have been a very educational/learning period, both personal and professional. On the educational side, it was the first time an area completely gained my interest u to the point I was overwhelmed by passion on the subject and which will certainly continue to keep my interest for its further development. On the personal side, these last five months have been the best time of my student life.

Of course I would like to thank many people who have contributed to my achievement. First I would like to thank the Institute of Transport Studies in Sydney and more particularly John Rose for giving me the great opportunity to conduct my research at the institute. John, thank you so much for your guidance, your enthusiasm, and for all your patience. Thanks to you I enjoyed my research from the beginning till the end and I could have spent even more time on it. Only now I can say I know how to use Excel, Ngene, Nlogit, and even Access. Besides that, I discovered a beautiful country which is Australia and had an "awesome" time in Sydney. After staying there for more than four months, it was difficult for me to return to Europe. Who knows, maybe I'll come back soon!

From Delft University of Technology, I would like to thank Michiel Bliemer as first. Michiel, alias Mike, without you, all what happened this last year could never have been realised. I would like to thank you for all what you have done for me. Thanks for your patience during all these months, your tips, and the good times in Sydney. Together with John, you have been the key elements in the accomplishment of my work. I cannot be more grateful to you, you are an excellent teacher and hope you will attain your objectives.

Also, I would like to thank Professor Bovy, Eric Molin and Professor Santema for their support and valuable comments at each state of my research. Having been away for the half of the time did not facilitate your accompaniments. Despite the circumstances, and the fact that I have been a virtual graduating student for Professor Santema, your accompaniments have been a great contribution.

Finally, I would like to thank Jun Zhang for all his work in coding the internet survey and everyone that I did not mention above and who supported me from the beginning of this exiting experience which I will never forget.

Reinier Beelaerts | June 2008

Table of Contents

List of Figures

List of Tables

Terminology & Definitions

AAS:	Amsterdam Airport Schiphol
Air service:	additional facilities offered by an airline within a specific flight
Airline:	a commercial enterprise that provides scheduled flights for passengers and air transport services
Airport location	distance between the arriving airport and the final destination
Alternatives:	options containing specified levels of attributes
Attribute balanced design:	a design in which the levels of any given attribute appear the same number of times as all other levels for that particular attribute
Attribute:	characteristics of an alternative
Attribute levels:	a specific value taken by an attribute
Attribute-level label:	the narrative description corresponding to an attribute
Blocking:	the use of an additional design column to assign sub-sets of treatment combinations to decision makers
Choice set:	the set of alternatives over which an agent makes a choice
Coding:	the use of numbers to designate a particular state of an attribute
Degrees of freedom:	the number of observations in a sample minus the number of independent constraints imposed during the modelling process
Design degrees of freedom:	the number of treatment combinations required to obtain the necessary degrees of freedom
Discrete choice:	the selection of one alternative among a set of mutually exclusive alternatives
Egress time:	additional travel time needed for travelling from the arrival location to the final destination point
Egress price:	additional price for travelling from the arrival location to the final destination point
Experimental design:	the specification of attributes and attribute levels for use in an experiment
Itinerary:	proposed route of a specific trip
Observation:	a choice made by an individual in a choice setting
OTA:	online travel agent
Panel data:	data incorporating multiple observations per sampled individual
Parameter:	a unique weight used to describe the systematic contribution of a particular element in a model

Route:	path from a point of origin to a specific destination
Significance level:	a value representing the probability that the analyst is willing to reject the null hypothesis when in fact the null hypothesis is correct
Stated choice experiment:	an experiment involving hypothetical choice scenarios and research-specified attributes and attribute levels
Flight transfer:	additional stop during allowing travellers to change of aircraft
Utility:	the level of happiness that an alternative yields to an individual
Utility maximization:	the act of seeking the alternative that yields the highest level of utility
WTP:	willingness to pay

Part I - Demarcation and Conceptualisation

This first part consists of a demarcation and conceptualisation of relevant concepts for this research. It provides the background of this research and defines the issues of the problem while discussing the actors and stakeholders involved.

Chapter 1 Introduction

1.1 Regional airports: background and literature review

During the last decades of the 20^{th} century, the demand for air travel grew at an average rate of 5% per year, and despite the impacts of the global economic downturn and the events of 9/11, annual growth levels of 5.1% are forecasted for the next 20 years (Hess & Polak 2005). While the growth in traffic has been accompanied by a high increase of seat-kilometres, runways and terminal, capacities have not been enlarged proportionally.

Many major airports are congested and overcrowded which creates delays and other nuisances for the aircraft, passengers, and cargo, and this could be seen as an actual transportation problem. Additionally, these airports are these days surrounded by established communities that are badly affected by the pollution, and noise produced by the airport exploitation (De Neufville 2000). Some of those problems can be solved rapidly. For example, traffic congestion may be reduced by establishing regulations, pricing strategies, etc. On the other side, the growth in traffic results in an environmental degradation, but these can also be mitigated by various forms of regulation. Sophisticated landing and take-off patterns have been used, and the way of propelling the engines has also been evolved in order to minimize the noise effects. Nevertheless, airports cannot handle infinitely with this traffic growth and some long term alternatives have to be found. Several options are feasible, namely supplying the travel needs by other modes (e.g. high-speed rail), increase the airport size by enlarging the terminals and adding runways, or finally building a secondary regional airport.

How additional airport capacity should be built is a major policy issue for both local and national governments and business interests, but the more obvious question is whether the additional airport capacity should be located at the existing major airport or elsewhere, at a secondary regional airport (De Neufville 2000).

The construction of secondary airports can have great advantages. First, it may have heavy commercial interests for the region (employment facilities, welfare and economic growth) and they have a geographically defined niche market. However, a secondary airport will only be a success if it is sufficient attractive. Passengers and airlines will not use a secondary airport when they can get a better service elsewhere (De Neufville 1995). In the passenger's perspective, the geographically accessibility, the high frequency of departures, and the low fares are the most attractive reasons for using secondary airports. For the airlines' perspective, a secondary airport is commercially attractive only if it provides a good market. Observable is that airlines can be competing according to different strategies. Some are competing on costs and may choose to reduce their level of services e.g. low cost carriers (Barrett 2000), whereas others are competing on quality and may choose to operate at airports with the highest access and egress and high slot availability e.g. network carriers. Hence, secondary airports offer many opportunities for airline attempting to reduce costs on level services.

Nevertheless, building a second regional airport requires a very good understanding in travellers' behaviour and in demand forecasting. Because the decision-making process in airport expansion depends on the passengers' demand, many studies on modelling travellers' choice behaviour have been conducted. Although this area of research has attracted increased activity the last ten years (Veldhuis *et al.* 1999, Pels *et al.* 2001, Basar & Bhat 2004), the development of an understanding of airport choice is still at an early stage and particularly compared to other dimensions of travel choice.

Observable is the fact that carriers only view passengers as their customer group, whereas airports both regard airlines and passengers as their key customers (Graham 2001). Though, passengers choose among different airports based on locality, they also choose on a series of airport and airline level-of-service attributes (fare, travel time, ease of access or egress, comfort, ect). Although, the fare is generally the most important factor influencing the decisions-making process, other attributes appear as important too. These attributes can be classified into two categories, namely the flight characteristics (flight frequency, flight travel time, number of stops, aircraft type) and the airport characteristics (airport access time, access costs, lounge, parking facilities) see (Harvey 1987, Loo 2007). Harvey (1987) and Ashford & Bencheman (1987) find in their study that airport access time and flight frequency are both significant attributes whereas Pels et al.(2001) indicate in their research for airport and airline choice that travellers are more likely to switch between airlines than between airports. This would mean that travellers may not be willing to use secondary airports. On the other hand, Basar & Bhat (2004) showed in their research that flight frequency (which is higher at secondary airports than at main airports) was the most important aspect and that it dominated the access-time factor.

As can be seen from all the previous studies, passengers' airport choice in multi-airport regions is an important research topic in the field of transportation. Despite all studies on passengers' behaviour, a trade-off will always have to be made between the three most important attributes: price, travel time, and frequency.

1.2 Problem demarcation

1.2.1 A policy management problem

As seen in the above literature review, a dilemma exists between extending (as far as possible) current airports and developing new secondary regional airports. Typically, the development of secondary airports highly depends on the operating airlines and on travellers. If airports are not attractive for airlines nor for travellers, implementing a secondary airport will have no benefits. Additionally governmental authorities see the development of secondary airports as a trade-off between a benefit for the regional economy and travel opportunities against the deterioration of the local environment. In fact, in case of success, the implementation of a secondary regional airport could be a good opportunity for companies to relocate into the region as well as it would lead to an economical regional welfare favourable for the local merchants.

The literature, nowadays, gives none information on the influence of airport location in travellers' itinerary choice. This means that if decision makers chose to develop a secondary airport instead of extending the actual main airport, uncertainty exists in the earning capacity of such projects.

When an individual decides to undertake an air travel trip (from an airport of origin to any destination), he usually considers all possible route alternatives before making his final itinerary choice. His itinerary choice may depend on several factors such as transport services or the existing traffic network. In this case, airports (e.g. airport location, airport services, etc) and airlines (e.g. company image, flight characteristics, etc) may both have an impact on individual's choice. Figure 1-1 represents the constitution of the airline market and its most important stakeholders influencing it. The three boxes at the bottom give an overview of what research & development could add if the research is focussed on the travellers, the airline market, or the airlines and airports. In this study the focus will be set on boxes 1 and 2.

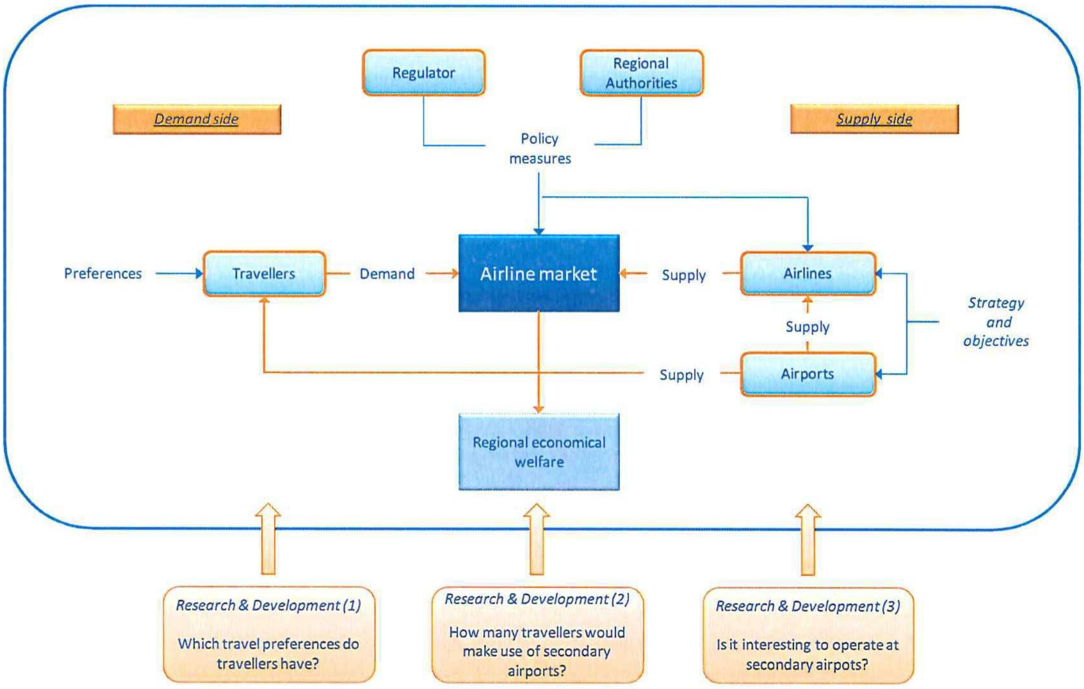


Figure 1-1 Airline market and its stakeholders

1.2.2 Introduction of a new methodology

Question 1 and 2 (mentioned in the R&D boxes of Figure 1-1), are questions that refer to demand analysis, which can be solved by modelling travellers' travel behaviour.

Since their introduction by Hensher and Louviere (1982), designed choice experiments have grown in popularity and specially stated choice (SC) methods are now a widely accepted data paradigm in the study of behaviour response of agents (Hensher 2003). This has lead to an increase of interest in the design of choice experiments in the transportation field and whereas a lot of books handle the topic of discrete choice modelling (Ben-Akiva & Lerman 1985, Louviere *et al.* 2000, Hensher *et al.* 2005), no books exists for designing SC experiments (Bliemer & Rose 2006). Furthermore, theoretical developments in and estimation of discrete choice models have had a large impulse and have lead to the expansion of several new techniques for determining designs for SC experiments, but such procedures have never been used in practice yet (Bliemer & Rose 2006). Therefore, air travellers' behaviour in travel choices is a rich field for applying this new methodology and analyse the performance of such methods compared to the old ones.

1.3 Research objectives

Most of the studies in the literature review focus on travellers' airline preferences, when departing or arriving at the main airport situated closely to the city. Additionally, most of studies are based on the trade-offs between access time versus flight frequency. However, no research has been conducted on egress time, which is highly dependent on the landing destination. Recently, Loo (2007) noticed that ticket price is the most important airport level-of-service attribute. Most noticeable is that, nowadays, secondary airports are only used by low-costs carriers. Their strategy targets to attract the largest amount of customers by offering the lowest prices and a minimum of services. They attempt reducing costs by using online reservation facilities instead of call centres and reservations employees, and by paying less airport taxes (due to the further away location). These strategy choices permit low-costs carriers to offer customers interesting flight options at low prices. Travellers during their decision making in choosing a flight itinerary, have to make a trade-off between low prices coupled with a long egress time and higher prices coupled with a shorter egress time. Figure 1-2 shows the situation in which travellers have to deal with this dilemma.

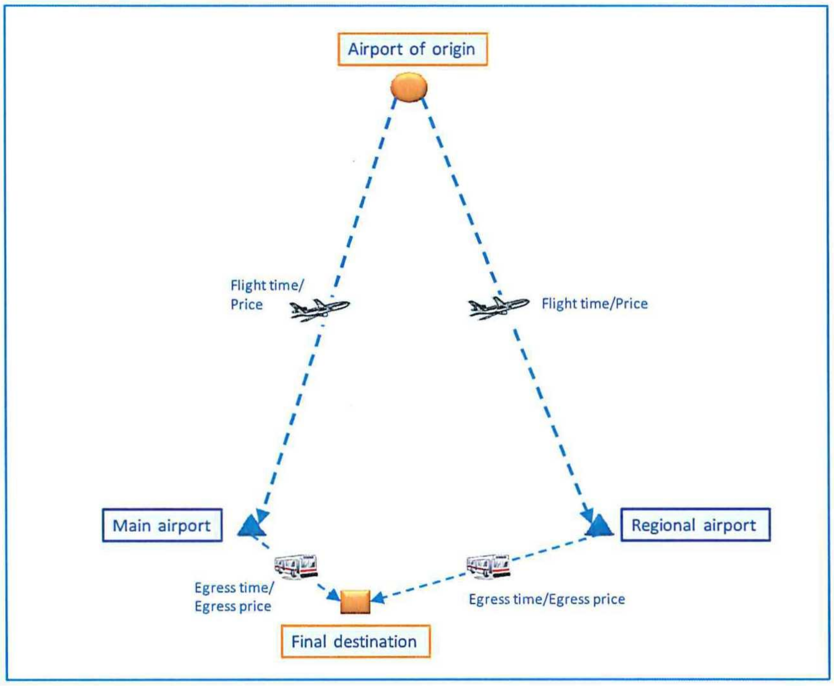


Figure 1-2 Trade-off between travel time and costs

By applying discrete choice modelling on air itinerary choice, this study gets a twofold aim. Firstly, as mentioned earlier, the purpose of this research is to provide information and more knowledge on travellers' behaviour to airports and airlines. Secondly, as secondary regional airports are an important topic, this study aims to secure new knowledge related to the influences of airport location in travellers' itinerary choices. Hence, from this empirical perspective, the objectives of this study are:

a. To determine which factors are currently influencing air travellers in selecting flight itineraries and to what extent.

b. To learn more about the effects of airport location on travellers' travel choices.

Stated choice experiments are often used in transportation studies for estimating and forecasting behaviour of travellers, though the reliability of the estimated parameters can be relying on the estimating models used or on the underlying experimental designs. Also more and more researchers realise that so-called efficient designs are able to produce more efficient data in the sense that more reliable parameter estimates can be obtained with an equal or lower sample size (Bliemer & Rose 2006). By applying the new knowledge on airline choices, the methodological perspectives of this study are:

c. To design a stated choice experiment based on different types of experimental designs.

d. To prove that the use of efficient experimental design results in more reliable parameter estimates than orthogonal experimental designs.

In other words, this research aims (i) to determine the influence of airport locations on air travellers' itinerary choice and (ii) to determine which type of experimental design in stated choice experiments result in the most reliable parameter estimates.

1.4 Research contribution

The first outcomes of this research will result in information concerning air travellers' behaviour in itinerary choices and especially in the attractiveness of regional secondary airports relative to main airports. As airports and airlines both have an impact on air travellers' itinerary choice, investigating air traveller's preferences may assist airports and airlines to adjust their strategy on travellers' needs in order to gain a larger market share.

Since Amsterdam Airport Schiphol (AAS) reaches its capacity, the Dutch government attempts to find other solution than expanding the airport and increasing the already highly noise emissions, and thus, developing a secondary airport could be an interesting resolution. Two secondary airports located close to AAS could be developed, namely Rotterdam airport and Lelystad airport. Because Rotterdam airport attained its capacity, the issue would be to develop Lelystad airport.

Since a lot of studies have already been done on airport access time (Ashford & Bencheman 1987, Harvey 1987, Pels *et al.* 2003, Basar & Bhat 2004), we choose to focus this research on the influence of the arriving airport location on travellers' itinerary choice. Additionally, inhabitants of a country have knowledge about access mode option and access times and prices, whereas all these familiarities are mostly unknown when arriving in a foreign country. This argues why the research focus is set on the effects of egress time and egress price on travellers' itinerary choice. Note that, the outcomes of this research can only be representative to foreign passengers travelling to the Netherlands. Consequently, this research partially contributes in giving insight in air travellers' behaviour in itinerary choices, which is necessary into the decision process of implementing a secondary airport. Furthermore, we will see that the decision of developing or implementing a secondary airport not only depends on the travellers. Other stakeholders may have different perceptions on this issue and can easily be a threat during the decision making process. Analysing the influence of these involved actors may contribute in facilitating the progression of taken some critical decisions.

The second outcome of this research relates to the methodological part of this study. Since the introduction of new techniques for designing SC experiment, they have never been applied so far. By generating different types of experimental designs and comparing their data estimation outcomes, this thesis research will contribute in the practical validation of the benefits of new designing methods for SC experiments. If it can be proven that efficient designs result in better parameter estimates than orthogonal designs, applying these new designing methods will permit analysts to improve their research.

1.5 Research approach

1.5.1 Discrete choice modelling and perspective on data-collection

In order to achieve the formulated objectives, using discrete choice models is a proven way to obtain insight into travellers' behaviour. However, the outcome of a particular choice process is based on a specific data set. When considering the issue of data for travel behaviour analysis, and in particular the analysis of travellers' choices for travel alternatives, it can be argued that there are three categories of data-types that together represent the bulk of the theoretical and empirical research efforts in this field (Chorus 2007). These data types can be simulated data, Stated Preference (SP) data or Revealed Preference (RP) data. The most frequent data-types in use are RP and SP data, and will be discussed in the following paragraphs.

RP data relates to people's actual choices in real-world situations. The data is obtained by asking travellers to report about their choice situations before having undertaken a certain trip (most of the time, this trip is their latest trip made). RP data has the advantage that it reflects actual choices. For this reason they widely used in travel behaviour research (e.g. Hensher & Bradley (1993), Polydoropoulou & Ben-Akiva (2001), ect). According to Train (2003), such data is limited to choice situations and attributes of alternatives that currently exist or have existed historically. In addition, only the chosen alternative is observed and the non-chosen alternatives have to be constructed by the researcher, which is not an easy task and results in limited value range and in correlation between the attributes.

On the other hand, SP data is data collected in experimental or survey situations where respondents are presented with hypothetical choice situations. Then the respondents state their needs, willingness to pay, or preferences for the alternatives. The advantage of SP data is that the experiments can be designed to contain as much variation in each attribute as the researcher thinks appropriate (Train 2003). Furthermore, the SP approach enables the evaluation of the demand for products and services that are not yet available in the market at the time of the investigation. The main disadvantage of this method is its limited external validity, the analyst can never be sure that the observed hypothetical behaviour resembles the behaviour in real life. What people say they will do is often not the same they actually do. This discrepancy is called hypothetical bias.

While the majority of studies of air travel choice behaviour make use of RP data, an increasing number of analyses are now carried out on SP data (e.g. Bradley 1998; Adler et al. 2005; Hess et al. 2007). As it has been seen in the precedent paragraphs, SP data has the advantage of being based on accurate records of all information, which is generally not the case with RP data. Hence, SP studies are generally more successful in retrieving significant effects for crucial factors having a huge influence on the choices (Hess 2007). However, by combining SP and RP data the advantages of each can be obtained while mitigating the limitations. SP data provide the needed variation in attributes, while RP data ground the predicted shares in reality (Train 2003).

Because of its advantages compared to RP experiments and because no RP data is available, in this research an SP experiment will be set up; sampled respondents will be presented with a number of hypothetical choice situations that differ on a number of attribute dimensions. These respondents are asked to specify their preferred alternative from the proposed set of alternatives.

1.5.2 Case scenario

In order to perform the SP experiment, a hypothetical choice situation has to be assumed. While conducting the case scenario, there has been chosen to deal with real departure location and a real destination instead of choosing two random locations. The reason is that, respondents may better conceive the choice situations when alternatives look more realistic instead of being subjective (e.g. people may better visualize a travel itinerary between Amsterdam and Barcelona than between location A and location B). Making the SP experiment look as real as possible (real departure/destination location, real airlines brands, etc) may result in better estimation of travellers' preferences than in the case of having a non-existing scenario frame (destination A/B, Airline 1/2/3, etc).

In this case scenario there has been chosen for a holiday trip between Amsterdam Airport Schiphol (AAS) and Barcelona city, assuming that Barcelona is one of the most famous European leisure destinations having a main airport and a regional airport. The context of the scenario is as follows:

"Imagine making a holiday trip (7 days) to Barcelona, leaving from AAS. In the following experiment hypothetical search results from an online travel agent website will be showed and the respondent will be asked to choose his most preferred flight.

There are two airports near Barcelona city, namely Barcelona Airport and the further away located Girona Airport (see map Figure 1-3). In order to reach the city, additional travel is necessary from both airports (additional travel time and travel price to reach the city)."

Respondents have been asked to review the different travel options (see Figure 1-4) and to make a choice.

Figure 1-3 Map airports around Barcelona city

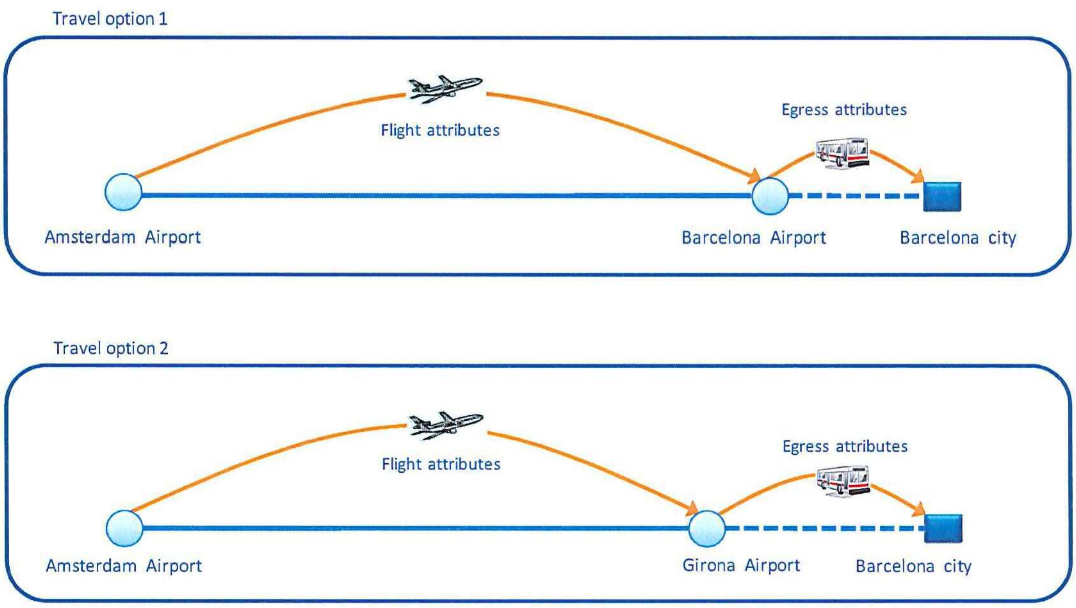


Figure 1-4 Alternatives in the Amsterdam-Barcelona case study

1.5.3 Report structure

The report is outlined in three different parts.

Part I is the so-called demarcation and conceptualisation of relevant concepts for this research. In chapter one, an overview of the literature background is given and the problem definition as well as the research objectives have been formulated. Chapter 2 consists of an actor analysis where the objectives and instruments of each actor involved in the airline market will be discussed. This chapter leads to a demarcation of the research and offers a broader insight for who the outcomes of this study can be helpful. In chapter 3, the concept of travel choice analysis will be introduced. This chapter will end with the set up of a choice model needed later on in the research.

Part II of the study will present how to generate a stated choice experiment for an airline case. Chapter 4 discusses the model specification needed to conduct an experimental design. Herein, will be discussed the choice set composition, the number of alternatives, and the number of attributes included in the model. Chapters 5 and 6 provide a description of the generation of two different types of experimental designs, namely an orthogonal design and an efficient design. The last chapter of this part gives an insight how to make a survey questionnaire on the basis of which data will be collected.

Part III is the last part of the report. The focus in this part will be set on the analysis of the collected data. Chapter 7 provides an overview of the sample of respondents that observed the survey. In chapter 8, the estimating models and their outcomes will be discussed. Chapter 9 presents the comparison of different generated experimental design followed by chapter 10 which forms the conclusions of the report.

In the last pages a bibliography is presented including all sources used in the main text as reference. Additionally an appendix has been added where some background or additional information can be read on some specific topics discussed in the report.

Chapter 2 Stakeholders of the Airline Market

2.1 Introduction

As seen in chapter 1, a trade-off exists between expending current main airports and developing secondary regional airports. It has been explained that a secondary airport is only profitable if it gets attractive for travellers and airlines. However, not all involved stakeholders have the same point of view on such policy implementations, and those can easily be obstructed or delayed if the interest of all stakeholders are not considered, as resources needed are spread across multiple actors and not exclusively in the hands of the government (Van De Riet 2003). In order to facilitate the decision process, the interest and objectives of affected stakeholders is not negligible and will therefore be discussed into the next sections. However, the next section will first discuss the concept of multi-actor complexity.

2.2 Multi-actor complexity

A multi-actor policy setting is characterised by a multi complexity (Van De Riet 2003). This complexity arises from the diversity in problem perceptions among the actors involved. The interest of actors (outcomes of interests) determines their objectives. The existence of multi actor complexity puts additional demand on the policy analysis. Van De Riet identifies three classes' requirements that a multi actor context puts on the research. This can be visualised in Figure 2-1.

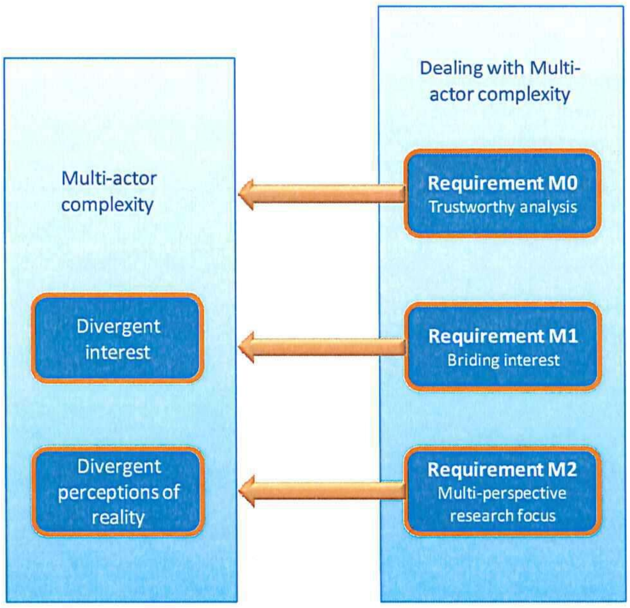


Figure 2-1 Additional requirements on analysis due to multi-actor complexity (Van De Riet 2003)

M0 requirement: trustworthy analysis

Stakeholder analysis is an input for the M0 requirement. "Trust can be enhanced by giving stakeholders a voice in the analysis. Giving them a voice leads stakeholders to perceive the analysis process as trustworthy. In addition it broadens the acceptability of the results in that it makes actors more willing to live with policies that are, in their view, less than perfect." (Van De Riet 2003).

M1 requirement: bridging interests

"Bridging stakeholders' interests is essential in multi-actor policy settings if the aim is to provide useful knowledge. The effort of bridging interest implies a multi-dimensional search for a solution in which the interests of all stakeholders are taken into account and the occurrence of losers is avoided as much as possible" (Van De Riet 2003).

M2 requirement: Multi-Perspective Research Focus

"The final requirement is that the research takes a multi-perspective research focus in analysing the problem and assessing the effects of the policy options" (Van De Riet 2003). The multi-perspective research focus demands that the problem is explored from the reality perspectives from a multi-actor point of view. This chapter identifies these perspectives and proposes three stakeholder disciplines with their respective perspectives.

2.2.1 Stakeholder typology

A first categorisation of stakeholders can be made on the level they operate. However, the problem is different on the various levels. At national level, the problem is the main airport's small capacity which threatens its functionality. The considered option is to use a secondary regional airport, where traffic can be diverted to.

The problem is that, such solutions are beneficiary for some stakeholders tough, there are disadvantages for others too. Involved stakeholders around the regional airport will see this as a benefit as it creates new welfare opportunities. On the other hand, it might increase the air traffic which forms the core of the problem for stakeholders with environmental issues. For example, airlines may have to shift their operations to another location which may create additional problems (e.g. airlines may become less attractive due to the further distance of airports to major cities). Obvious is that, further distinction of stakeholders type (i.e. governmental agencies, interest groups, regulators, suppliers and users) and their perspectives is needed.

Van De Riet and Turk (2006) described different possible points of view that an individual could have on an infrastructure, namely the society point of view, the user's point of view, and the supplier's point of view. According to Van De Riet and Turk, the society view captures the perspective of the community or all stakeholders affected by the infrastructure. The user view represents the perspective of each separate actor that uses or receives the services provided. In opposite of it, the supplier view captures the perspective of diverse organisations that provide the desired services, including production, trade, transport, and distribution.

Notice, that this actor typology is an abstraction. In fact there are many possible perspectives toward airport use and planning. Other perspectives can be seen as combinations of these three basic actor views. For example, airports (seen as the suppliers) incorporate the users' goals in their corporate strategy in order to please potential clients (airlines, travellers...) and maximise their market share. However, in order to deal properly with the multi-actor complexity, these three mentioned perspectives will be discussed to sort the stakeholders involved in the airline market.

The next paragraph puts forward which stakeholders are relevant for the system. Those stakeholders are the one whom might be interested in the research outcomes of this report.

2.2.2 Stakeholders analysis

As explained in Chapter 1, the Dutch government is interested in the development of a secondary regional airport close to Amsterdam. For this reason, there has been chosen to develop the stakeholders' analysis based on the Dutch airline market. The interest of this chapter is to get more knowledge concerning the stakeholders playing a role in the decision processes concerning the development and extension of a secondary regional airport in the Netherlands, namely Lelystad airport.

By asking the following questions the identification is facilitated and actors are not forgotten (Enserink *et al.* 2002)

- What actors are actively involved in the problem?
- What actors can be involved in either the origin or the solution to the problem?
- What actors have resources that can be important to the problem?
- What actors can be expected to have desire to be involved in the problem?
- What other stakeholders, not actively participating in the problem, will be influenced by the problem?

2.3 Stakeholders with a societal perspective

2.3.1 National authorities

Currently, the government's decisions concerning airports extension depend on the noise contour constraining factor that is allowed around the airports. On national level, the directly involved ministries are Ministry of Transport, Ministry of Economy and Ministry of Economic Affairs. Those are the three principal ministries responsible for aviation policy and airport planning into the Netherlands. However some ministries may be involved in second stage. For example, in Lelystad (where a regional airport is located) ministry of Agriculture, Nature, and Food Quality is involved as a Dutch national park is situated in the same province. As an illustration, Figure 2-2 represents the influence of the government on the Dutch aviation system.

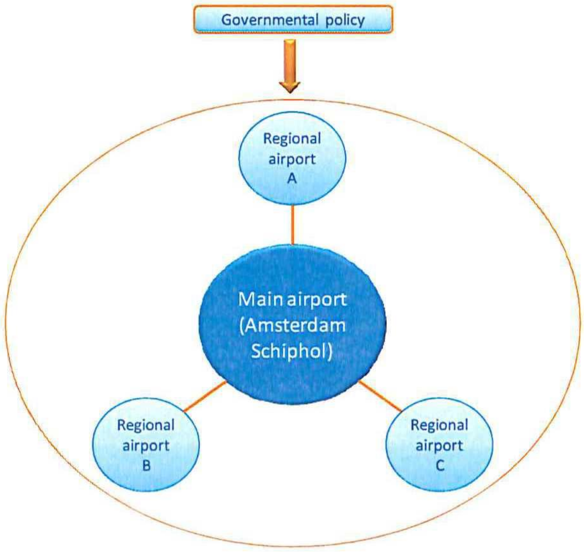


Figure 2-2 Governmental influence on the Dutch aviation system (Heerkens Thijssen 2007)

The role of the government is to assure stability in the aviation system and expects a high performance of it. Therefore, the government can use some measures to evaluate the performances of the airline market. In his research Gosling (2000) proposed some aviation performance measures. The one listed below are those with which the government can control the stability and the performances of the aviation system.

- *Mobility and accessibility*: under mobility and accessibility, four different categories are distinguished. The considered groups are the travel time, delay (difference between actual travel time and travel times in optimal conditions), access to desired destinations, and the access to airport system.

- *Economic well-being*: noticeable is that, what matters here, is not the share of transportation final demand but rather the productivity of the transportation sector. The problem is that the productivity of the airline market cannot simply be reviewed by the input provided by airport authorities but must also consider the airport related inputs by the airlines as well as the air traffic control system. These last two are hard to estimate.

- *Sustainability*: two aspects play an important role for the ability of future generations to meet their transportation needs. The first aspect is the dependence of the transportation system in oil-based fuels whereas the second aspect is the issue of deferred maintenance and renewal of the transportation infrastructure.

- *Environmental quality*: several environmental standards have been established concerning measures for commercial air service, aircraft noise exposure, and emission of pollutants.

- *Safety and security*: although aviation accidents are very low compared with the accident rates for other mode of transportation, they are of great concern of air travellers and especially potential terrorist attacks on airports or aircrafts remain a immense policy concern.

Obvious is that, policy regulations on the airline market are based on these described performance measurements. At the end, the government is the final decision maker who decides about the implementation of new policies. Therefore, such performance measures might be used to evaluate if stakeholders stick to governmental policies.

2.3.2 Regional authorities

In the Netherlands, government's aim is to make the Dutch provinces the leading authorities for regional airports (Ministry of Transport 2006). That way, provincial authorities may behave autonomously concerning their economical interests and may accept the consequences of their actions on travel growth and environmental impacts. Nevertheless, constraints with regards to environmental and external safety will be set by the government as well as the safety of the airports remains the government's responsibility.

As shown in Figure 2-3, three provinces and municipalities may have their importance in the traffic diversion from the main airport to the regional airports. Those provinces are the province of South Holland where Rotterdam airport is located, the province of Flevoland where Lelystad airport is sited and the province of North Holland where Amsterdam Schiphol airport is placed.

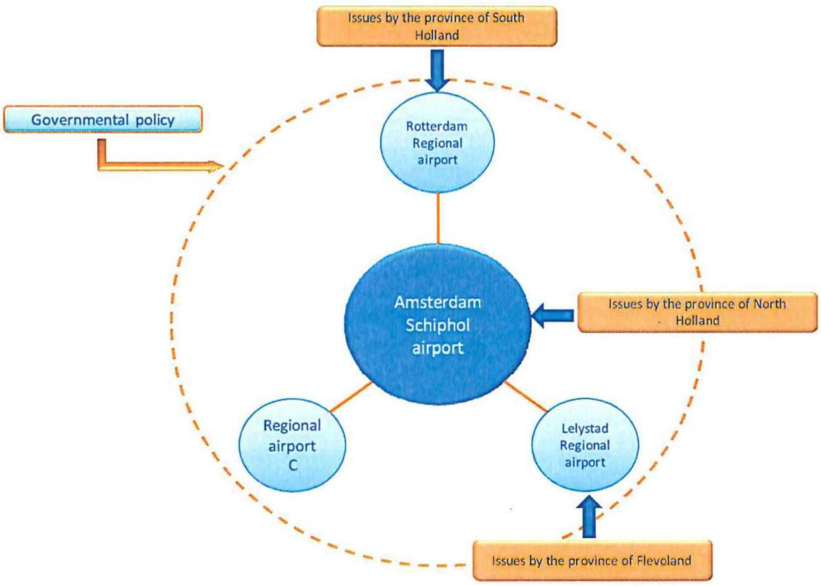


Figure 2-3 Provincial influences on the Dutch aviation System (Heerkens Thijssen 2007)

Noticeable is that Rotterdam airport is operating at capacity. This means that diverting flights from AAS to Rotterdam airport is not an issue. Also, airline carriers located on AAS will not switch to the regional airport of Rotterdam as the operating capacity is already attained. At the moment Rotterdam airport is used by low-cost carriers; the principal airlines operating at Rotterdam airport are Transavia, VLM airlines and KLM Cityhope and they only serve European destinations. Lelystad airport on the other hand is planning to extent the runway for offering future capacity. Because of the small density in that area, the affected population due to noise impact and air pollution will be limited. In the future, Lelystad might obviously become the second largest airport and would take over AAS's market share.

However, because of this potential growth, municipalities as environmental interest groups around Lelystad may have an important counting opinion and therefore have to be involved in the policy decision process rapidly. Notice that the municipality of Rotterdam not only has an interest in the Rotterdam airport but also in AAS as this last one is owned by 75% by the national government, 2,4% by the municipality of Rotterdam himself, and 21,8% by the municipality of Amsterdam (Schiphol 2008).

In the future, due to its vicinity to AAS and the catchment area of AAS, which is roughly the Randstad[1] area, Lelystad might become the most growing regional airport of the Netherlands. Obviously, Rotterdam is located closer to the Randstad area but in contrast to Lelystad airport cannot increase anymore because of the densely populated area. Therefore, powerful stakeholders such as provinces or municipalities have a high interest in the development of regional airports and help the airport financially for a better growth. Moreover, the Dutch ministry of Transportation has set a part of its focus on the infrastructural development of the province of Flevoland and especially around Lelystad; this is reported in the Dutch coalition agreement report of 2007 (Dutch government 2007).

This analysis show that Dutch authorities (national and regional), see the development of secondary airports as a positive aspect since it might solves the capacity problems at the main AAS and results in an improvement of regional economies.

2.4 Stakeholders with a supplier perspective

2.4.1 Airports

On its website, Schiphol airport precisely clarifies its objectives:

"The mission of Schiphol Group is to create sustainable value for its stakeholders by developing AirportCities and by positioning Amsterdam Airport Schiphol as the leading AirportCity. [...] It views an airport as a city that offers tailor-

[1] The **Randstad** (Rim City, i.e. a city at the rim of a circle, with empty space in the centre) is a conurbation in the Netherlands. It consists of the four largest Dutch cities (Amsterdam, Rotterdam, The Hague and Utrecht), and the surrounding areas. With its 7.5 million inhabitants (almost half of the population of the Netherlands; when other conurbations connected to this area are also taken into consideration, it would have a population a little over 10 million, almost 2/3 of the entire Dutch population) it is one of the largest conurbations in Europe. Its main cities are Almere, Amsterdam, Delft, Dordrecht, Gouda, Haarlem, Hilversum, Leiden, Rotterdam, The Hague, Utrecht, and Zoetermeer. The cities of the Randstad more or less form a crescent or chain. This shape has given the Randstad its name (*rand* means rim or edge and *stad* means city or town). The area that is enclosed by the larger cities is called the *Green Heart* (Groene Hart).

made facilities to a broad range of users. [...] It is the company's ambition to become one of the world's leading airport companies" (Schiphol 2008).

Nevertheless, to attain their objectives AAS would have to come in agreement with some policy constrains. The problem facing AAS is that it is running out of capacity and will not be able to accommodate growth in the near future. The forecast capacity shortage is largely a result of governmental constraints on aircraft movements, the number of passengers, the amount of cargo and the amount of noise emissions (Walker *et al.* 2001). A question that could be asked is until when AAS may accommodate the demand growth? In addition, authorities may not be willing to invest in more capacity and are interested in redistributing passengers over different airports (Van Eggermond 2007). For those reasons, diverting air traffic to secondary regional airports may be considered. Hence, if this really happens, competition between airports will exist, especially concerning the European flights (as most of the airlines operating at regional airports do not go overseas). Additionally, AAS risks losing an important part of its incomes since 35% of its total revenues stem from non-aviation business activities e.g. car-parking, concessions, advertising, etc (Schiphol 2008). Hence it sees the development of a secondary airport as a threat for its annual turnover when actual airlines would shift to the new airport.

The aim of developing regional airports (e.g. Lelystad airport) can be defined as maximizing the contribution of airports to their local economies, while also relieving pressure on congested main airports (Graham & Guyer 2000). However, due to their unattractive location, travellers have preferences for the centrally located main airport (Tron *et al.* 2007). For this reason, most airlines operating at secondary airports have a different strategy approach than airlines operating at main airports (see 2.5.1).

Concerning the airport demand, from a marketing perspective, both airlines and travellers can be seen as consumers. The airlines are the one that buy the airport facilities and the travellers are more the one that consume or utilize the airport product. If passengers' choice will rely on the nature of offered air service, for the airlines, the factors having the most importance are related to the nature of catchment area. As the flying route has to be attractive for the airlines, the catchment area has to be attractive for business and tourism. In his research about 'Airport managing', Graham (2001) developed the factors for airlines as well as for travellers that affect their choice in airports. Those factors are listed in Table 2-1 below.

Table 2-1 Consumers' factors affecting the choice of airport

Passengers	**Airlines**
Destination of flights	Catchment area and potential demand
Flight fare	Slot availability
Flight availability and timings	Competition
Frequency of flights	Network compatibility
Image and reliability of airline	Airport fees and availability of discount
Airline alliance policy and frequent flyer programme	Other airport costs (e.g. fuel, handling)

Surface access cost to airport	Range and quality service
Ease of access to airport	Ease of transfer connections
Car park costs	Maintenance facilities
Range and quality of shops, catering and other commercial facilities	Environment restrictions
Image of airport and ease of use	

Source (Graham 2001, p. 184)

In order to increase their revenues, main airports as well as regional airports need to adapt their strategy on travellers' needs. Hence, gaining knowledge in air travellers' needs is one of the research objectives of this thesis.

2.4.2 Regulators: Air Traffic Control the Netherlands

As an Independent administrative body, Air Traffic Control the Netherlands (LVNL) provide air traffic control by order of the Minister of Transport, Public Works and Water Management (Air Traffic Control the Netherlands - LVNL 2008).

AAS is a complex runway system, designed in a time with small traffic volume. Nowadays, the traffic volume is so large that the regulator (LVNL) handles more than 100 flights an hour. However to accomplish the wishes of AAS, LNVL supports the mainport's objectives and tries to optimize the airspace within the limiting conditions for safety and environment.

For regional airports, the same safety criteria apply as for international airports whereas the amount of traffic is much smaller. Beside AAS, LVNL is also responsible for the development of flight procedures and routes of Rotterdam airport, Groningen airport and Maastricht airport. Especially for airports in vicinity of each other, it might be that air traffic routes interfere with each other. Therefore capacity planning on the ground is not negligible and has to be controlled in cohesion with available airspace possibilities.

Depending of the overlap of the TMA[2] areas, conflicting air traffic routes or flight paths will result. Rotterdam airport and Lelystad airport, which are relative small airports, already overlap and interfere with the TMA of AAS. An increase in traffic at Lelystad airport will also affect the overlap in both TMA's. Therefore, regulators as LVNL use the TMA as indicator for reviewing the airspace capacity. Figure 2-4 illustrates the relation between the AAS's TMA and the different control zone (CTR). The CTR is the controlled airspace set up to protect air traffic operating to and up that airport. It can be seen that a traffic growth from and to Lelystad airport will demand a high security level as both TMA's could interfere in uncontrolled areas.

[2] Every airport needs a Terminal Manoeuvring Area (TMA) within which incoming aircrafts are directed to the airport and outgoing air traffic is send in the right direction. The conflicts between incoming and outgoing traffic are generally solved within the TMA

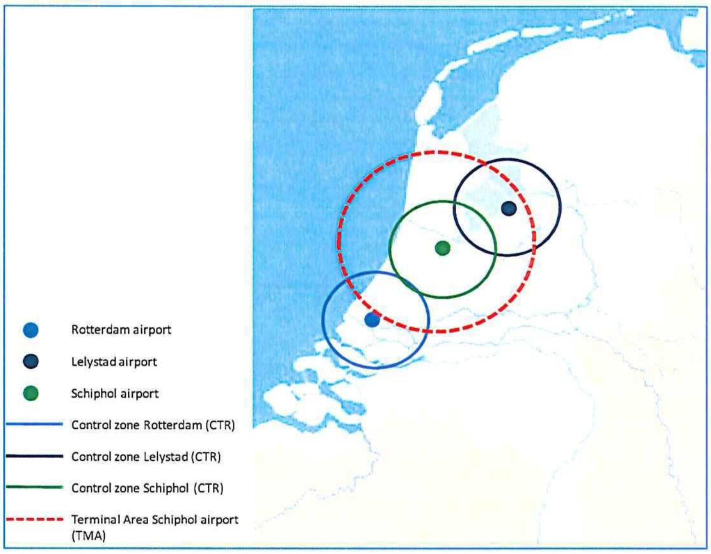


Figure 2-4 Schiphol's TMA and airports'control zone (AIS 2008)

Globally while optimizing the use of the airspace above the Netherlands, LVNL tries to find a balance between its three outcomes of interest; namely safety, efficiency and the environment. Therefore, it plays an important role regarding the objectives and future functions of AAS as well as those of regional airports experiencing rapid growth (e.g., Lelystad airport).

2.5 Stakeholders with a user perspective

2.5.1 Airlines

Airlines are the first users of the airport capacity and therefore have in this research a user perspective. Since the airline deregulation, a competition between airlines and airlines alliances exists (Button 1991). As a result of it, economy fares have fallen and have tended to increase the importance of the airport costs in the average fare (Barrett 2000). Since that moment, two different types of carriers with different strategies could be distinguished namely the already existing network carriers (NC) and the upcoming low cost carriers (LCC). In the coming paragraphs, the mission of both different types of airlines will be discussed.

Network carriers (NCs)

In order to determine the mission of the network carriers the case of KLM (a Dutch NC's) will be discussed. On its internet site KLM explains its objectives which are:

"KLM wishes to play an active part in setting the criteria necessary to realize its objectives: growth opportunities at the Schiphol home base, access to all markets that add to the quality of the network, and a level playing field for all. KLM seeks to balance the company's interest with those of its local surroundings. [...] KLM's strategic goal is profitable and sustainable growth. Together with Air France, it will achieve this through the further development of its three core activities in the most attractive markets, through cooperation within SkyTeam and through further reductions in unit costs" (KLM 2008).

Franke (2004) argues that at this time, NCs try to draw more and more traffic to their hubs, since they could create a disproportional increase in connections at incremental cost. Also he explains that the NCs predominant optimization targets are a coverage of as many demand categories as possible and connectivity in the hub which can be found back in KLM's objectives. The hup-and-spoke network has enabled the airlines (and alliance partners) to offer a wide range of products to all kinds of customer needs (e.g. business or leisure customers; continental or intercontinental routes, rebooking facilities, in-flight services...) and at the same time allows them to bundle the traffic flows and thereby increase cost efficiency. Though, this strategy has some important negative aspects which reflect a loss of convenience for the passenger who would prefer direct flights, and a high cost penalty for the airline on the operational side. Franke, named some consequences of those negative aspects: temporary congestion (reduced airside productivity), time-critical connection (special process required), and strongly fluctuating utilization of ground handling facilities. All this consequences could additionally lead to poor punctuality performances which as a summation might be decreasing the travellers' perception of network carriers. Although NCs are unwilling to switch their services to secondary airports (De Neufville & Odoni 2003), airlines are focussed on a high frequency of service, which is easier to maintain when operating at secondary airports. Typically, LCCs align their strategies on such kind of issues and permit to be competitive to NCs on medium-short distances.

Low Cost Carriers (LCC)

It is with a completely different business model that LCCs have entered the airline market. The so-called LCCs have successfully designed a focused operation providing them with a significant cost advantage. Experts estimate that they operate with up to 60% lower unit costs than network carriers (Hansson *et al.* 2003). During their growths, LCCs have created a unique value of proposition through product and process design that enables them to propose low fare rate in exchange to no service. Gillen & Morrison (2005) proposed several examples of service feature trade-offs: less frequency, no meals, no free or any alcoholic beverages, more passengers per flight attendant, no lounge, electronic tickets, and less legroom. Also LCC are usually operating at secondary airports which are further located to the main cities. Although the connectivity for accessing or feed the airport is not often as good as at main airports, the aircraft taxing times are much cheaper than at main airports. Mason (2000) argues that the operational strategy of secondary airports allows airlines to achieve utilization of its aircraft of two hours per day more than a network carrier based on a main airport.

Transavia which is one of the biggest LCCs in disserving European destinations applied this above mentioned business model; this is noticed in their strategy approach:

"The strategy of transavia.com is aimed at positioning the company as a "web-based" travel brand with flights as its key activity and at close collaboration with business partners in the expansion of its package of products. transavia.com would like to distinguish itself in the marketplace by basing its activities on "low-cost, low fare with individual service". It is also crucially important to transavia.com's healthy future that the company is successful in its selective growth, aimed at healthy returns and based on a carefully devised route network and revenue management" (Transavia 2007).

In Figure 2-5, are aligned the drivers on which LCCs base their strategy and reduce their costs.

Although airline attributes have the big impact on travellers' travel choice, airport location seems to play an important role as well. For example, in his study on airport choice behaviour, Bondzio (1996) showed the travellers' value for travel time plays an important role regarding airport choice behaviour and specifically the access time to the airport. Also, he demonstrated that access time is more important for business travellers than for leisure, which can be understandable. The same conclusions have been repeated in the research done by Pels et al (2003). These results show that, travellers not only focus on airlines' low fares or prices but that, good infrastructural facilities are needed as well in order to make the access to the airport attractive for passengers.

2.6 Network analysis

In this chapter, different stakeholders affecting the airline market have been described. Important is that those stakeholders can be sorted in different groups, namely in the one having a social perspective, a supplier perspective and a user perspective. All of those stakeholders have their own perception on the market and try to influence the market in order to attain their objectives.

Given the stakeholders of the aviation market, it is essential to see whether the involved stakeholders are critical for the future development and growth of the system. If it seems that their resources are not replaceable and their dependency has a high level, then the involved stakeholder is called a critical actor for the system. Resource may include expertise, money, asset and lobby. The stakeholders' dependency is determined by three levels, namely low, medium or high level. Notice that, stakeholders' position and influence on the system might change depending on the scenario and the topic of the discussion. Another important factor is the power of each involved party. The power permits the involved actors to obstruct or to advance the processes. It might happen that, at the end of long discussion processes and after having obtained some cooperation and consensus, people do not wish that other dominant parties became a hindrance for the decision process. In this situation all the named stakeholders, excepting the travellers, do have a critical position. In Table 2-2, answers on previous formulated questions by Enserink et al. ((2002) are listed.

The outcomes of this analysis shows that, despite of their importance due to their production power (demand in aviation), travellers are not critical actors for the market. In fact travellers may have the choice to use other transportation modes to undertake their trip (car or train) unless the destination is not reachable with other transportation modes. This results in the fact that making destination choices might constrain the mode choice. Van Zuylen (2005) demonstrated in his framework that the freedom of choice depends on the time. That means that, once a decision is undertaken, some other choices are not possible anymore, and depending on the time period, choice possibilities get restricted. So, if travellers' destination is only accessible by plane, their only choices result in airline choices and in airport choices.

Concerning the other stakeholders of the market, their resource, power and dependency determine if they will be critical or not in some new decision-making processes. Noticeable is that all other mentioned actors are defined as critical and play an important role in the maintenance of the aviation system. Notice that the implementation of a regional secondary airport will be a benefice for the local economy. However, local companies are not critical in the decision process of implementing a secondary airport. Because they can only perceive the consequences of a secondary airport in their region, they might be able to subsidy a part of the implementation costs. Nevertheless they are not critical in the decision process.

Table 2-2 Network analysis

Stakeholder	Resources	Power	Dependency	Critical
Governmental authorities (Ministry of transportation, Ministry of Economic Affairs)	• Expertise • Subsidies • Policy rules	• Productive power • Start decision • Law and enforcement (directives)	High	Yes
Provinces	• Area plan • Subsidies (money for area development)	• Law and enforcement • Blocking power (area plans)	Medium	Yes
Airports	• Money • Expertise • Assets • lobby	• Productive power • Blocking power	High	Yes
Regulator	• expertise	• law and enforcement	Medium	Yes
Airlines	• lobby • expertise	• Productive power • Blocking power	High	Yes
Air travellers	• indirect production power	• Boycotting	Low	No
Local economy (companies)	• subsidies	• productive power	Low	No

Since stakeholders have different perception on this issue, it is interesting to analyse if stakeholders encourage or not the development of a secondary airport. Table 2-3 presents an overview of the stakeholders' opinion on the implementation of a secondary airport. Obvious is, that the government and the air regulator are favourable to the development of a secondary airport. A reason is that it would solve the capacity problems at AAS and could be a nice opportunity to increase the concerning regional welfare. On the other hand, the development of a secondary airport would be a threat for AAS which would prefer to expand its airport instead of developing a secondary airport. Concerning the airlines, operating at secondary airports is not necessarily a bad thing as the operating costs would be lower than at the main airport. However, airlines depend on the travellers demand. Despite the fact that air travellers have no critical status in the air market, they do have an important impact on the decision process of implementing a secondary airport. Since travellers are the users of airports and airlines, without travellers, airports and airlines will not make any profit. For this reason, having more insight air passengers travel preferences is necessary in order to analyse if secondary airports are attractive or not. Based on the advantages and inconveniences of implementing a secondary airport and on the stakeholders' perception on it, the government will finally have to take a decision as he is the decision maker of this problem.

Table 2-3 Stakeholders' perception on the development of a secondary airport

Stakeholder	Perception on the development of a secondary airport
Governmental authorities (Ministry of transportation, Ministry of Economic Affairs)	Pro: secondary airport will solve AAS's capacity problem while offering new route possibilities. Also it enables the economical development of a new region.
Provinces	Pro: enables the increase of the regional welfare
Airports (AAS)	Against: sees the development of a regional airport as a threat and as a lost of profit.
Regulator	Pro: permits to divert the traffic which leads in more traffic security
Airlines	Unclear: see regional airports as an opportunity to decrease the operational costs. However, staying at the main airport takes out all the risk of non attracting people.
Air travellers	Unclear: secondary airports usually are cheaper but located further away from the city centres. This result in more travel time. More insight in travellers' perception will be ginven by the conclusions of this research.
Local economy (companies)	Pro: the implementation of a secondary airport would signify that the region will become more attractive which is good for the local economy.

2.7 Conclusion

In this chapter different stakeholders of the airline market have been analysed. In this research the societal interests have been set on the parameters influencing air travellers' in their travel choices. Therefore, the outcomes of this study are mostly important for the airlines, which may adapt their strategy on travellers' needs, but also for the airports as travellers and airlines form the airport demand. Looking at the airline market (Figure 1-1), the outcomes of this research should give answers to the first and second research boxes which should be beneficent for the involved stakeholders to improve their objectives. In Table 2-4 below, all the discussed stakeholders' goals and instruments (possible actions) have been enumerated as a concise overview.

Concerning regional airports, stakeholders of the aviation market have different perceptions on their implementation. On one hand, governmental authorities perceive this as a solution to solve AAS's capacity problems, and as a way for improving the local economical welfare. On the other hand, AAS sees the development of a secondary airport as a threat to its market share, and will try to obstruct governmental decisions to develop a secondary airport. Finally, airlines see secondary airports as an opportunity to increase their flight frequency and decrease their operating costs. However, good access to, and from, the airport is necessary in order to attract travellers. Air travellers' preferences in itinerary choice are therefore an important aspect in the decision process of implementing a secondary airport.

Table 2-4 Objectives and instruments of the stakholders involved in the airline market

		Objectives		Instruments
Government and regional authorities	G_O 1	Solve capacity problems at AAS	G_I1	Extent AAS
			G_I2	Implement a regional secondary airport
	G_O3	Increase the public value of regions having airports	G_I3	Investment in infrastructure and subsidies
	G_O4	Assure a healthy, clean and safe living environment (passenger and environmental)	G_I4	Implement policy restrictions concerning noise emissions, safety and security
Airports	AP_O1	Increase the number of airlines operating	AP_I1	Reduce the operating costs and taxes
			AP_I2	Assure faster operating services
	AP_O2	Increase the number of travellers	AP_I3	Increase the number of operating airlines deserving more destinations
			AP_I4	Reduce airport fees (tax)
			AP_I5	Improve services and transfer facilities (shopping facilities, decrease in waiting times...)
			AP_16	Improve airport accessibility (public transportation, bus shuttle),
	AP_O3	Increase non-aviation business revenues	AP_I7	Improve commercial facilities (car-parking, shops, restaurants, hotels...)
Regulators	R_O1	Create capacity for strengthening the aviation network	R_I1	Optimise the airspace controls and the interfering airspace areas
	R_O2	Limit the regional hindrance	R_I2	Assure an optimal airspace safety and capacity
	R_O3	Create a clear perspective for spatial development	R_I3	Optimise the airspace control methods and techniques
Airlines	A_O1	Reduce airports costs	A_I1	Operate at secondary airports
			A_I2	Reduce airport services
			A_I3	Fly at off-peak hours
	A_O2	Reduce flight costs	A_I4	Reduce in-flight services
			A_I5	Reduce kerosene consumption
			A_I6	Optimise aircraft occupation rate
	A_O3	Reduce operation costs	A_I7	Use electronic ticketing

			A_I8	Reduce personnel
	A_O4	Increase travellers volume	A_I9	Increase routes/flight destinations
			A_I10	Reduce ticket prices
			A_I11	Increase marketing operations
Travellers	T_O1	Minimise travel costs	T_I1	Choose itinerary with lowest fare
			T_I2	Choose airline with low fare
			T_I3	Choose airport with low fares
			T_I4	Travel less (frequent)
	T_O2	Minimise travel time	T_I5	Choose for itinerary with minimum travel time
			T_I6	Choose airline with short travel time (no transfers)
			T_I7	Choose airport with short access and egress times
			T_I8	Choose other travel mode
	T_O3	Maximize comfort	T_I9	Choose airline with high quality service
			T_I20	Choose airport with high quality service

Chapter 3 Travel Choice Modelling

3.1 Travel choice in the field of transportation

Trip making is the result of individual choice behaviour. An individual person decides whether he will leave home to do an activity elsewhere or not, where he will do that activity, and how he will travel to that destination (Bovy *et al.* 2006). All these decisions are dependent on many factors, such as the availability of the transport services or the existing traffic network. A person who chooses to go to work by car, using a certain route has made these decisions based on his knowledge, preferences, and constraints (such as time and budget constraints).

Schoemaker et al. (1999) developed a framework (Figure 3-1) that represents the characteristics and the composition of transport systems.

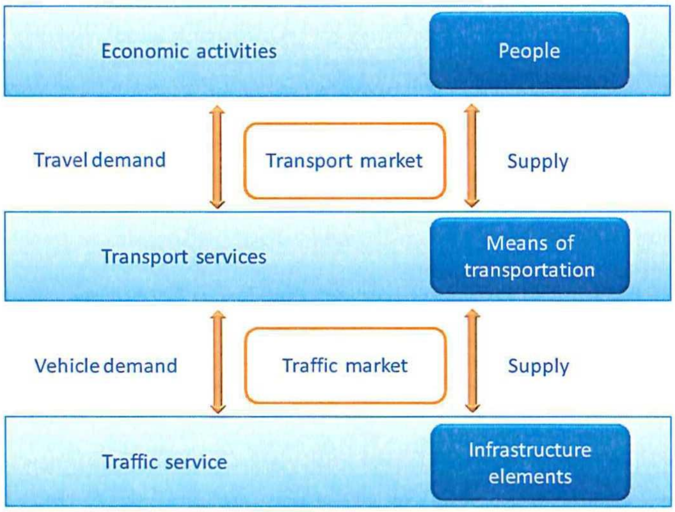


Figure 3-1 Transportation layer model (Schoemaker *et al.* 1999)

The activity-layer relates to the activities performed by people, companies or organisations. Hence, typical activities can be living, work, study, shopping, etc. Since activities have different locations in space and time, people have to make trips. The transport services-layer offers transport facilities to people. Just as in the activity-layer, many actors are involved (from individuals to organisations) each offering transport services to facilitate the travel demand. In fact, this layer provides a supply outline in space and time for the transportation of people. Typical characteristics of this supply are the levels of service, prices, and quality (Schaafsma *et al.* 2001). It is important to notice that the actual travel demand is not equivalent to the desired travel demand. Travellers might choose activities that are different to the most preferred chosen activity because of the long travel time or the high prices related to the transport services. In addition, when a traveller decides to undertake a trip, it can be assumed that his activity location choice has already been determined, and that the remaining choices concern the route choice, the mode choice, and the departure time choice. When modelling traveller's choice behaviour, this aspect might include some uncertainty and it would be always be difficult to guarantee that travellers behave in reality as at observed in a model.

In air transportation, when a traveller decides to make a trip, his activity location choice and the mode choice are fixed and a choice has to be made between several flight alternatives, where alternatives are mutually exclusive possibilities for making a trip of which only one can be chosen at the same moment (Bovy *et al.* 2006). The flight could be characterized by a certain amount of attributes which will influence the traveller to make a rational choice among the different proposed alternatives. Therefore, the traveller decides according to his own personal views and preferences and tries to optimise his own personal situation by choosing the alternative that suits him the best. Furthermore, travellers become sufficiently informed about the product category, which they are looking at, to form preference ordering which involves trading off product attributes possessed by product alternatives. Hence, travellers develop a preference ordering for alternatives, and depending on budget or other considerations, make decisions which alternative to purchase (Louviere *et al.* 2000).

One of the goals of this research is to gain better understanding of the needs and wants of individual air travellers and this can be provided by focusing on the relationship between travellers' choice and the factors that influence them to make a specific trip. Based on the travel choice framework developed by Proussoulaglou (1999), and in order to provide a better insight in travellers' airline choice, a conceptual framework for air travel choice is presented in Figure 3-2.

This framework shows that by searching and learning, the traveller gains knowledge on what satisfies his needs. Consequently, the traveller forms an ideal alternative based on his preferences and begins to compare the value of his initial alternative to the available set of alternatives. In the case that none of the proposed alternatives matches with the initial preferred alternative, the traveller has to adapt and/or reform his initial preferences according to the available set of alternatives. If the traveller decides to purchase an activity, the decision making process will end. The chosen alternative is the one that is most preferred, given the available options.

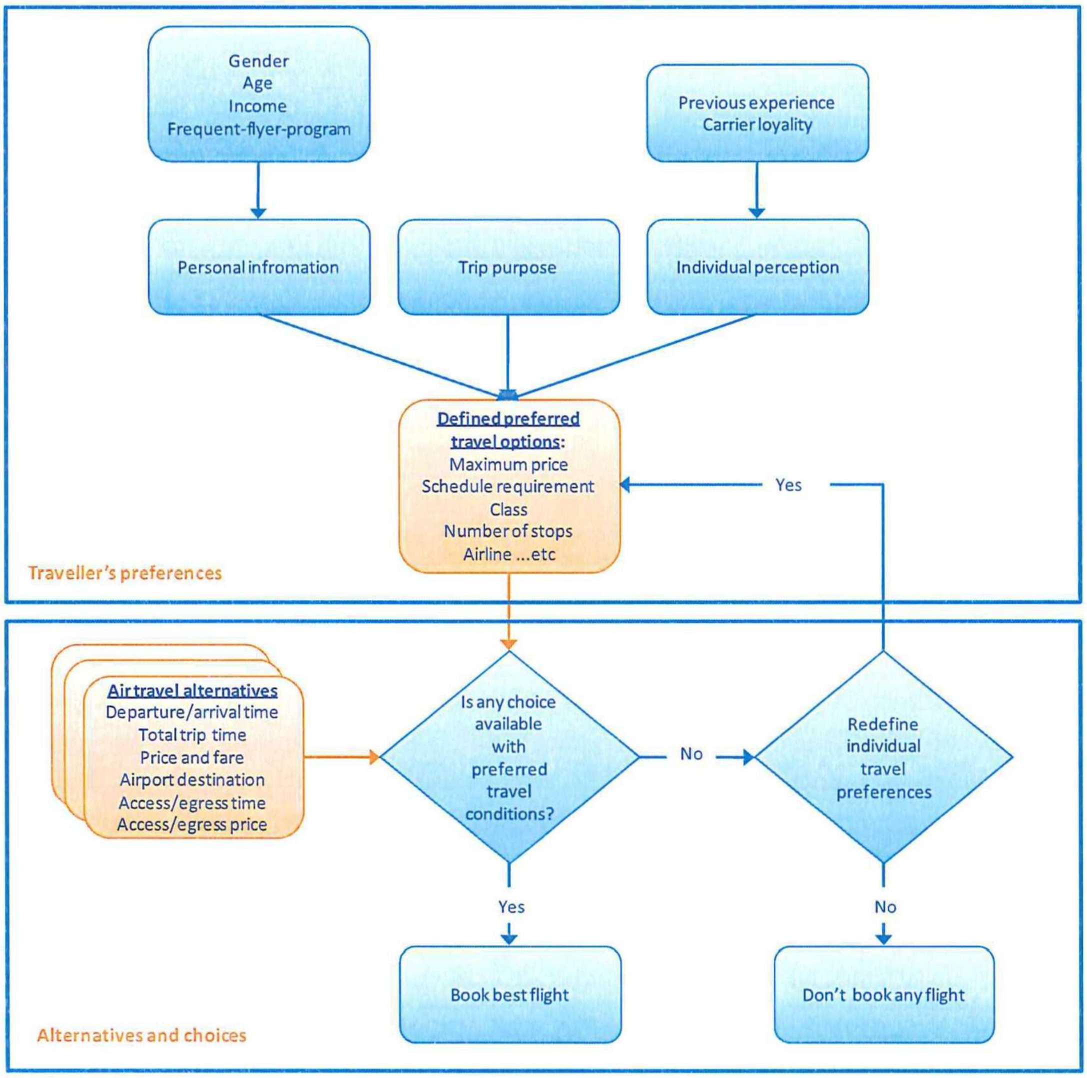


Figure 3-2 Air travel choice framework

3.2 Choice modelling based on utility maximisation

Modelling travel behaviour is a key aspect of demand analysis, where aggregate demand is the accumulation of individuals' decisions. The individual decision-making entity depends on the particular application; hence, the decision maker can be a person or group of people (such as an organisation for example). Beside the decision-maker, the framework for a discrete choice model is presented by a set of general assumptions about alternatives, which determine the options available to the decision-maker, and attributes, which measure the benefits and costs of an alternative to the decision-maker (Ben-Akiva & Bierlaire 1999).

The individual's preferences for consuming goods or for services can be expressed in "utility". Notice that the amount of utility is represented by a relative and abstract value and cannot be observed nor measured by itself. Satisfaction and happiness toward a consumed good or service will be translated into a positive utility, whereas dissatisfaction will be expressed as a negative utility. Each individual is assumed to behave rationally, such that he or she will choose that alternative which gives the largest relative utility.

Although the decision-maker has a perfect discrimination capability, the analyst is assumed to have incomplete information, and therefore when determining individual's utility, uncertainty has to be taken into account. Manski (1977) identifies four different sources of uncertainty: (i) unobserved alternative attributes, (ii) unobserved individual characteristics, (iii) measurement errors, and (iv) proxy (or instrumental) variables. Hence, to reflect this uncertainty, the utility is composed of two parts. The first part is the observed deterministic part (V) and the second part is the non-deterministic, unobserved part of the utility modelled as a random variable (ε). More specifically, the utility U_{iq} that individual q associates with alternative $i \in C_q$, where C_q is the set of available alternatives for individual q, is given by:

$$U_{iq} = V_{iq} + \varepsilon_{iq}, \quad \forall i \in C_q, \qquad (3.1)$$

where the deterministic observed part V_{iq} is described by a function $f(\beta, X_{iq})$, where β is a vector of taste parameters and X_{iq} is a vector of attribute levels for alternative i that can be measured or observed (such as price and travel time). In addition, socio-demographic attributes of decision-maker q (such as gender or income) can be included in the deterministic part of the utility function. The non-deterministic non-observable part of the utility function ε_{iq} is assumed to follow a given random probability distribution (Ben-Akiva & Lerman 1985).

As mentioned before, it is assumed that all decision makers aim to maximise their utility, and select the alternative that has the highest utility among the alternatives in the choice set. The probability that alternative i is chosen by decision-maker q from choice set C_{iq} is therefore:

$$P(i \mid C_q) = P[U_{iq} \geq U_{jq} \forall j \in C_q]. \qquad (3.2)$$

It is important to realize that only the differences between utilities are relevant here and not utilities themselves. This can be seen by rewriting equation (3.2) as follows:

$$P(i \mid C_q) = P[U_{iq} - U_{jq} \geq 0 \forall j \in C_q]. \qquad (3.3)$$

Applying this general utility maximising theory to the problem of travel choice in air travel, some decisions have to be made about the function $f(\beta, X_{iq})$ and the random variable ε_{iq}. In this research, every alternative i will represent a trip to Barcelona city. Also, in the next paragraphs will be described how to generate a set of alternatives C_q. Furthermore, the alternatives i and attributes X_{iq} that will be used will be discussed, as well as the assumptions that have to be made about the random part of the utility ε_{iq}.

3.3 Alternatives

In this study, each alternative represents a trip from AAS to Barcelona city. As seen previously, the aim of the research is to get more knowledge about the attributes influencing travellers' travel choice. Each alternative is composed of a flight part and of a ground transportation part.

Alternatives may be characterised by being labelled or unlabelled. An alternative is defined as labelled if a "name" card is attributed on it. In the case of air travel choice, labelled alternatives would have been: Air France, KLM, Iberia, etc. In the same case, unlabelled alternatives are alternatives without specific names i.e. trip A, trip B, trip C, etc. The advantage of unlabelled alternatives is to offer more flexibility in creating alternatives. For example, a choice set with labelled alternatives would only have considered the comparison between Air France, KLM, and Iberia, whereas a choice set with unlabelled alternatives also allows the comparison between KLM (1), KLM (2), and Iberia. Unlabelled alternatives have disadvantages though, it restricts the researcher to use generic parameters. Parameters are called alternative-specific when the parameter for a certain attribute differs across the different alternatives (e.g. the parameter for travel time may be different for different travel modes). If parameters are the same across all alternatives, parameters are called generic. The disadvantage of using unlabelled alternatives is that only are generic parameters meaningful whereas labelled alternatives may contain alternative-specific parameters only and/or with generic parameters.

3.4 Choice model specifications

3.4.1 Set of alternatives: C_q

In their framework (Figure 3-3), Bovy & Stern (1990) introduced various sets of alternatives, called choice sets. In choice behaviour modelling, the choice set is mostly described according to the traveller's point of view. However, it might be that this choice set differs if it is seen from another viewpoint e.g. the researcher's perspective. In the following part, both different perspectives will be discussed.

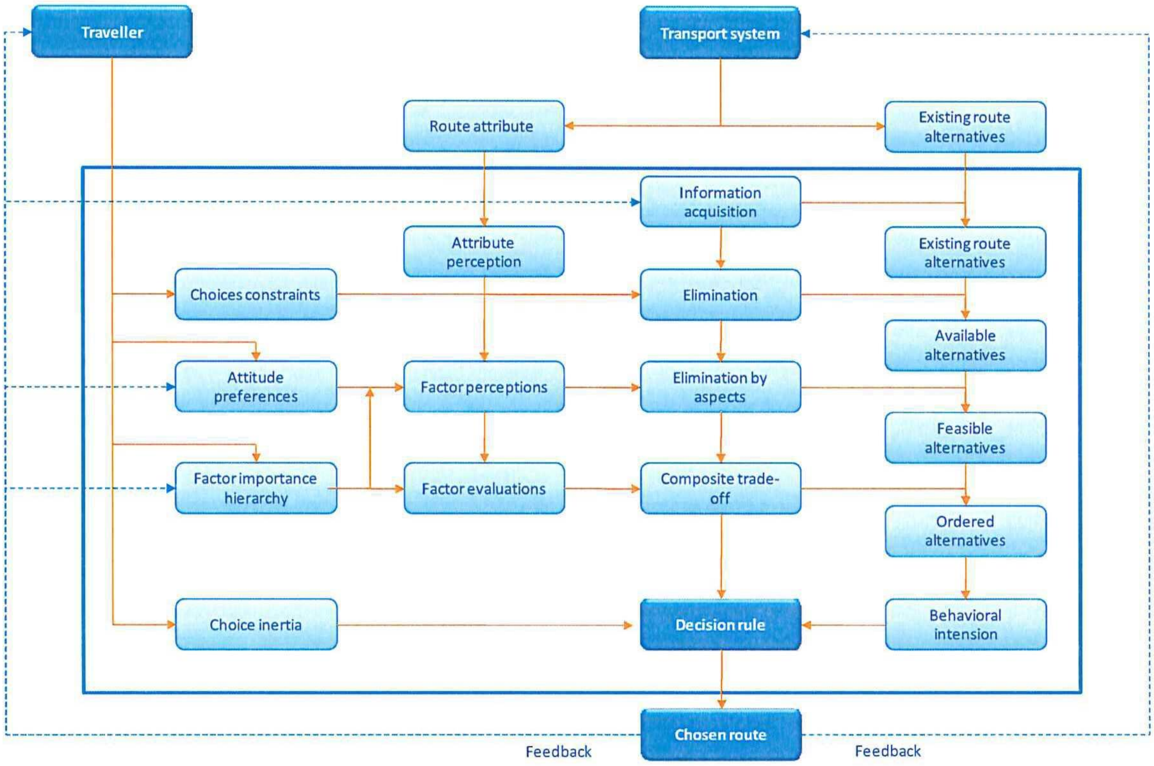


Figure 3-3 Conceptual framework for choice set formation of an individual traveller (Bovy & Stern 1990)

Set of alternatives from the traveller's perspective

Travellers and researchers have different knowledge of the transport system, and hence of alternatives available for a specific trip. Although the number of trip alternatives might be large, only a subset of them is known to, feasible to, and considered by the individual (Hoogendoorn-Lanser 2005). In addition, Bovy (2008) argued that only the known alternatives that satisfy the traveller constraints will form the choice set available from the perspective of the traveller. The traveller's perception of relevant alternatives and their attributes are incomplete and inaccurate and again are linked to his travel experiences and travel preferences. These individual traveller's properties that influence his decision-making can be found listed on the left side of the framework.

In her PhD thesis, Hoogendoorn-Lanser discussed the different choice sets which can be distinguished from the perspective of an individual traveller. Those are defined in Table 3-1. Furthermore, she argued that psychologists suggest that the number of alternatives that most travellers are able to consider in their final choice process is limited to more or less 7 alternatives. Figure 3-4 shows the relationships between the trip characteristics (existing, known, feasible…) and the corresponding set of alternatives.

Table 3-1 Choice set notions from the travellers' perspective

Terminology	Definition
Universal choice set	Set of all existing route alternatives in a given network from an origin to a destination.
Actual master set	Subset of the individual choice set containing all route alternatives known by an individual, travelling between an origin and a destination and leaving at a preferred departure time satisfying his or her travel needs.
Actual subjective choice set	Subset of the individual choice set containing all route alternatives known by and feasible to an individual, travelling between an origin and a destination and leaving at a preferred departure time satisfying his or her travel needs. (e.g. cost and time sensitive travellers)
Actual consideration set	Subset of the subjective choice set containing all route alternatives considered by an individual in the choice process. Those remaining alternatives are constructed by the traveller's preferences (specific carrier preference, departure interval, type of aircraft…)
Actual chosen alternative	Alternative that is actually chosen and is part of the consideration set.

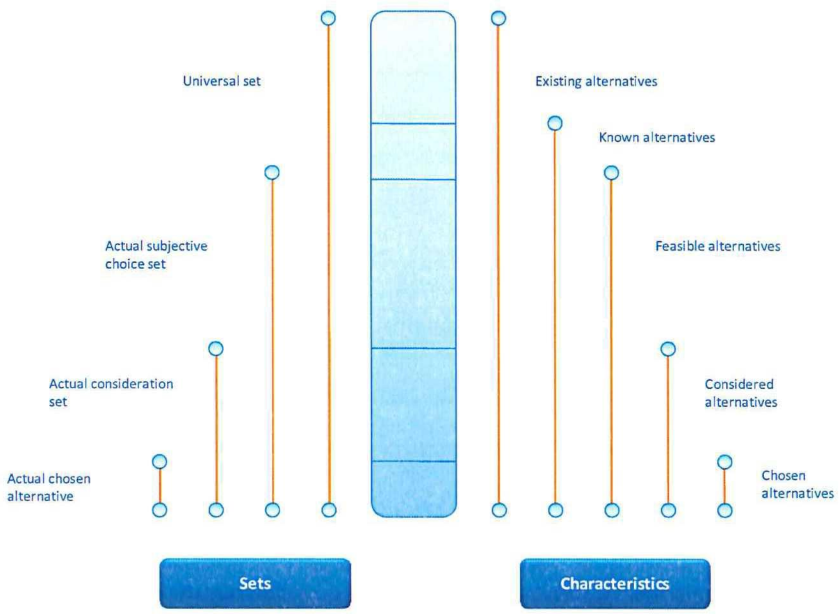


Figure 3-4 Relationship between the alternatives' characteristics and the corresponding sets

Set of alternatives from the researcher's perspective

Generally, a researcher may generate choice sets for four different kinds of applications namely: (i) analysis of available alternatives, (ii) estimation of parameters in utility functions or other choice models, (iii) prediction of choice probabilities, and (iv) data completion (Hoogendoorn-Lanser 2005). In this investigation, generated choice sets will be used in analysis and estimation applications.

Generation of choice sets is a complex task, since travellers and researchers often have different information about the situation and researchers do not precisely know travellers' preferences and considerations. Table 3-2 lists the different choice set notions from the researcher's perspective.

Table 3-2 Choice set notions from the researchers' perspective

Terminology	Definition
Universal choice set	Set of all existing routes in a given network from a specific origin to a specific destination. In this case it consists of all the alternatives between the moment that a trip generation has been decided and the moment of departure.
Observed/generated objective choice set	Subset of the universal choice set observed/generated by a researcher containing route alternatives that are assumed to be logical and feasible to an individual, leaving within a preferred departure time interval satisfying their travel needs.
Observed/generated subjective choice set	Subset of the universal choice set observed/generated by a researcher containing route alternatives that are assumed to be known to an individual, leaving at a preferred departure time that satisfy their travel needs.
Observed/generated consideration set	Subset of the universal choice set observed/generated by a researcher containing route alternatives that are assumed to be considered by an individual in the choice process.
Observed/generated chosen alternative	Alternative observed/generated by a researchers and assumed to be chosen by the traveller

Within travel choice modelling and for the estimation of parameters in utility functions, the generated choice set of alternatives has to fulfill some conditions, which are:

- The observed chosen alternative needs to be included in the generated choice sets
- Choice sets should not include dominant alternatives
- Choice sets should show sufficient variation in choice attributes
- Generated choice sets need not be exhaustive, but rather be a subset of all relevant alternatives.

As a summary, Figure 3-5 shows the relationship between the observed/generated choice sets from the researcher's perspective and the choice sets from a traveller's view.

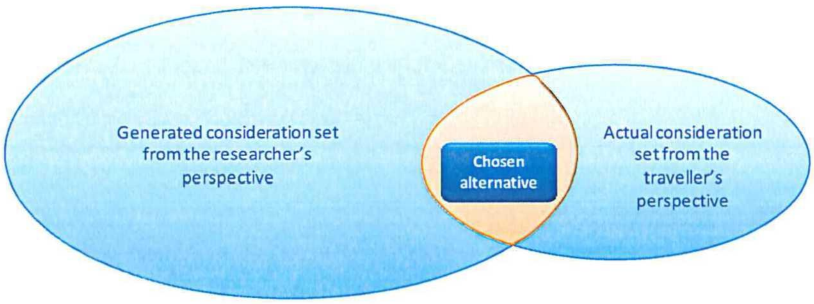


Figure 3-5 Relationship between researcher's and traveller's choice sets

3.4.2 Attributes related to each alternative: X_{iq}

Identification of attributes

Now that choice sets of alternatives have been formulated, it is necessary to find out what characterizes the alternative diversity. In fact, alternatives are described by their corresponding attributes.

According to Benckendorff (2006), identified attributes can be classified into three groups based on their relative importance, namely basic attributes, pivotal attributes, and supplementary attributes. The basic attributes represents the basic level of service. Hence, from a practical perspective, it includes all the information about flights, schedules, and booking functions needed to satisfy the travel demand. Pivotal attributes will enhance satisfaction if delivered and will cause dissatisfaction if not present. An example of such attribute is the management of frequent flyer programs which might result in frustration if unavailable (and particularly if offered by competitors). At last, supplementary attributes are attributes that may lead to satisfaction but will not cause dissatisfaction if they are not present, e.g. the value added service (legroom, entertainment facilities, food choice) falls into this category.

Based on previous literature of traveller's behaviour for airline choice (Proussaloglou & Koppelman 1999, Theis *et al.* 2006, Garrow *et al.* 2007, Hess 2007, Park 2007), Table 3-3 gives an insight into the attributes that might be taken into account by travellers in their itinerary choice. Notice that the airline choice only depends on attributes implicated with the airline, whereas in the itinerary choice the attributes related to the airport play an important role as well.

Table 3-3 Classification of all attributes playing a role in air travellers' itinerary choice

Basic	Pivotal	Supplementary
Flight price	Availability for use of Frequent flyer program	Entertainment
Departure time	Switching costs	Food choice
Arrival time	Passenger satisfaction	Legroom
Total time in air	Airline (airline image)	Service expectation
Total trip time	Airport Location	Type of aircraft

Number of stops/transfer time	On time performance	Travel insurance
Egress time Airport/city centre	Reliability (on time performance)	Accommodation
Egress Price Airport/ city centre	Egress mode choice	Seating comfort

In choice modelling, an increase in the number of attributes will have an important impact on the model size, though decision makers require a minimum number of attributes on which founding their choice. Foerster (1979) explains that individuals rank the attributes in order of decreasing importance and selects the alternative with the most preferred performance on the first attribute. Also if alternatives are identical in terms of this attribute, the decision maker would repeat the procedure based on the second attribute and so on, until a single alternative is selected. Obvious is, that a trade off has to be made when generating the choice model. On one hand a high number of attributes increases the model dimensions (DeShazo & Fermo 2002, Hensher 2004b, Caussade *et al.* 2005), but on the other hand a minimum amount of attributes are needed for the decision maker to make his choice. Relating to the amount of number of alternatives, the literature stays loose about the subject. First Louviere et al (1997) argued that increasing the number of attribute associated with a given number of alternatives will not significantly affect choice outcomes. Secontly, more recent studies showed that the number of attributes do have some impact in the variance of the error term (Dellaert *et al.* 1999, DeShazo & Fermo 2002, Hensher 2004b). Based on this last outcomes and having in mind that it has an impact on the model dimensionality, there has been chosen to select a subset of attributes among all the above mentioned.

Those selected attributes are all the standard attributes that are necessary to make a choice when booking a trip online (departure time, price, airline, etc). Added to those attributes, are the attributes related to the effects of airport location, which might have any influence on travellers' itinerary choice. Regarding to the Amsterdam-Barcelona case, the retained attributes are listed here below; observable is that the four first attributes belong to the flight characteristics (AAS to arrival airport) whereas the three last one are feature of the ground transportation from the arrival airport (Barcelona or Girona) to Barcelona city. In this case, the attribute 'Airport location' is seen as the distance between the arriving airport to the final destination. Notice that the chosen attributes mostly belong to the 'basic' category. Attributes of other category would be important if the analyses were focussed on airline choice. Attributes as entertainment or food choice would then be essential to analyse travellers' preferences in airline choice. Given the focus of this research on itinerary choice and due to the restricted numbers of attributes that can be included into the experiment, other attributes than those retained seem not having an important impact on travellers' itinerary choice.

- Airline (AIR)
- Flight price (PR)
- Departure time (DPT)
- Transfer time (TRT)

- Egress price (EGP)
- Egress time (EGT)
- Airport location (DEST)

3.4.3 Basic choice model

The complete specification of the utility functions of the choice model can be determined from the previously described alternatives and corresponding attributes. The model consists of a choice set C_{iq} with i alternatives, each composed of six attributes. The utility function of the choice model can be defined as:

$$U_{iq} = \underbrace{\beta_{AIR} \cdot AIR_{iq} + \beta_{PR} \cdot Pr_{iq} + \beta_{DPT} \cdot DPT_{iq} + \beta_{TRT} \cdot Tr_{iq} + \beta_{egp} \cdot EGP_{iq} + \beta_{egt} \cdot EGT_{iq}}_{V_{iq}} + \varepsilon_{iq} \quad \forall i \in C_q \tag{3.4}$$

Where β represents the estimated parameter, and ε_i the random error component. AIR will be treated as a categorical variable. Furthermore, other attributes may be expressed using dummy coding to include possible non-linear effects in estimation later on (see Chapter 6).

As can be seen from the utility function (3.4) each alternative i is composed of an unobserved part ε_i, and therefore some assumptions are needed concerning this random error term. Based on the assumption that the random terms ε are independently and identically distributed (i.i.d) following an extreme value type I (EVI) distribution, McFadden (1974) proposed the so-called multinomial logit (MNL) model. The choice probabilities of each alternative i from choice set C_q can be calculated as:

$$P_{iq} = P(i \mid C_q) = \frac{e^{V_{iq}}}{\sum_{j \in C_q} e^{V_{jq}}} \tag{3.5}$$

The MNL model is the most widely used discrete choice model due to its simple mathematical structure and ease of estimation. However, some of the assumptions made on ε are rather restrictive have lead to the development of extension models to the MNL model, such as the mixed logit (ML) model.

ML is a highly flexible model that can approximate any random utility model (McFadden & Train 2000). It obviates the three limitations of standard logit by allowing for random taste variation, unrestricted substitution patterns, and correlation in unobserved factors over time (Train 2003). Whereas in MNL models, the point parameters β are estimated, in ML models, the parameters of the distributions of the parameters are estimated. Hence, according to Train (2003), the expected choice probabilities are not anymore as described in equation (3.5) and may be calculated as following:

$$E(P_{iqs}) = \int_{\beta} \frac{e^{(V_{iqs})}}{\sum_{j} e^{(V_{jqs})}} f(\beta) d\beta \; , \tag{3.6}$$

where f is a multivariate distribution function over vector parameters β. The estimation methodology of these models can be found in Appendix C.

Important is to remember that this experiment is an unlabelled experiment, and therefore all alternatives consist of generic parameters. However, when estimating the model, alternative specific constants will be added to the model. This is a common way to prevent bias in the parameter estimates due to the ordering of alternatives in each choice situation (see Chapter 9).

3.5 Choice model application

By applying discrete choice methods, the aim is to estimate the parameters in the above formulated utility function (3.4). A stated choice (SC) survey will be conducted in order to obtain the choice data from respondents. As illustrated by Figure 3-6, the next part of the report will describe the generation of the SC experiment and in the part after that, the model parameters will be estimated based on the SC data.

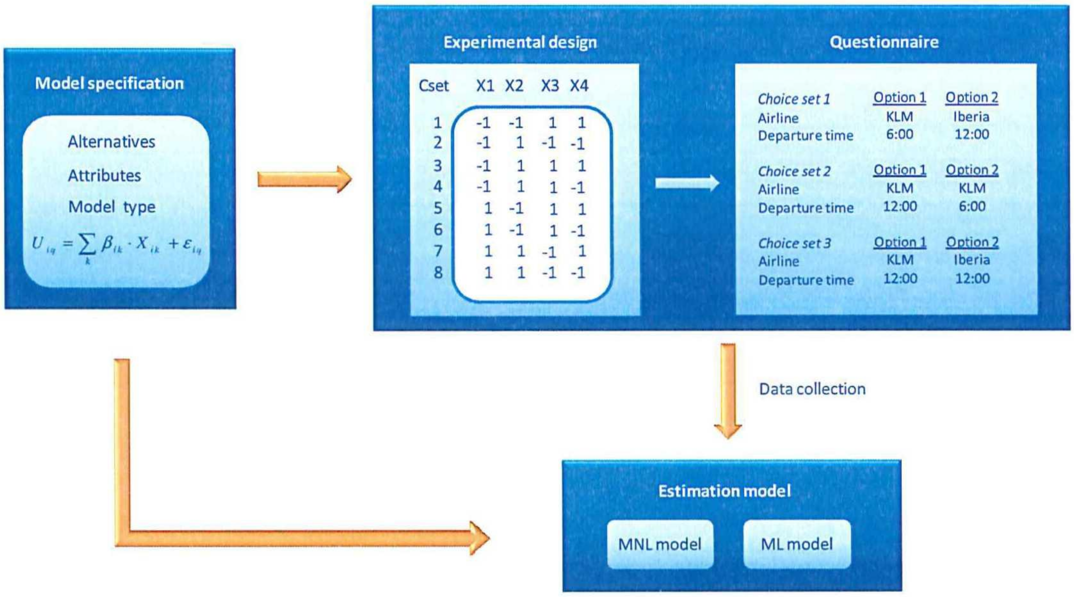


Figure 3-6 Steps for conducting a SC experiment and estimating the models

Part II – Designing Stated Choice Experiments

In this second part of the thesis, a stated choice experiment will be set up. After designing experimental designs, a web survey will be conducted and data will be collected.

Chapter 4 Generating Experimental Designs

In this chapter a Stated Choice (SC) experiment will be set up. Bliemer & Rose (2006) define a SC experiment as presenting a sample of respondents with a number of hypothetical scenarios, consisting of a universal but finite number of alternatives that differ on a number of attribute dimensions. In this research the alternatives consist of trips from AAS to Barcelona city. Figure 4-1, shows an example of such alternatives.

KLM								
	From	To	Departure	Arrival	Stops	Travel Time	Price	Choose this ticket
Flight	Amsterdam	Barcelona Airport	12:00	14:00	0	2 hr 00 min	€ 100	
Ground Transportation	Barcelona Airport	Barcelona City				0 hr 30 min	€ 3	

Transavia								
	From	To	Departure	Arrival	Stops	Travel Time	Price	Choose this ticket
Flight	Amsterdam	Girona Airport	18:00	21:00	1 - (1 hr 00)	3 hr 00 min	€ 100	
Ground Transportation	Girona Airport	Barcelona City				1 hr 20 min	€ 12	

Figure 4-1 Hypothetical scenario consisting of two alternatives

The attribute levels (such as departure time 6:00 or price €100) in each choice situation in the survey come from an underlying experimental design. As presented in Figure 4-1, an experimental design consists of a matrix of numbers where each row of the matrix represents a single choice situation, and each column an attribute. A combination of several attributes characterizes an alternative. The numbers in the matrix correspond to the different levels assigned to the attributes in each choice situation. When generating an experimental design, the analyst typically uses design coding (e.g. 0, 1, 2) or orthogonal coding (e.g. -1,0,-1) to code the different attribute levels. This coding is not necessarily the same coding that respondents will face in the survey. As such, after having generated an experimental design, the analyst may have to convert the design coding into the real level values.

There are many different ways of assigning combination of attribute levels to the attributes in each situation. Hence, the purpose of the experimental design is to determine these combinations in the best possible way in order to gain the maximum amount of information out of it. Bliemer and Rose quoted different questions that should be answered before setting up the experimental design. These questions are the follow:

- How many alternatives are included in each choice set?
- How many levels per attribute are used?
- What are the attribute level ranges?
- What type of design should be used?

In order to attain the research objectives these above mentioned questions will be discussed in the following of this chapter.

4.1 Number of alternatives

An important aspect is how individuals deal with the amount of information that they are asked to evaluate. Therefore, the number of alternatives that included in each choice situation plays a critical role.

Realism in travel choice modelling arises from the fact that respondents are asked to undertake similar actions as they would do in real markets. However, for an individual respondent, even knowing that SC experiments are hypothetical scenarios, realism may be lost if the described alternatives do not realistically portray that respondent's experiences or alternatives are considered as not credible (Rose & Hensher 2005). Furthermore, some researchers suggest that an increase in choice set complexity will compromise choice consistency (Heiner 1983, DeShazo & Fermo 2002). Heiner argues that increasing choice complexity would widen the gap between an individual's cognitive ability on the demand of decision; which would lead to a restriction of the range of decisions considered. Also, DeShazo and Fermo showed in their study (2002) that on from behavioural perspective, an increase in the quantity of information provided increases the variance with which individuals make their choices. However, from a statistical perspective, if one increases the number of alternatives in a choice set up to a threshold number, the variance decreases and then increases. Therefore, it is important that choice sets have to be generated including a strict definite number of alternatives such that the individual does not get overloaded with information which could result in some losses in the estimation accuracy.

Despite all the completed studies about the impact of design dimensionality and number of profiles on behavioural response of respondents (Arentze *et al.* 2003, Hensher 2004a, Hensher 2004b, Caussade *et al.* 2005), Rose et al. (2007) argued that no evidence has been found that the number of choice tasks had more than a marginal impact upon the behavioural responses of respondents. In 2005, Caussade et al. tested some hypotheses concerning the influence of each design dimension in a SC experiment and concluded that if the designs were ranked in terms of numbers of alternatives, those with four alternatives would come first, followed by the five alternatives and at last the one with three alternatives (Caussade *et al.* 2005).

When comparing, for the same trip purpose, different Online Travel Agents (OTA), strikingly is that all of them show (on each page) different number of alternatives (e.g. Vliegwinkel 8 alternatives, Ebookers 12 alternatives, Cheaptickets 8 alternatives, Budgetair 6 alternatives...).

Since in a SC experiment the same respondent faces multiple choice situations, and because the experiment is unlabelled, it is not necessary to show all alternatives in one screen, but rather can be distributed over multiple choice screens. Hence presenting the respondent 2, 3, or more alternatives per choice situations would have no impact on the parameter estimates. Choosing for presenting the respondents a high number of alternatives in each choice situation (as Ebookers or Cheaptickets) would have a large impact on the design dimensionality. Therefore a trade-off has been made between design dimensionality and realism according to the OTA's, and there has been chosen to generate choice sets in which each respondent faces five different alternatives.

4.2 Attribute level refinement

4.2.1 Number of attribute levels

As seen previously, the numbers in the experimental design correspond to the different levels assigned to an attribute. Essentially, each attribute will have two or more levels. If nonlinear effects are expected for a certain attribute, then a minimum of three attribute levels are needed in order to estimate these nonlinearities.

Due to the property of balancing the attribute levels (see section 4.2.3), it is wise not to mix too many different numbers of levels, as this would increase the number of choice situations. Furthermore, choosing for all even or all odd numbers of levels will typically restrict the number of choice situations needed.

Different studies (Dellaert *et al.* 1999, Caussade *et al.* 2005) proved that increasing the number of levels had a negative effect on the error variance that becomes higher. 6 airlines have been included in the survey (Air France, KLM, Iberia, Vueling, Transavia, and Easyjet), hence for the attribute AIR six levels are used. For all the other attributes three level are used. By choosing 6 levels for the airline attributes, and 3 levels for all the other attributes, it enables nonlinear effects estimations, but also keeps the number of levels limited for lower error variances in estimation. Additionally combining attributes having 3 and 6 levels (where 6 is a multiple of 3) facilitates the finding of a relative small design.

4.2.2 Range of attribute levels

According to Bliemer and Rose (2006), using a wide level range is statistically better than using a narrow range as this leads to more reliable parameters estimates with a smaller standard error. On the other hand, when generating a SC experiment, despite alternatives being completely hypothetical, better results are obtained from a behaviours point of view if the combinations of levels in each alternative have some realism. Therefore, choosing for a wider range may only be statistically better, but not in practice. For example, if the price range is really wide (€1-€1,000), this might have a dominant effect on traveller's choice regarding to other attributes. Actually travellers evaluate levels of one attribute against levels of other attributes by weighting and trading off all attributes. Thus, a trade-off exists between the statistical preference for a wide level range and the practical considerations that may limit the level range.

In this research the airport location has to have a decisive role in traveller's airline choice. For this reason the price attribute, which is usually the most important factor in travel choices (Loo 2007), has been assigned a relatively narrow level range. This will avoid the price parameters having a too dominant influence in travellers' choice and will permit more reliable estimation of the influence of other attributes. Table 4-1 illustrates all the retained attributes and their respective levels and range.

Table 4-1 Attribute levels and level range

Attribute	Level	Unit	Range
Airline	6	-	Airline name
Flight price	3	€	(50-100)
Departure time	3	Hours	(6:00-18:00)
Transfer time (number of stops)	3	Minutes	(0-120))
Egress price	3	€	(1-5) or (9-15)
Egress time	3	Minutes	(20-40) or (60-100)

Observable are the two different level ranges assigned for egress price and egress time. The reason is that, egress price and egress time may change according to the airport at which travellers land. In this case, from Barcelona airport to the city, travel time varies between 20 and 40 minutes, and egress price alters the values between €1 and €5. From Girona airport to the city of Barcelona, travel time is expected to be between 60 and 100 min whereas travel costs are between €9 and €15.

Noticeable is that when generating the different designs, a destination attribute has to be added to the model such that all egress time and price attributes are observed an equal numbers of times. The destination attribute (with 2 levels) will only be needed for the creation of the design, and will not be retained into the model estimations as it already is correlated with egress price or egress time.

4.2.3 Attribute level balance

The attribute level balance property means that each attribute level appears an equal number of times for each attribute in the experimental design, such that high levels and low levels are all represented well. However, application of this property is mitigated. On one hand, according to Bliemer and Rose (2006), the attribute level requirement is imposing constraints on the problem of minimising the design's efficiency error, and more efficient designs may be found when this assumption is not considered.

On the other hand, attribute level balance ensures that parameters can be well estimated on the whole range of levels which is better than having data points at only some levels. For this reason, attribute level balancing is generally considered as desirable when conducting experimental designs. In this research on air travel choice, the generated designs will maintain the attribute level balancing property. It can be perceived that, to obtain an attribute level balanced design, the number of choice situations needs to be a multiple of the number of attribute levels. In other words, the number of choice situations needs to be divisible by 3 and 6. This is the reason by mixing numbers of levels may lead to large designs.

4.3 Maximum and minimum number of choice sets

When generating an experimental design a minimum number of choice situations are required. This is required to estimate the models properly (see chapter 9). The minimum number of choice situations required equals the degrees of freedom of the model that will be estimated. A degree of freedom is defined as the total number of parameters (excluding the constants), plus 1. All constants are accounted for in that added degree of freedom (Bliemer & Rose 2006). Here, the smallest number of choice situations needed to satisfy this property is 11+1=12.

The maximum number of choice sets can be found by generating a full fractional design. Next paragraph discusses how to conduct such kind of designs.

4.4 Types of designs

4.4.1 Full fractional and fractional factorial designs

Referring to the literature (Hensher & Louviere 1982, Bliemer & Rose 2006), several experimental designs can be considered. The type of generated design mostly depends on the preferred statistical properties, on the available information and on the preferred size of the design. According to these three points, two sorts of designs can be constructed, namely full factorial designs or fractional factorial designs. Full factorial designs are designs in which all possible choice situations are enumerated i.e. each possible combination of the attribute levels. In this research this would mean that the total number of combination is $6^5 \cdot 3^{5x5} \cdot 2^5 = 4{,}89 \cdot 10^9$ which is an incredibly high number of combinations. Obviously, it is not feasible to let respondents face all these choice situations. In order to reduce the number of choice situations, different possibilities exist: it may be possible to block the design such that each respondent does not need to face every choice situation, or to create a fractional factorial design. Blocking will not help, as there are just too many choice situations. Hence a fractional factorial design will be created.

Fractional factorial designs are designs that only use a subset of choice situations from the full factorial design. This is mostly the easiest way to reduce the number of choice situations for each respondent. Again, different types of fractional factorial designs exist. So-called orthogonal designs and efficient designs are the main types. Also the aim of this research is to compare those two design types. This is discussed in the next section.

4.4.2 Comparing designs

Estimating data coming from different design types lead to different parameter estimates. One of the objectives of this research is to determine which type of design results in the most reliable parameter estimates. In order to make the comparison possible, it is necessary that all the designs have been observed an equal number of times such that each design has the same number of observations (see Chapter 6).

In order to analyse which design results in the most reliable parameter estimates, it suffices to compare the standard errors and the t-ratios of the parameter estimates for different sample sizes. Because the parameter estimates are asymptotically converging to one single point, the main difference that can be observed between the designs is the rapidity with which useful and reliable information can be obtained. As an example, Figure 4-2, illustrates two different designs (design 1 and design 2). Here, the figure shows that design 2 needs much more observations to attain the same outcomes of design 1.

This means that for a certain amount of observations, design 1 results in better and more reliable parameter estimates than design 2.

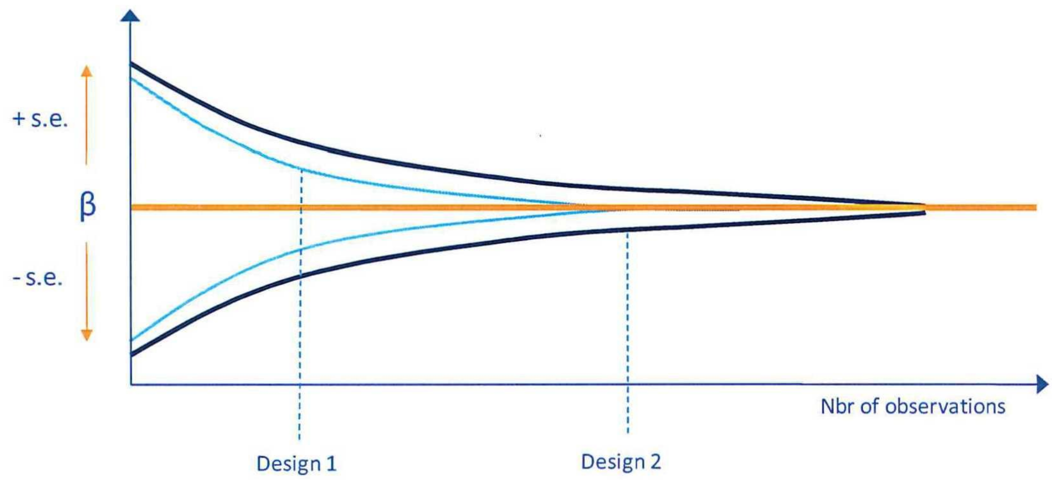


Figure 4-2 Convergence of the parameter estimate β in function of the sample size

The next coming chapter describes the most well-known design, namely the orthogonal design. Later on, Chapter 6 will discuss the efficient design.

Chapter 5 Orthogonal Designs

For a long time orthogonal designs have been used in the experimental design literature (Hensher & Louviere 1982, Proussaloglou & Koppelman 1999, Adler & Golany 2001, Caussade *et al.* 2005, Grammig *et al.* 2005, Garrow *et al.* 2007). However, despite the upcoming new techniques, leading to better reliable estimate outcomes, orthogonal designs still remain the main form of design used today (Bliemer & Rose 2006).

5.1 Definition of orthogonality

According to Bliemer and Rose, the definition of orthogonality is as follows:

"A design is said orthogonal if it all parameters are independently estimable. This translates into the definition that the attribute levels for each column in the design need to be uncorrelated."

An orthogonal design requires at least equal distance between the attribute levels. If attribute levels are not at equal distance from each other, orthogonality may be violated because of the apparition of correlation between attribute levels. Hence, the levels are assumed equally spread over the attribute level range, e.g. price will have levels €50, 75, and 100.

5.2 Generating an orthogonal design

5.2.1 Number of choice situations

The most difficult part in generating a fractional factorial orthogonal design is to find a design that satisfies design degrees of freedom, attribute level balance and orthogonality. Regarding to the airline model (equation 4.1), an orthogonal design has to be found for seven attributes (6+destination) having respectively six, three, three, three, three, three, and two levels. Concerning the degrees of freedom requirement, the smallest number of choice situations possible to satisfy the property equals 12, assuming dummy-coded attributes AIR and Dep.

Due to the attribute level balance property, the number of choice situation should be divisible by six, three and two. Hence the minimum number of choice situations that respects both properties equals 18, but unfortunately an orthogonal design, with such a number of choice situations, does not exists (or has not been found). Even orthogonal designs with 36 or 54 choice situations could not be found. This results that, only a design including 108 choice situations as the minimum number of choice situations could be found. This study shows that it is not always easy to find an orthogonal design satisfying all the required conditions. Additionally, adding an attribute or an attribute level may lead to much larger designs. In the worst case, no orthogonal design can be found.

5.2.2 Design blocking

When an orthogonal design has been found, it may occur that the design is still too large to give all the choice situations to a single respondent. By using the blocking method, the design is split into several smaller designs and a single respondent faces only a subset of choice situations from the orthogonal design. Each block by itself is not orthogonal, only the combination of all blocks is orthogonal.

Nevertheless, attribute balance is respected in each block such that respondents do not only face attributes with only high or low levels.

The orthogonal design of 108 choice situations found before, has been blocked in 18 blocks by determining an extra orthogonal column with 18 levels. Hence, each respondent will face six different choice situations wherein five different alternatives will be presented. Also a minimum of 18 respondents is required to fulfil the entire design constituted of 108 choice situations.

While orthogonal designs can only be created manually for the smallest problems, the final design in this report has been created using a software programme called Ngene, which automatically generates a design once the model specifications have been specified (number of alternatives, number of attributes, attribute levels, number of choice sets, number of blocks, utility function, etc). The syntax used to compute the design and the final created orthogonal design can both be found in Appendix A.

5.3 Orthogonality issues

One of the reasons why orthogonal designs are so much liked is because of their property of uncorrelated attribute levels which enables estimation of unbiased parameters. However, orthogonality is easily lost while collecting the data. Furthermore, the property of uncorrelated attribute levels is important for linear models and does not translate to nonlinear model such as the MNL model.

Firstly, it is important to understand that parameters are estimated from the collected data set and not from the underlying experimental design. In case of non-response, some choice situations will be missing which will lead to a loss of orthogonality within the data. Furthermore, if the design is blocked, such that the respondents do not have to face all choice situations, and some respondents do not correctly fulfil the survey, some blocks will be missing and the data will loose its orthogonality.

Secondly, if the attribute levels are not equidistant, as mentioned before, the design will lose its orthogonality once the orthogonal design coding is replaced with the real labels.

Thirdly, orthogonality in the data only holds if linear effects are to be estimated. If nonlinear effects using dummy or effect coding are estimated, the corresponding data is no longer orthogonal.

The above demonstrated properties of orthogonality may not be that important for discrete choice models and will likely be lost in the data anyway (due to dummy-coding). In the next chapter, efficient designs will be introduced, which theoretically should outperform the orthogonal design in terms of the reliability of the parameter estimates.

Chapter 6 Bayesian Efficient Designs

6.1 Introduction

A characteristic of orthogonal designs is the non-correlation structure between the attribute levels of each choice alternative. However, because of the non-linearity of discrete choice models (Train 2003), it is the correlation of the differences in the attribute levels which should be of concern when generating designs. Neglecting orthogonality results in an aim that does not try to minimize the correlation in the data (as in orthogonal designs), but tries to result in data that generates parameters estimates with the smallest possible standard errors (i.e. the square roots of the diagonal elements of the asymptotic variance-covariance (AVC) matrix).

By conducting efficient designs, the aim is to generate designs that (i) improve reliability of the parameters estimated from SC data, (ii) reduce the sample size required to produce a fixed level of reliability in the parameter estimates, and (iii) can rule out dominant alternatives in choice situations as much as possible (Bliemer *et al.* 2007). For example, with orthogonal designs, it might be possible to face a choice situation where one of the alternatives is dominant, e.g. flight 1, non-stop flight, €50, against flight 2, 1 hour transfer, and €100. Obvious is that every individual would choose the first flight, and the analyst would gain no additional information out of this choice situation. When generating efficient designs, dominant alternatives are ruled out due to some degree of utility balance[3] (Huber & Zwerina 1996), which leads to the construction of more efficient choice situations.

6.2 Definition of efficiency and efficiency measures

6.2.1 Efficient designs

Bliemer and Rose (2006) define efficient designs as follows:

"An experiment design is called efficient if the design yields data that enables estimation of the parameters with as low as possible standard errors".

The asymptotic[4] variance-covariance (AVC) matrix is an output of the estimation procedure. The roots of the diagonals of the matrix are the so-called standard errors. It turns out that these standard errors can already be determined for a given design before conducting the survey (Huber & Zwerina 1996). In order to minimize all the standard errors simultaneously, different efficiency measures have been proposed and will be discussed in the next subsection.

---

[3] A design is more utility balanced if there are no dominating alternatives in the choice situations, and all the alternatives have more or less an equal observed utility.

[4] The term asymptotic refers to the fact that it is consistent in large samples, or it is representative as an average for all samples when the survey would be repeated many times.

6.2.2 Efficiency measures

The efficiency of a design can be derived from the AVC matrix. The most widely used efficiency measure is called the D-error, which takes the determinant of the AVC matrix Ω scaled to the number of parameters K (Huber & Zwerina 1996). The D-error is given by:

$$D\text{-error} = \det(\Omega)^{1/K}. \tag{6.1}$$

A design with the lowest D-error is called D-optimal. Other efficiency measures have been proposed, such as the A-error, which minimizes the sum of the diagonal elements of the AVC matrix. However, this measure is sensitive to scaling of the attribute levels, hence the D-error is the preferred efficiency error measure.

The AVC matrix Ω dependents on the parameter estimates (see section 6.3). Because these parameter estimates are unknown, some prior parameters, used as best guesses of the true parameters (see section 6.4), are used to construct Ω . As these priors are never known beforehand with certainty, Sándor and Wedel (2001) propose to use prior parameter distributions instead of fixed values. Different sets of prior parameter values lead to different D-errors as in equation (6.1). Therefore, assuming probability distributions of prior parameter values means that the D-error will also be stochastic. Following Sándor and Wedel (2001), the mean (expected) D-error will be used as efficiency measure, called the Bayesian D-error. The Bayesian D-error is given by:

$$\text{Bayesian } D\text{-error} = \int_{\beta} \det(\Omega(\beta \mid X))^{1/K} f(\beta \mid \mu, \sigma^2)\, d\beta, \tag{6.2}$$

where, f is a multivariate normal probability density function of the parameters. Other distributions can be used, however the normal distribution is the most widely used.

This integral can only be approximated by simulation, increasing the computation time for evaluating the efficiency of a given design. Since thousands if not millions of designs need to be evaluated for determining an efficient design, computation time is an issue. By using smart techniques such as Gaussian quadrature and Halton draws, the time for calculating the Bayesian D-error can significantly be reduced (Bliemer *et al.* 2007).

Bliemer and Rose (2005) introduced another efficiency measure related to the required sample sizes to estimate each parameter to a statistically significant level. Bliemer and Rose argue that, if some parameters need much higher samples sizes than others to be estimated, it may be better to focus on the parameter that has difficulties to be estimated with high levels of statistical significance. This measure looks at the asymptotic t-ratio for each parameter, extracting the theoretical minimum sample size needed for all parameters to be statistically significant. The design is then optimised for sample size, and the design with the lowest so-called S-estimate needed is the most efficient design.

Since the S-estimate only focuses on the minimisation of one parameter, instead of all the parameters of the matrix diagonal, there has been chosen to use the Bayesian-D-error as efficiency measure for this research, which is the most widely used measure since 1994 (see (Bunch *et al.*).). The AVC matrix Ω is input for this measure and will be derived in the next section.

6.3 Derivation of AVC matrix

As illustrated by Figure 6-1, data required to estimate model parameters is obtained by performing a survey which depends on the underlying experimental design.

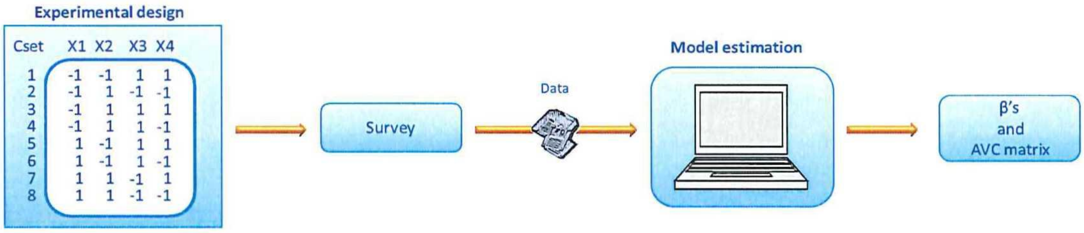


Figure 6-1 Steps to obtain parameter estimates and AVC matrix

However, it is possible to determine the AVC matrix beforehand, and without conducting any survey. The derivation of the AVC matrix will be described in the section below..

As illustrated by equation (6.3), the AVC matrix is the negative inverse of the Hessian matrix, also called Fisher information matrix (Train 2003), where the Fisher matrix is the second derivative of the log-likelihood function.

$$\Omega_N(X,Y,\beta) = -[E(I_N(X,Y,\beta))]^{-1} = -\left[\frac{\partial^2 LL_N(X,Y,\beta)}{\partial\beta\partial\beta'}\right]^{-1}, \tag{6.3}$$

where $I_N(X,Y,\tilde{\beta})$ is the Fisher information matrix with N respondents, and $LL_N(X,Y,\tilde{\beta})$ is the log-likelihood function in case of N respondents.

The log-likelihood function (see Chapter 9) can be calculated as in following equation (6.4):

$$LL_N(X,Y,\beta) = \sum_{n=1}^{N}\sum_{s=1}^{S}\sum_{j=1}^{J} y_{nsj}\log\left(P_{nsj}(X,Y,\beta)\right), \tag{6.4}$$

where $Y = [y_{jsn}]$ equals one if respondent n chooses alternative j in choice situation s and zero otherwise, and $P_{nsj}(\beta)$ is the MNL probability in equation (3.5). By taking the second derivative of the log-likelihood function, Y drops out of the equation (see (McFadden 1974, Huber & Zwerina 1996)). Hence, the AVC matrix directly depends on the choice probabilities of the alternatives. Notice that each different design will yield a different AVC matrix since the choice probabilities of the alternatives will change. In order to compute this probabilities, prior values for β are needed. As mentioned before, the next section will discuss how to find those prior parameter values.

6.4 Prior parameters

A main difference with orthogonal designs is that efficient designs rely on prior values for the parameters. Adding this information enables the researcher to make the design more efficient. Knowledge of only the sign of a parameter (e.g. price has typically a negative parameter) already provides information. Since the objective of the survey is to return these values, the exact parameter values are clearly unknown. Nevertheless, several approaches exist to obtain prior parameters.

Some methods are to realise a literature research, focus groups, or to ask for expert judgement (Sándor & Wedel 2001). Another approach is to perform a pilot study. It consists of generating an SC experiment where the estimated parameter estimates will later be used as prior parameters for constructing a more efficient SC experiment. Notice that the model estimated in the pilot study has to be the same model as the one estimated in the main experiment. This method of executing a pilot study is one of the most accurate ways to find some prior parameter, and is also used in this research.

6.5 Generating Bayesian efficient designs

6.5.1 Pilot study

In order to find some prior estimates, a pilot study has been conducted. An orthogonal design, as described in the previous chapter has been constructed. The design consisted in 108 choice situations of 5 alternatives each composed of the 6 previously mentioned attributes (Air, Pr, Dep, Tran, Egp, and Egt). The 108 choice situations have been blocked into 18 blocks such that each respondent faces 6 choice situations. The design used for the pilot study is exactly the same orthogonal design as constructed in Chapter 5 (see appendix B). In total 36 respondents have participated at the survey. Hence, only two complete designs have been completed by all respondents together, but as this is a pilot study, it is enough to get some information needed about the parameters (in total 216 choice situations have been observed).

The parameter estimates of the pilot study have been estimated according to a MNL model (see equation 3.5)). Table 6-1 shows the outcomes of this pilot study.

Table 6-1 Outcomes of pilot study

Attribute	Unit	Parameter Beta	Standard error	t- value
Air France (AIR1)*	-	-.516	.323	-1.596
KLM (AIR2)*	-	.194	.293	.661
Iberia (AIR3)*	-	-.218	.298	-.729
Vueling (AIR4)*	-	.125	.294	.426
Transavia (AIR)5*	-	-.547	.310	-1.765
Price (Pr)	Euro	-.040	..004	-8.444
Departure time 6:00 (DT0)*	-	.215	.221	.974

Departure time 12:00 (DT1)	-	.614	.213	2.883
Transfer time (Tr)	Minutes	-.006	.002	-3.298
Egress price (EGP)*	Euro	-.027	.030	-.904
Egress time (EGT)	Minutes	-.018	.005	-3.265

R-squared=.190, LL=-266, * non significant parameters at 95% level.

These parameter estimates will be used as prior parameters when conducting the efficient designs. Notice that for the generation of Bayesian efficient designs, a distribution of the prior parameters is needed. Therefore, it will be assumed that all parameters are distributed according to a normal distribution $\beta \sim N\left(\mu, \sigma^2\right)$ where the mean μ can be seen as the estimation value of the parameter and where the standard deviation σ is the standard error of the parameter. If the parameter estimates are non-significant, the prior parameters have been put equal to zero with the corresponding standard error as value for the standard deviation (e.g. $\beta_{EGP} \sim N\left(0, 0.03^2\right)$).

6.5.2 Number of choice situations

The generation of efficient designs requires the same characteristics as when generating an orthogonal design. However efficient designs loose the property of orthogonality, and do not necessarily require attribute level balance. This last property has been kept as it assures that all levels are uniformly represented. Because one of the objectives of this research is to compare orthogonal designs with efficient designs, the model used for the construction of the efficient design is the same as the one used for the construction of the orthogonal design. In other words, the model consists in 5 alternatives, each composed of the 6 previously mentioned attributes and their corresponding levels (see equation 3.5).

Clearly, more choice situations result in more data per respondent, (at an increased burden), and the efficiency will automatically increase with more choice situations. However, by normalising the efficiency error, it does not make that much difference how many choice situations are chosen (Bliemer & Rose 2006). Therefore, the minimum required number of choice situations only depends on the degrees of freedom needed to estimate the model and the property of attribute level balance.

When generating the orthogonal design, the minimum number of choice situations due to the degree of freedom was 12 (assuming that AIR and Dep were dummy-coded attributes). In this case, the construction of an efficient design still requires 12 degrees of freedom. However some estimation models, other than the MNL model (mixed logit model), may require more degrees of freedom (see Chapter 9). In order to be able to estimate the parameters later on with different estimation models, the minimum degrees of freedom required becomes 18.

According to the above characteristics, two different efficient designs will be conducted. The first one consists of an efficient design of 18 choice situations, which is the minimum number of choice situations required according to the degrees of freedom. Additionally, a second efficient design will be created with the difference that it will consist of 108 choice situations, which is the same number of choice situations as the orthogonal design has. The objective by creating two different efficient designs is to acquire more knowledge not only in how better efficient designs outperform orthogonal designs, but also in the impact of sample size (18 versus 108 choice situations) on the parameter estimates.

6.5.3 Design blocking

In order to make a comparison possible between orthogonal and efficient designs, it is necessary to block the efficient design such that respondents would face the same number of choice situations independently of which design they would be faced with. Thus, the efficient design of 18 choice situations has been blocked in 3 blocks, and the design of 108 choice situations is blocked in 18 blocks, such that each respondent again faces 6 choice situations.

Both efficient designs have been generated by using the software programme Ngene, and designs have been optimised according to the minimised Bayesian D-error. The syntax used to compute the designs as well as the created efficient designs can be found in Appendix B.

Chapter 7 Construction of the web-Survey

7.1 Introduction

Given the model to be estimated and the experimental designs, the survey needs to be constructed such that observations can be collected. Different kinds of SC surveys can be conducted, such as pen & paper surveys, CAPI (computer aided personal interviewing) surveys, or internet surveys. In this research we have chosen to perform an internet survey because of its advantage compared to other methods. Internet surveys are very flexible and easy to execute (only have to send a link). Also the data is directly readily available. A disadvantage is, that the survey is only accessible by respondents with a computer and internet, which may potentially lead to a restricted sample. However, nowadays the penetration rate of the computers and internet in households is considered sufficiently high to avoid sampling problems. Hence performing an internet survey is the easiest way to collect respondents' observations. The data coming from these observations will later on be estimated (see Chapter 9). The estimation outcomes will be interpreted and more insight on attributes influencing air travellers' itinerary choice will be obtained.

This chapter will discuss how the questionnaire has been constructed after having established the model to be estimated and aftere having designed the related experimental designs. If the reader is interested in the considerations for the construction of the survey in terms of how and when to ask the right questions, then, he is advised to consult sources on this subject (see (Dillman 2000)).

7.2 Context scenario

The construction of a survey demands some specific attention as this is the only part respondents will be faced to. Respondents will have to answer the questionnaire without having any background on how the survey has been generated. Therefore, it is important that the analyst creates a decision context in which respondents will have to make their choices. For example, when choosing between two alternatives for travel choices, respondents might give different response depending on the context scenario (e.g. the choice might change if the trip has a business or a leisure purpose). After the scenario description, the respondent will be asked to observe several choice sets of different alternatives and to choose the most preferred one.

For this research, respondents have to imagine making a trip from AAS to Barcelona city. Therefore, they have the choice between flying to Barcelona airport, which is the main airport, or to Girona airport, which is the secondary airport. In both situations, additional ground transportation is needed to access the city of Barcelona. Independent of the airport destination, respondents have to assume that flights from AAS to both airports take 2 hours (direct flight). It might be that some travel time has to be added to those two flight hours in case of an possible transfer through another airport. The prices shown in the travel tickets are all single trips (one way). A screen shot of the proposed case scenario can be found in Figure 7-1.

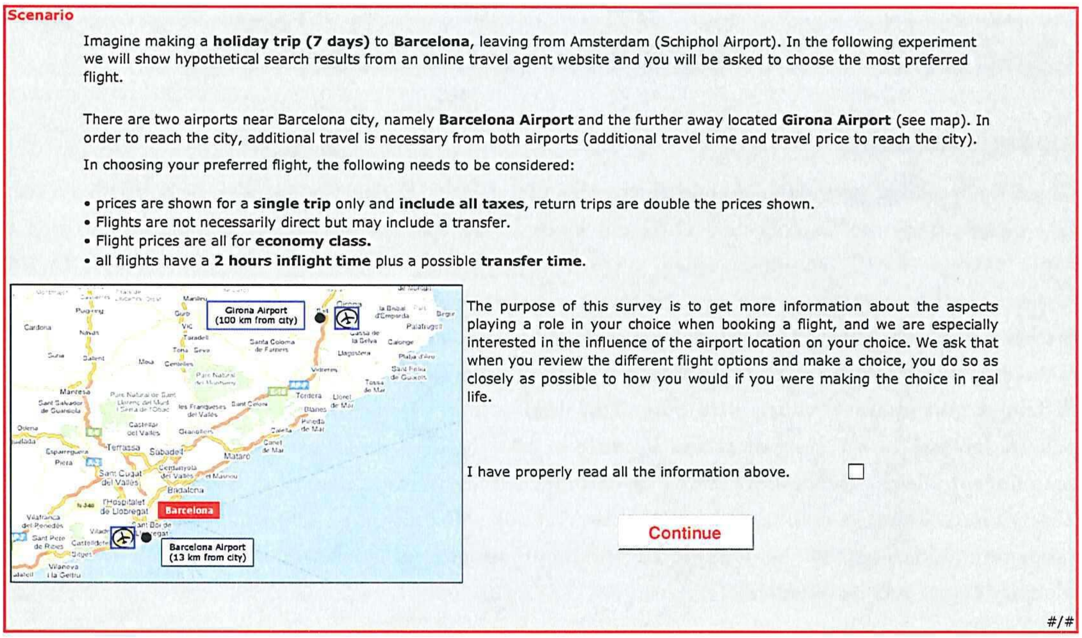


Figure 7-1 Screen shot of the case scenario

7.3 Realism of the survey

As described in Chapter 1, it has been shown that realism in the survey may affect the way how respondents may answer the survey. Even knowing that SC experiments are hypothetical scenarios, realism may be lost if the described alternatives do not realistically portray that respondent's experiences or alternatives are considered as not credible. Therefore, one of the aims of this study is to perform a web-survey that looks identically as the website of OTA's. So, respondents have the possibility, as in reality, to use a set of tools in order to facilitate them to find their preferred flight. Respondents were able to select or deselect a show button, that showed or masked the attributes in which they were interested (or not) in. Additionally to the six attributes represented in the choice model, the arrival time and the travel time have been added to the alternatives. Notice that since the in-flight time has been fixed to two hours, those variables only depend on the departure time and on the transfer time. Observable is, that at the beginning, the default setup showed all the attributes. Two additional attributes were presented but not marked by default, namely the total price and the total travel time. Those two attributes give the sum of the flight price and the egress price / flight travel time and egress travel time. Furthermore, respondents were able to rank the presented attributes on any of the showed attributes. A screen shot of those command buttons can be seen in Figure 7-2.

Attribute	Show	Sort by	Information
Flight			
Airline	☑	○	
Departure Time	☑	○	
Arrival Time	☑	○	Depature time + inflight time + transfer time
Number of Stops / Transfer Time	☑	○	Time spent waiting at the stop(s)
Travel Time	☑	○	Travel time needed from Amsterdam to Barcelona airport or to Girona airport (including inflight time and transfer time)
Price	☑	◉	Ticket price (flight only) including fees and taxes
Ground transportation			
Travel Time	☑	○	Additional price for traveling from the arrival airport to the city centre
Price	☑	○	Additional time needed to travel from the arrival airport to the city centre
Total Travel Time	☐	○	Total travel time (including inflight time and ground transportation travel time) from Amsterdam airport to Barcelona City
Total Price	☐	○	Total price (including flight price and ground transportation price) from Amsterdam to Barcelona City

Figure 7-2 Command options for finding most preferred flight

Due to its possibilities to show or mask some attributes, and to sort the presented flight on any of the attributes, the survey looks pretty similar to the OTA's websites. In order to make the survey even more realistic, the different travel alternatives are presented as a list from top to bottom. As a comparison between the survey and a real OTA, Figure 7-3 shows the screenshot of an existing OTA whereas Figure 7-4 illustrates the conducted web-survey.

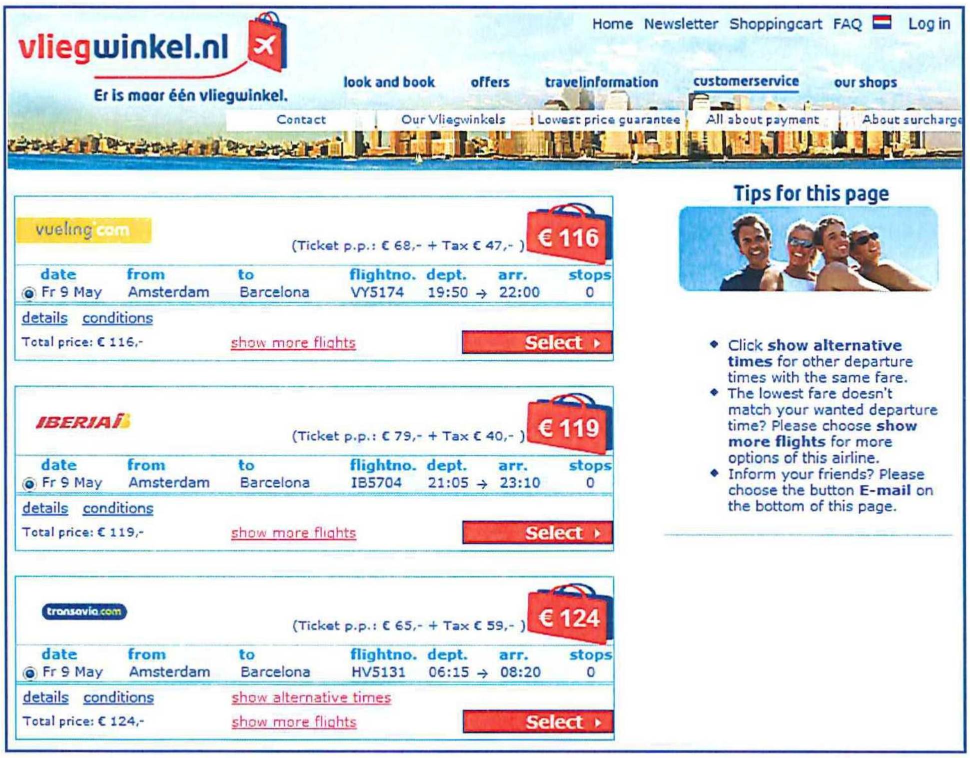
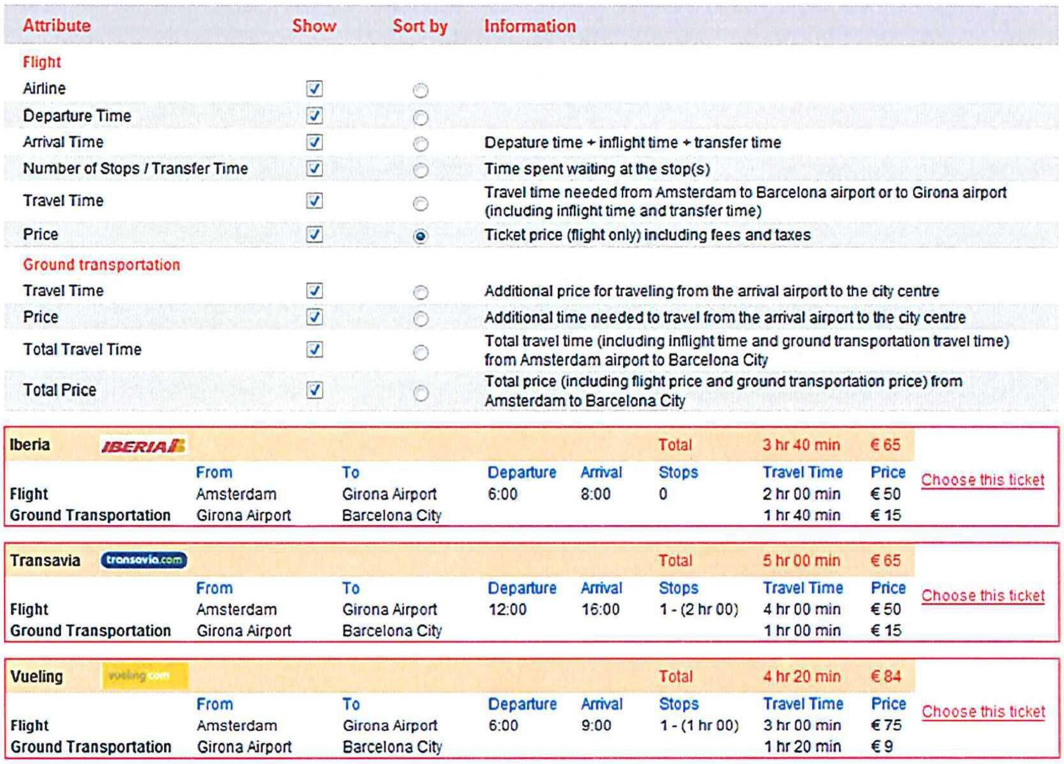


Figure 7-3 Screen shot of a Dutch OTA



Attribute	Show	Sort by	Information
Flight			
Airline	☑	○	
Departure Time	☑	○	
Arrival Time	☑	○	Depature time + inflight time + transfer time
Number of Stops / Transfer Time	☑	○	Time spent waiting at the stop(s)
Travel Time	☑	○	Travel time needed from Amsterdam to Barcelona airport or to Girona airport (including inflight time and transfer time)
Price	☑	◉	Ticket price (flight only) including fees and taxes
Ground transportation			
Travel Time	☑	○	Additional price for traveling from the arrival airport to the city centre
Price	☑	○	Additional time needed to travel from the arrival airport to the city centre
Total Travel Time	☑	○	Total travel time (including inflight time and ground transportation travel time) from Amsterdam airport to Barcelona City
Total Price	☑	○	Total price (including flight price and ground transportation price) from Amsterdam to Barcelona City

Iberia					Total	3 hr 40 min	€ 65	
	From	To	Departure	Arrival	Stops	Travel Time	Price	Choose this ticket
Flight	Amsterdam	Girona Airport	6:00	8:00	0	2 hr 00 min	€ 50	
Ground Transportation	Girona Airport	Barcelona City				1 hr 40 min	€ 15	

Transavia					Total	5 hr 00 min	€ 65	
	From	To	Departure	Arrival	Stops	Travel Time	Price	Choose this ticket
Flight	Amsterdam	Girona Airport	12:00	16:00	1 - (2 hr 00)	4 hr 00 min	€ 50	
Ground Transportation	Girona Airport	Barcelona City				1 hr 00 min	€ 15	

Vueling					Total	4 hr 20 min	€ 84	
	From	To	Departure	Arrival	Stops	Travel Time	Price	Choose this ticket
Flight	Amsterdam	Girona Airport	6:00	9:00	1 - (1 hr 00)	3 hr 00 min	€ 75	
Ground Transportation	Girona Airport	Barcelona City				1 hr 20 min	€ 9	

Figure 7-4 Screen shot of the constructed web-survey

7.4 Survey issues

7.4.1 Randomization of choice situations

As explained in Chapter 4, each row in the experimental design translates into a different choice situation with attributes having each their attribute level. Each respondent will face one specific block of a certain design, where each block consists of 6 choice situations (see Chapters 5 and 6)

In the end, if a good comparison will be done between the three designs (orthogonal design, Bayesian efficient design with 18 choice situations, and Bayesian efficient design with 108 choice situations), it is important to assure that each block within each design is observed a same number of times. The two designs of 108 choice situations have been blocked in 18 blocks, whereas the efficient design consisting of 18 choice situations has been blocked in 3 blocks. Hence, these last ones have to be observed six times as much as the other blocks of the two other designs. Finally, each block within each design is faced by respondents an equal number of times in order to obtain the same number of observations in each design. The final survey is composed of 54 blocks, where blocks number 1 to 18 belong to the orthogonal design, blocks number 19 to 36 belong to the efficient design with 18 choice situations (the 3 blocks have been repeated 6 times), and the blocks number 37 to 54 fall to the efficient design consisting of 108 choice situations.

Notice that if a respondent logs in on the survey but does not complete it entirely, its data will not be saved and the corresponding design block will still have to be observed by another respondent. Because the blocks are distributed randomly (according to the number of times they have already been observed by previous respondents), respondents do not know which block of which design they are faced to. A user id is assigned to each respondent that logs in on the survey as this permits to keep track of which block of which design has been assigned to the respondent.

An important issue when performing a survey is to randomize the choice situations. By randomizing the order of choice situations for each respondent, the respondents will view the choice situations in different order. When multiple respondents observe the same block, because respondents may learn during answering the survey, randomizing the choice situations eliminates the order bias in the experiment. In this research not only the choice situation order has been randomized, but also the order of the alternatives. As alternatives are shown from top to bottom, the top alternatives may be preferred to the ones that are listed on the bottom. Changing the order of the alternatives rules out order bias as much as possible.

7.4.2 Transformation of the design coding into attribute levels values

The way how experimental designs are coded (design coding) does not correspond to the same code that respondents will face in the survey. Obviously, the matrix full of numbers is meaningless to a respondent, and therefore, the analyst has to convert the experimental design in choice situations with attributes having real levels. Table 7-1 illustrates which attribute level value corresponds to each design code.

Table 7-1 Attribute level corresponding to each design coding

Design code	Air	Pr	Dep	Tr	EGP	EGT
0	Air France	€50	6:00	Non stop	€1* or €9	20 min* or 1 hr
1	KLM	€75	12:00	60 min	€3* or €12	30 min* or 1hr20
2	Iberia	€100	18:00	120 min	€5* or €15	40 min* or 1hr40
3	Vueling					
4	Transavia					
5	Easy Jet					

* If destination is Barcelona, other values are for destination Girona.

As an example, Figure 7-5 illustrates the changes between the design coding and the final survey.

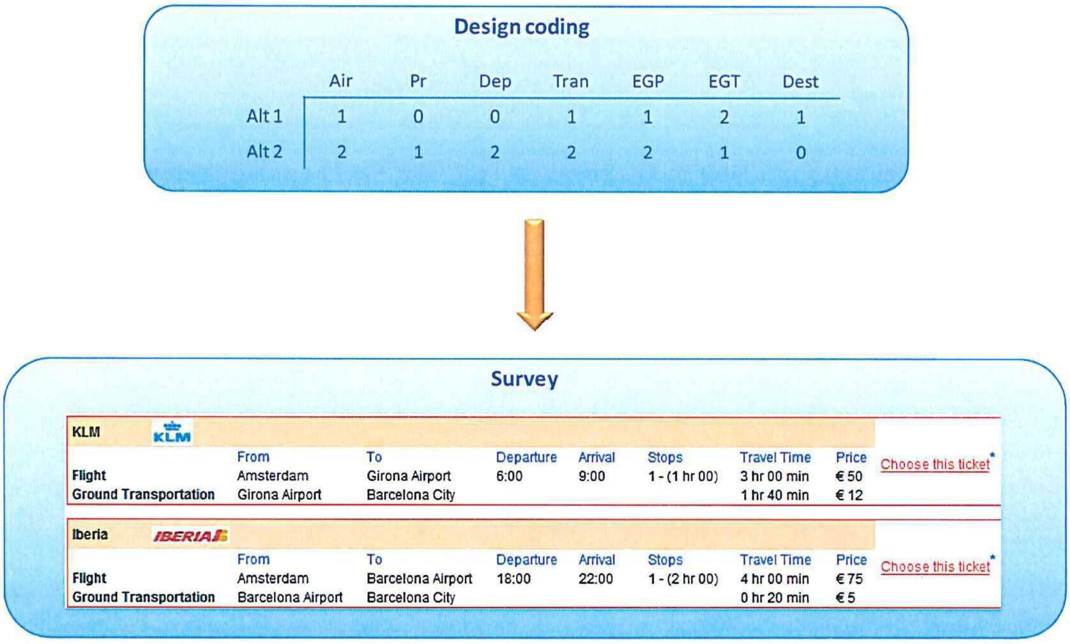


Figure 7-5 Conversion from design coding to survey

7.4.3 Sample respondents

After the construction of the web-survey, respondents have to be sampled. As the estimation outcomes will depend on the sample of respondent, it is necessary that the sample of respondents is representative for a whole population. Therefore, a nice spread of the population on gender or income is required (e.g. a full time working male will not have the same income as a part time working female, and so, might have different travel preferences). To control these issues, at the end of the survey, the respondents have been asked to answer some questions concerning their flight habits (frequent flyer programme membership, or how many times they have flown the last years, etc), and some personal questions (such as age class, gender and income).

In order to get a sample of respondents, nicely spread over the population, TeamVier, a Dutch market research company, has been asked to collaborate in this project by providing a sample of approximately 500 respondents in the Netherlands.

Given the collected data, the analysis and the model estimations can be effectuated. This will be presented in the last part of this report.

Part III - Estimations and Analyses

This last part focuses on the data estimation and on the analyses. First the characteristics of the data will be made followed by the model(s) estimations. In the end, conclusions related to the estimation outcomes will be presented.

Chapter 8 Data Statistics

As mentioned in the previous chapter, the Dutch market research company, TeamVier, provided a sample of respondents for this study. In this chapter, the collected data will be explored as well as its dispersal according to demographically aspects with the aim of showing that (i) the sample is heterogeneous and representative, and (ii) the collected data is suitable to perform the required analysis.

8.1 Collected data

8.1.1 Full data

In total 3993 emails have been sent out. In the end, a total of 6,450 observations were collected from 1,075 respondents. After excluding responses with insufficient data (uncompleted surveys), 3,384 observations, collected from 564 respondents, remained in the data set (response rate of 52%).

After analysing the data, as illustrated in Figure 8-1, it seems that 4 respondents needed more than an hour to respond the survey. Additionally, 10 respondents filled in the survey in 3 minutes or less. Even by observing the survey really quickly, it seems impossible to answer the survey rationally that fast. Because these observations (14 in total) are behaviourally irrational, these observations have been thrown out of the data set, such that estimations will not be biased. The cleared full data consists of 3,330 observations collected from 550 respondents.

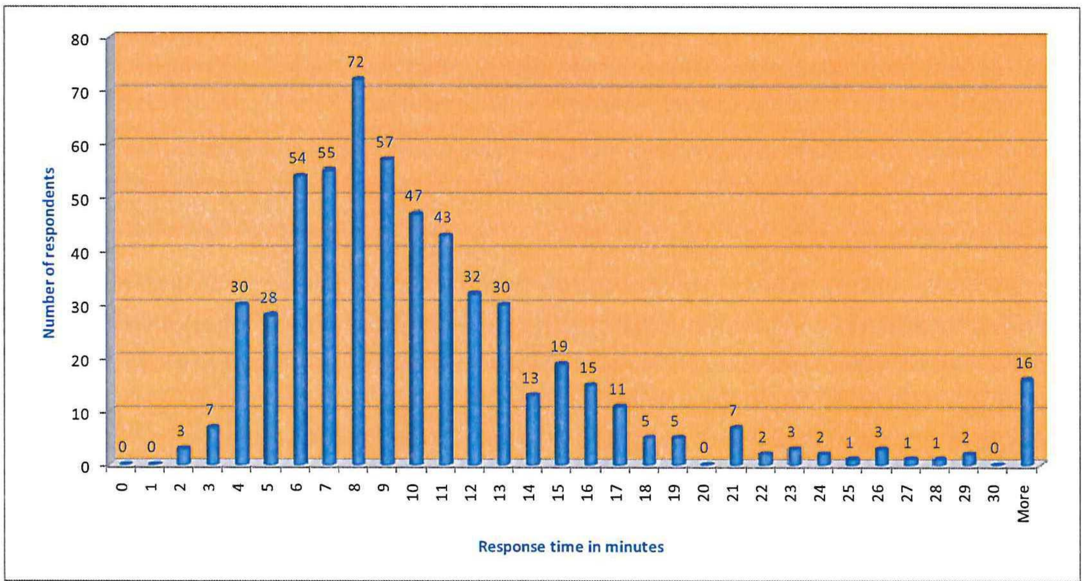


Figure 8-1 Time response frequency

Notice that the full data of 3,330 observations is composed of observations coming from all three generated designs (orthogonal design, efficient design with 18 choice situations, and efficient design with 108 choice situations). This full data will be used later on to obtain the parameter estimates of the model.

8.1.2 Specific data to each experimental design

As already mentioned in Chapter 4, comparing the designs is only possible if the number of observed choice situations is the same for all the designs. Also, it is necessary that each block of each design is observed the same number of times. Because each respondent observed one random block belonging to one of the three designs, the total number of observed blocks of each design had to be monitored, such that in the end, each block of each design would be observed an equal number of times.

Finally, the total number of respondents collected permits that each design was in total observed 8 times. Hence, for each experimental design (orthogonal, efficient 18, and efficient 108), a total of 972 observations were collected from 162 respondents.

8.2 Demographic sample characteristics

In the survey, the last questions concerned individual's background information (gender, age group, income, etc). Such typical demographical information is important for the analyst, such that its conclusions not only apply for a specific respondent's category, e.g. if the sample is not equally spread over all age groups, it might be that conclusions only apply for a certain age group of the population. Also, it is important with regard to the estimation outcomes, that the sample is properly distributed over all three designs, not that one design has been observed by only men with a high income and another design by women with a lower income. Hence, it is essential that all designs approximately have the same sample characteristics.

Figure 8-2 shows how the sample is spread over different designs. It can be seen that all three designs more or less have the same sample characteristics, especially regarding gender. Concerning respondents' age, in the efficient design with 108 choice situations, a shift occurs between groups age 18-25 to 35-45. Other age groups are equally spread over all designs. Regarding the number of respondents having a frequent flyer membership, approximately 20% of them have a Flying Blue membership (Air France and KLM), and almost none have another membership with any other airline or airline alliance (Iberia or Oneworld). The graph showing respondents' employment status indicates that only a few students answered the survey, the remaining respondents being full time, part-time, or even non working. This employment status dispersal has probably an impact on the respondents' income, which is relatively equally spread over the sample. Finally, it can be seen that only 20% of the respondents travelled with one specific airline in the last three years. This could be clarified by their loyalty towards a specific airline, but it might also be possible that those respondents only travelled once during that period (76% of the respondents flew less than 6 times during the last three years).

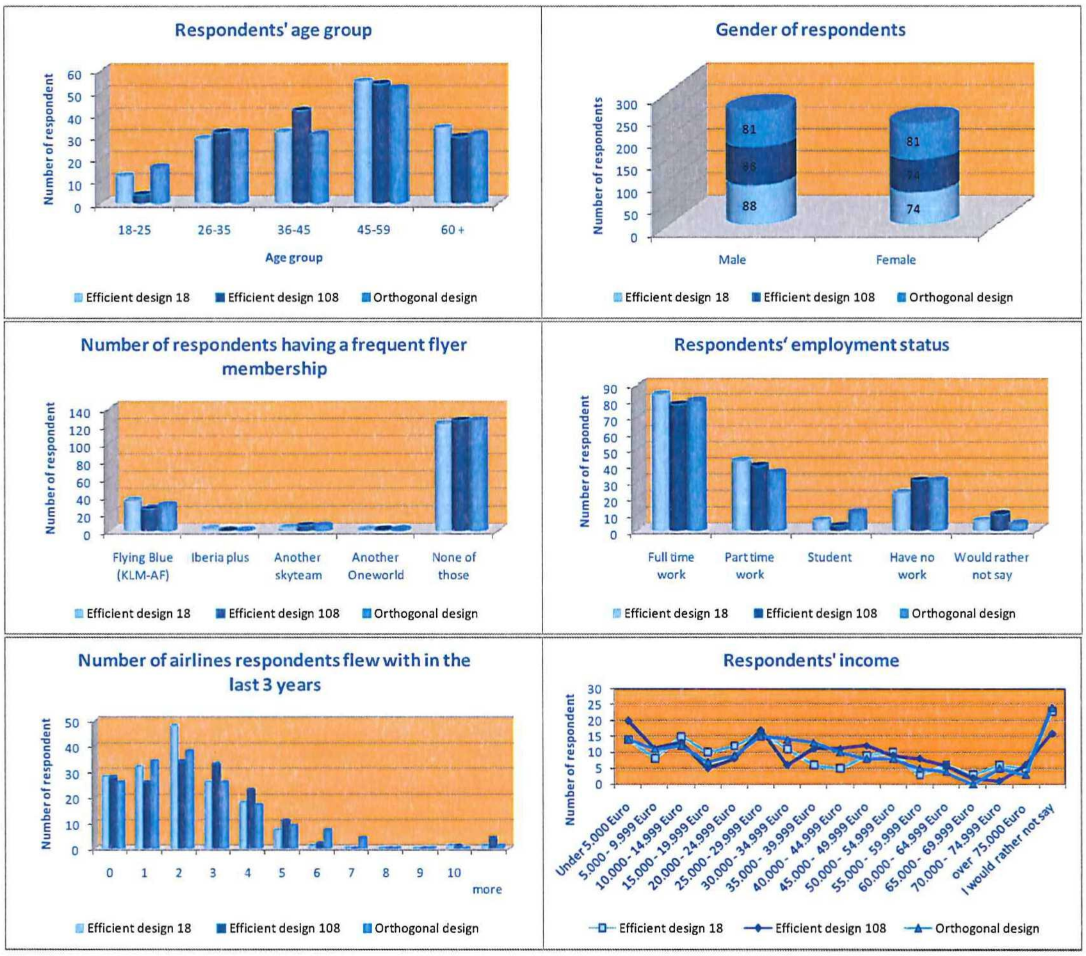


Figure 8-2 Demographic sample characteristics

8.3 Conclusion

This chapter has described the demographic characteristics of the collected data. First it has been showed that all three designs have been observed an equal number of times with each essential to compare the designs. Additionally, it has been seen that all three designs have more or less the same sample characteristics. This shows that the spread of respondents over the designs is heterogeneous as well.

Heterogeneity of the sample is essential for the interpretation and clarification of the estimation outcomes. This chapter has shown that respondents are spread approximately equally over the sample. There are no groups that are not represented which could influence the interpretation of the estimation outcomes.

Given a heterogeneous and representative data, the next chapter describes the different model types chosen to estimate the full data. Later in Chapter 10, all three generated designs will be evaluated and compared to each other.

Chapter 9 Model Estimations

This chapter provides the outcomes of different estimation models. Firstly, a simple MNL model is estimated, followed by 2 other models: another but more complex MNL model and an advanced ML panel model. All model estimations will be realised with the help of an integrated econometrics package called NLOGIT. For more information about this software package, the reader can visit the company's website (www.NLOGIT.com) or seek to literature on this software (see (Hensher *et al.* 2005)).

9.1 Simple MNL model

In Chapter 3, a basic utility function has been presented. Estimating the parameters of this utility function (see equation 3.4) can be done with a simple MNL model according to its log-likelihood function maximization (see Appendix C). As mentioned previously, some alternative-specific constants representing the weight of the ordering, have been added to the utility function (equation 3.4), this is to prevent some bias in the parameter estimates due to the ordering of the alternatives in each choice situation.

The outcomes of this MNL model are revealed in Table 9-1.

Table 9-1 Simple MNL results

Variable	Unit	β coefficient	t-value
Price (Pr)	Euro	-.0406	-30.102
Departure time 6:00 (Dep(0))	-	.2776	5.516
Departure time 12:00 (Dep(1))	-	.5914	11.497
Transfer time (Tr)	Minutes	-.0172	-37.184
Egress price (EGP)	Euro	-.0312	-4.733
Egress time (EGT)	Minutes	-.0162	-12.897
Air France (AIR1)*	-	.1357	1.790
KLM (AIR2)	-	.6178	8.524
Iberia (AIR3)*	-	-.0108	-.139
Vueling (AIR4)	-	-.2946	-3.758
Transavia (AIR5)	-	.2960	4.044
Constant 1 (C1)	-	.4845	7.421
Constant 2 (C2)	-	.3429	5.138
Constant 3 (C3)	-	.3561	5.268
Constant 4 (C4)	-	.1979	2.821
Log-Likelihood	-	-3887	

Adjusted ρ^2	-	.254

* Significance level of 5%

Most of the estimates are significant and of the expected sign. The results show that price and transfer time appear to be the most important attribute influencing travellers' itinerary choice, as they contribute the most to the utility (parameter * average attribute level). Both parameters have the highest t-value, which means they are the estimates with the highest accuracy. The fact that those parameters are the most important in travellers' itinerary choice is not a surprise since travellers mostly choose the itinerary with the shortest travel time and the lowest price. Other attributes such as departure times or airlines seem secondary attributes with which travellers' determine their preferences for a specific itinerary in case of multiple options. Because of the non-significance of some of the airlines, the NCs, Air France and Iberia, seem not to be relevant in travellers' choices and have the same attractiveness as the LLC, Easyjet. In contrast, travellers' have a high preference for the Dutch airlines (KLM and Transavia) compared to Easyjet. This can be clarified by the fact that the data has only been collected into the Netherlands. At last, Vueling seems less preferred compared to Easyjet as it has a negative parameter estimate. Concerning the departure times, travellers have a preference for travelling at noon or in the morning instead of in the evening (departure times have positive signs compared to the reference which is departing in the evening). Finally, the constants show that respondents were more likely to choose the first alternatives presented. This can be explained by the fact that respondents had the opportunity to sort the alternatives on their preferred attribute. In this case, the preferred attributes on which respondents might have sorted on are price or transfer time.

In their study, Rose and Bliemer (2006) mention that most discrete choice model estimation include covariates (i.e. interactions with socio-demographical attributes of the respondents). Including covariates into the utility function could change the value of the parameter estimates. In the next section, another MNL model including covariates will therefore be estimated.

Notice that the estimation outcomes described above will be used in Chapter 10 to compare the different designs. In the design generation, the covariates were not taken into account, although it is suggested they should be (Bliemer & Rose 2008). However, it is far from common practice to do this.

9.2 Advanced estimation models

As announced in the previous section, a more developed MNL model including covariates has been specified and estimated. Many different plausible interaction combinations could be specified between the socio-demographical attributes and the parameters characterising a trip (e.g interaction between age and price, gender and departure time, etc). When estimating some of those covariate combinations, some interactions appeared to be non-significant. This was the case for all interaction including gender, the interaction between income and price, or age and departure time. A positive correlation was expected between price and income. An explanation might be that Dutch consumers are typically stingy and highly price sensitive despite of the income.

The basic utility function formulated in equation (3.5) has been enlarged with the following covariates: airoine and frequent flyer program (AFF), age and price (APr), and age and travel time (ATr). The new extended utility function results in:

$$\begin{aligned} U_{iq} &= C_{iq} + \beta_{AIR} \cdot AIR_{iq} + \beta_{Pr} \cdot Pr_{iq} + \beta_{DPT} \cdot DPT_{iq} + \beta_{TR} \cdot Tr_{iq} + \beta_{egp} \cdot EGP_{iq} + \beta_{egt} \cdot EGT_{iq} \\ &+ \beta_{AFF} \cdot AIR_{iq} * FF_{iq} + \beta_{APr} \cdot AGE_{iq} * Pr_{iq} + \beta_{ATr} \cdot AGE_{iq} * Tr_{iq} + \varepsilon_{iq} \end{aligned} \tag{9.1}$$

where C_{iq} represents an alternative specific constant and ε_{iq} is the usual type I extreme value term, distributed identically and independently over alternatives and observations (see Chapter 3). Note that attributes AIR and DPT are expressed using dummy coding.

Two choice model types have been applied to estimate the defined utility function presented in equation (9.1). The first model is the well-known MNL model already used in the previous section. The second model is the more advanced panel mixed logit model and should result in better estimates (see Chapter 3).

With the help of NLOGIT, the MNL and panel ML models are estimated according to the theory of log-likelihood maximization (see Appendix C). As discussed in Chapter 4, the parameter estimates were estimated as generic across the SC alternatives. When estimating the panel ML model, the time parameters (transfer time, egress time and departure time) were allowed to be random assuming normal distributions. Other remaining parameters were estimated as non-random. Table 9-2 present the results of the MNL model and of the panel ML model.

Table 9-2 Outcomes MNL model and panel ML model with covariates

Variable	Unit	β coefficient	t-value	β coefficient	t-value
		MNL		**Panel ML**	
				Random parameters in utility function	
Transfer time (Tr)	Minutes	-.0154	-16.305	-.0233	-10.281
Egress time (EGT)	Minutes	-.0164	-12.899	-.0193	-10.176
Departure time 6:00 (Dep(0))	-	.2789	5.517	.2722	2.810
Departure time 12:00 (Dep(1))	-	.5841	11.302	.6478	7.929
				Non random parameters	
Price (Pr)	Euro	-.0312	-12.715	-.0378	-11.454
Egress price (EGP)	Euro	-.0307	-4.624	-.0530	-5.977
Air France (AIR1)*	-	.1155	1.413	.0985	0.952
KLM (AIR2)	-	.5823	7.490	.7633	7.597
Iberia (AIR3)*	-	-.0201	-.254	-.1634	-1.625

Vueling (AIR4)	-	-.2990	-3.798	-.4854	-4.903
Transavia (AIR5)	-	.2886	3.925	.3320	3.567
Air France/Flying Blue (A1FF1)*	-	.1206	.810	.2115	1.084
KLM/Flying Blue(A2FF1) *	-	.2424	1.805	.3047	1.754
Air France/Skyteam (A1FF3)*	-	-.1945	-.505	-.1698	-.359
KLM/Skyteam (A2FF3)	-	-.9200	-2.324	-1.0653	-2.242
Iberia/Iberia plus (A3FF2)	-	.8068	1.362*	1.3842	1.991
Iberia/Oneworls (A3FF4)*	-	-.4558	-.668	-.0664	-.081
Age(18-25)/Price (AG1Pr)	Euro	-.0216	-3.877	-.0224	-3.105
Age(26-35)/Price (AG2Pr)*	Euro	-.0229	-6.136	-.0296	-5.711
Age(36-45)/Price (AG3Pr)*	Euro	-.0125	-3.638	-.0185	-3.805
Age(45-60)/Price (AG4Pr)*	Euro	-.0053	-1.757	-.0052	-1.274
Age(18-25)/Transfer time (AG1Tr)*	Minutes	-.0006	-.276	.0004	.077
Age(26-35)/Transfer time (AG2Tr)*	Minutes	-.0017	-1.245	-.0037	-1.143
Age(36-45)/Transfer time (AG3Tr)*	Minutes	-.0048	-3.403	-.0081	-2.512
Age(45-60)/Transfer time (AG4Tr)*	Minutes	-.0017	-1.397	-.0023	-.806
Constant 1 (C1)	-	.4689	7.137	.6084	7.045
Constant 2 (C2)	-	.3353	4.996	.4388	5.068
Constant 3 (C3)	-	.3552	5.226	.4382	5.058
Constant 4 (C4)	-	.2032	2.882	.2744	3.069
				Derived standard deviation of parameter distributions	
Transfer time (NsTr)	Minutes	-	-	.0183	15.447
Egress time (NsEGT)	Minutes	-	-	.0192	11.248
Departure time 6:00 (NsDep(0))	-	-	-	1.6145	14.449
Departure time 12:00 (NsDep(1))	-	-	-	1.0313	9.746
Log-Likelihood (LL)	-	-3848		-3509	
Adjusted ρ^2	-	.259		.339	

* Significance level of 5%

The analysis reveals that most of the coefficients in the MNL model and in the panel ML model are statistically significant and of the expected sign. The standard deviation parameters for all attributes of the panel ML model are also statistically significant suggesting that significant taste heterogeneity exists among the respondents for these attributes. The only unexpected sign is the interaction between KLM and the FFP Skyteam. This value will be explained later on in this section. In terms of model fit, the panel ML model obtains a very significant improvement in log-likelihood (LL) of 339 units over the MNL model. However, the panel ML model has more parameters. The adjusted ρ^2 (which corrects for the used number of parameters) of the panel ML model is significantly higher than the one for the MNL model (.339 compared to .259) which reveals the better model fit of the panel ML over the MNL model.

Comparing parameters estimates of different models with each other is not allowed. Only by comparing the willingness-to-pay (WTP), parameters can be compared with each other. The WTP is the ratio between a parameter and another (cost) parameter, where the parameters can be fixed or random. In order to simplify its calculation, Train (2003) suggests to use fixed costs parameters instead of random.

Comparing the Willingness to Pay (WTP) of both models, the WTP for transfer time (tr) and egress price (EGP) slightly appear to change between the model estimations. In order to interpret the outcomes in terms of consumer preferences, only the model that fits the best will be discussed, i.e. the panel ML model

At first, an analysis has been performed in order to gain more insight into the contribution of each attribute to utility. As an example, two itinerary choice options have been assumed. As illustrated in Table 9-3, both itineraries arrive at the same airport and are assumed to be operated by the same carrier (Easyjet). Both utility functions have been determined and the itinerary with the highest utility would be chosen by the traveller (property of utility maximisation). The relative utility of both itineraries are illustrated in Figure 9-1.

Table 9-3 Value for chosen and non-chosen itinerary (focus on transfer)

	Chosen itinerary			Non-chosen itinerary		
	Description	Utility	Contribution	Description	Utility	Contribution
Flight Price	€100	-3.78	71%	€75	-2.835	58%
Transfer time	No transfer	-	-	60 min	-1.398	28%
Departure time	12:00	.647	12%	6:00	.023	1%
Egress price	€3	-.159	3%	€5	-0.265	5%
Egress time	40 min	-.772	14%	20 min	-0.386	8%
Total		-4.063	100%		-4.856	100%

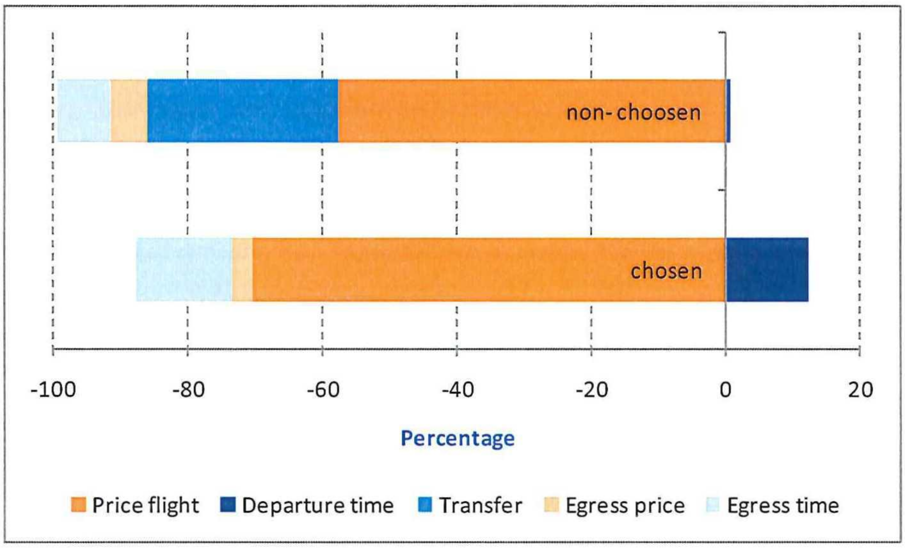


Figure 9-1 utility of chosen and non-chosen itinerary (focus on transfer)

In the chosen itinerary of the above illustrated case, fare contributes most to utility with 70%, followed by egress time (14%) and departure time (12%). In the non-chosen option, the fare contribution drops to 58%, due to the addition of a transfer in the itinerary. This shows that the contribution of transfer is almost 30% of the disutility. Additionally it shows that travellers have a preference for departing at noon instead of early morning (which is preferred towards departing in the evening). The departure time is in this case the only positive contribution to utility. In other cases where different airlines are involved, the airline may contribute positively to utility. We may conclude that price and travel time and departure time are the three most important attributes influencing travellers' itinerary choice.

In the next case, both itineraries would consist of a direct flight but will have a separated airport location. This will permit to get more insight into airport location's importance on travellers' itinerary choice. The proposed itineraries are as presented in Table 9-4. Again, the relative utility of both itineraries are illustrated in Figure 9-2.

Table 9-4 Value for chosen and non-chosen itinerary (focus on airport location)

	Chosen itinerary (to Barcelona)			**Non-chosen itinerary (to Girona)**		
	Description	Utility	Contribution	Description	Utility	Contribution
Flight Price	€100	-3.78	71%	€75	-2.835	52%
Transfer time	No transfer	-	-	No transfer	-	-
Departure time	12:00	.647	12%	6:00	.023	1%
Egress price	€3	-.159	3%	€12	-0.265	12%
Egress time	40 min	-.772	14%	1h40 min	-0.386	36%

Total	-4.063	100%	-5.373	100%

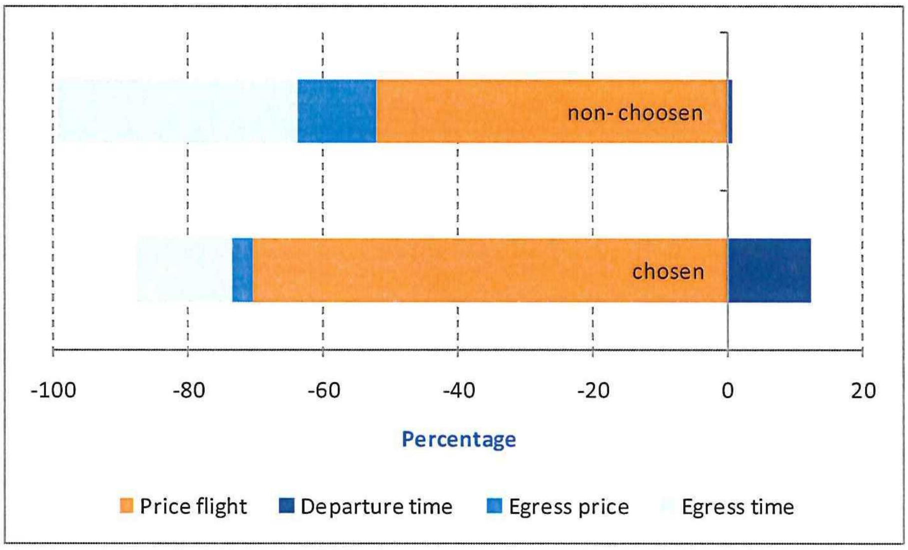


Figure 9-2 Relative chosen and non-chosen itinerary (focus on airport location)

In this case, it can be seen that arriving at a secondary airport (which is further located from the city) drops the fare contribution from 71% to 52%. This drop is due to the additional travel fare and the additional travel time needed to reach the city departing from the further located airport. Also it can be seen that contribution of egress time increase heavily (more than twice as much) when the airport is located further away from the final destination. The contribution of egress price is four times as much as in the previous case.

Both analyses have shown that travellers are price and time sensitive. Both attributes are the most contributing to utility followed by secondary attributes as departure time. The objective of these analyses is to determine to what extent the attributes are influencing travellers' choices. In the next section, the WTP of different parameters will be determined. This will result in more perceptible interpretations of the model outcomes.

As the panel ML model estimates the parameters of the distribution of the parameters, the WTP´s are probability distribution functions and not fixed values. Figure 9-3 illustrates the probability distribution function of the WTP for avoiding one hour of transfer time and one hour of egress time. It can be seen that the mean of the distribution function of the WTP for avoiding one hour of transfer time is estimated at €37 while WTP is estimated at approximately €22 for an hour of egress time. Hence, this is the amount of money travellers are willing to pay more in order to avoid one additional hour of travel time.

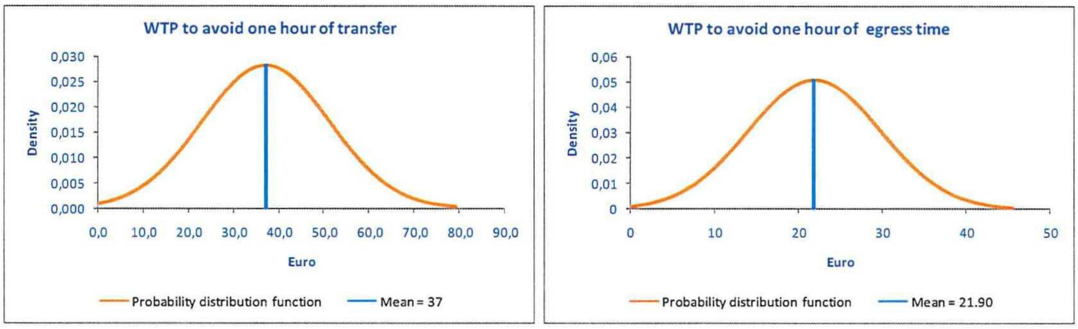


Figure 9-3 probability distribution function of WTP for one hour of transfer and egress time

In this case, Girona has an egress time of about 40 to 60 minutes more than Barcelona, therefore, travellers are willing to pay about €20 to travel to Barcelona instead of Girona.

Concerning the departure times, notice the extreme high value of the standard deviation parameters compared to the parameter estimates (the variation coefficient of β_0 is 600% and is 160% for β_1). This reveals that travellers' preference for departing at a specific moment of the day depends regarding to each individual. Additionally, the low t-value for Dep(0) showed that the model had difficulties in estimating the parameter value. Nevertheless the significant parameters permit to determine the WTP for departing at a certain moment of the day. Notice that the presented value is the mean of the distribution function of the WTP, which means it is averaged value. The WTP for departing in the morning (6:00) instead of in the evening (18:00) is approximated on €7,20 whereas the WTP for departing in the afternoon (12:00) instead of in the evening is estimated on €17,15.

The outcomes of the model show that Air France an Iberia have no more attractiveness compared to Easyjet (the estimates are non significant). It could be that some of the respondents had a bad experience with one of those airlines in the past and have no preference anymore in a specific airline. However, the results show respondents' preferences for the Dutch airlines. On one hand, the willingness to pay for travelling with KLM instead of Easyjet is around €20 and is estimated on approximately €13 for Transavia. On the other hand, Vueling has a negative impact on travellers' choice and the WTP to travel with Easyjet rather than with Vueling is estimated on €9.

Surprising are the interactions between airlines and FFP. Roughly, the results show that having a FFP does have none significant impact on travellers' itinerary choice. Apparent is the non significance of Iberia's parameter whereas the correlation between Iberia and Iberia's FFP is positive and significant. This means that Iberia only becomes more attractive than any other LCC if travellers are in possession of an Iberia FFP. Additionally, the results demonstrate that travellers prefer travelling with Easyjet rather than with KLM while having a skyteam FFP (which is the same as having a KLM FFP). Obviously, this outcome is strange and seems not completely reliable. The really low t-value of this parameter estimate demonstrates that the significance of this parameter could be a coincidence, and thus, no interpretation on this parameter will be given. Again, the constant estimates show that respondents were likely to choose the first alternatives presented. The sorting options could clarify this phenomenon, as well as the fact that European people read from left to right and from top to bottom. In this survey the alternatives were presented from top to bottom.

Finally, the interaction between age with price and age with transfer time result in the following outcomes: young travellers are more sensitive for price than older travellers. Travellers of the first age group (18-25) are 59% more price sensitive than travellers of the last age group (60+), travellers of age of the second age group are 78% more sensitive, and travellers of the third age group are 48% more sensitive. Even if the interaction between income and price was non-significant, the differences of price sensitiveness according to the age groups might be explained by peoples' income as it generally increases with age. Generally, younger travellers are more likely to spend more time while travelling than spending money. For this reason they mostly choose for low budget tickets.

9.3 Conclusions

In this chapter, different outcomes of estimation models have been discussed. It has been seen that panel ML models have a better model fit compared to MNL models ($\rho^2_{PanelML} > \rho^2_{MNL}$) which leads in better estimations of the parameter estimates. With the outcomes of the panel ML model, statements can be done regarding to travellers' behaviour in travel choices, and thus, the first two research questions of Chapter 1 can be answered. As seen in the previous section, the outcomes of the analyses show that price and travel time are the most important attributes influencing travellers' travel choice. Also, young travellers are more sensitive for price than older people and therefore are willing to spend more time travelling in exchange for a lower travel price. Generally, travellers prefer to travel in the morning or at midday instead of later on the day. Although this really depends according to individual preferences, this might be explained by the context scenario which involved undertaking a holiday trip. This would mean that travellers prefer to leave earlier in order to enjoy their trip at maximum and gain an extra day.

As observed in the model's outcomes, travellers are sensitive for egress times. Because secondary airport are located further away of the city than main airports, travellers have a preference for arriving at the main airport due to the less egress time which is related to the airport location. As well, the results show that travellers are willing to pay approximately €22 to avoid one hour of egress time. Since the airline type (NC or LCC) does not really affect travellers' choice (Iberia or Air France are not significantly preferred to EasyJet and Transavia is as LCC significantly preferred to other NC's), the attractiveness of secondary airport for travellers will mostly depend on the offered travel prices.

Regarding the airlines, operating at secondary airports is not necessarily a disadvantage. Because airport taxes are less than at main airports, it permits to offer cheaper travel prices to travellers. As seen in Chapter 2, reducing ticket prices is one of the airlines' instruments to attract more travellers and travel prices are the most important attribute for travellers' travel choice. Additionally, the high frequency service of secondary airports might even be an advantage for airlines to increase their traffic volume.

Now that the empirical research questions have been answered, the next chapter will discuss the methodological perspectives of this research. After having designed different experimental designs (see Chapter 5 and 6), those will be compared in order to determine the estimates' reliability of each design.

Chapter 10 Designs Comparison

10.1 Introduction

As seen in chapter 5 and 6, three different experimental designs have been generated: an orthogonal design, an efficient design with 18 choice situations (efficient 18), and an efficient design with 108 choice situations (efficient 108). Bliemer and Rose (2007) mention two benefits of efficient designs compared to orthogonal design: (i) the improvement of the parameter estimates' reliability and (ii) the reduction of required sample size to produce a fixed level of reliability in the parameter estimates. Since these statements have never been confirmed in practical cases, the aim of this chapter is to compare all three experimental designs and verify the validity of the theory.

The first section of this chapter discusses the designs' comparison methodology, whereas the second section presents the comparison outcomes and a discussion on the results' interpretations.

10.2 Comparison methodology

Comparing the designs with each other can be done by estimating the collected data derived from each different design. Since each experimental design has been optimised for an MNL model, this type of estimation model suits the best to determine the designs' performances.

When estimating models with different data, the model fit (ρ^2) usually permits to determine which of the models reveal the best reliable parameter estimates. However, since the data differ for each design, this criterion is not valid for determining the model performances. So, in order to determine which of the three designs result in the best parameter estimates, three comparing criteria will be analysed between different designs: (i) the parameter estimates over different sample sizes, (ii) the standard error over different sample sizes and (iii) the t-ratios over different sample sizes.

The parameter estimates of different designs are compared to the parameter estimates of the simple MNL model (see Chapter 9) since this estimation has been done over the full data. However, because the estimates are asymptotically converging, it might be that the most reliable parameter estimate has not completely converged yet, which could mislead the analyst in his conclusions. Comparing the standard errors of the parameter estimates is a more accurate comparison criterion, since that the most reliable parameter estimate is the one with the lowest standard error (see Chapter 6).

The second benefice of efficient design relative to orthogonal design should be the reduction in sample size needed without losing any level of reliability in the parameter estimates. In order to analyse this statement, the t-ratios over different sample sizes are compared. Since $t = \beta / s.e.$, and in order to compare the t's, the standard error (s.e.) is computed per design, and β is kept fixed at the reference i.e. the reference is the β of the simple MNL model (estimation with the most data).

In the next section, the outcomes of the models estimations are presented and discussed.

10.3 Results and interpretations

The results of the estimation models are presented in three different graph types representing respectively the estimated parameter value, the standard error of the parameter, and the t-value (significance) of the parameter, each in function of the number of respondents having observed the survey. The outcomes are presented in Figure 10-1 and Figure 10-2 below.

In order to determine which design yields the most reliable parameter estimates, the parameters' standard deviation of each estimation model have to be analysed. In Figure 10-1 and Figure 10-2, the graphs in the second row represent the parameters' standard deviation in function of *N* (number of respondents having answered the survey). When analysing the standard error outcomes for each design, the two efficient designs seem, for five of the six parameters, to perform better than the orthogonal design. This can be explained by the fact that, both efficient designs have been optimised on the base of prior estimated parameters (see chapter 6). However, for some parameters (i.e. price and Dep(1)), the orthogonal design seems to perform as good or even better than both efficient designs. This means that those parameters were easy to estimate compared to other parameters. This can be seen when looking to the t-values of the parameters. Statistically, a parameter estimate becomes significant if its t-value exceeds 1.96. However, the results show that, the price and transfer parameter estimates are already significant after 18 observations. Figure 10-3, represents the t-value outcomes of the parameters needing less than 18 observations to become significant. Even then, the graphs show that the price and transfer parameter estimates are directly significant for the orthogonal and the efficient 18 design. These outcomes support the fact that those parameters are easy to estimate compared to other parameters as EGP or Dep(0).

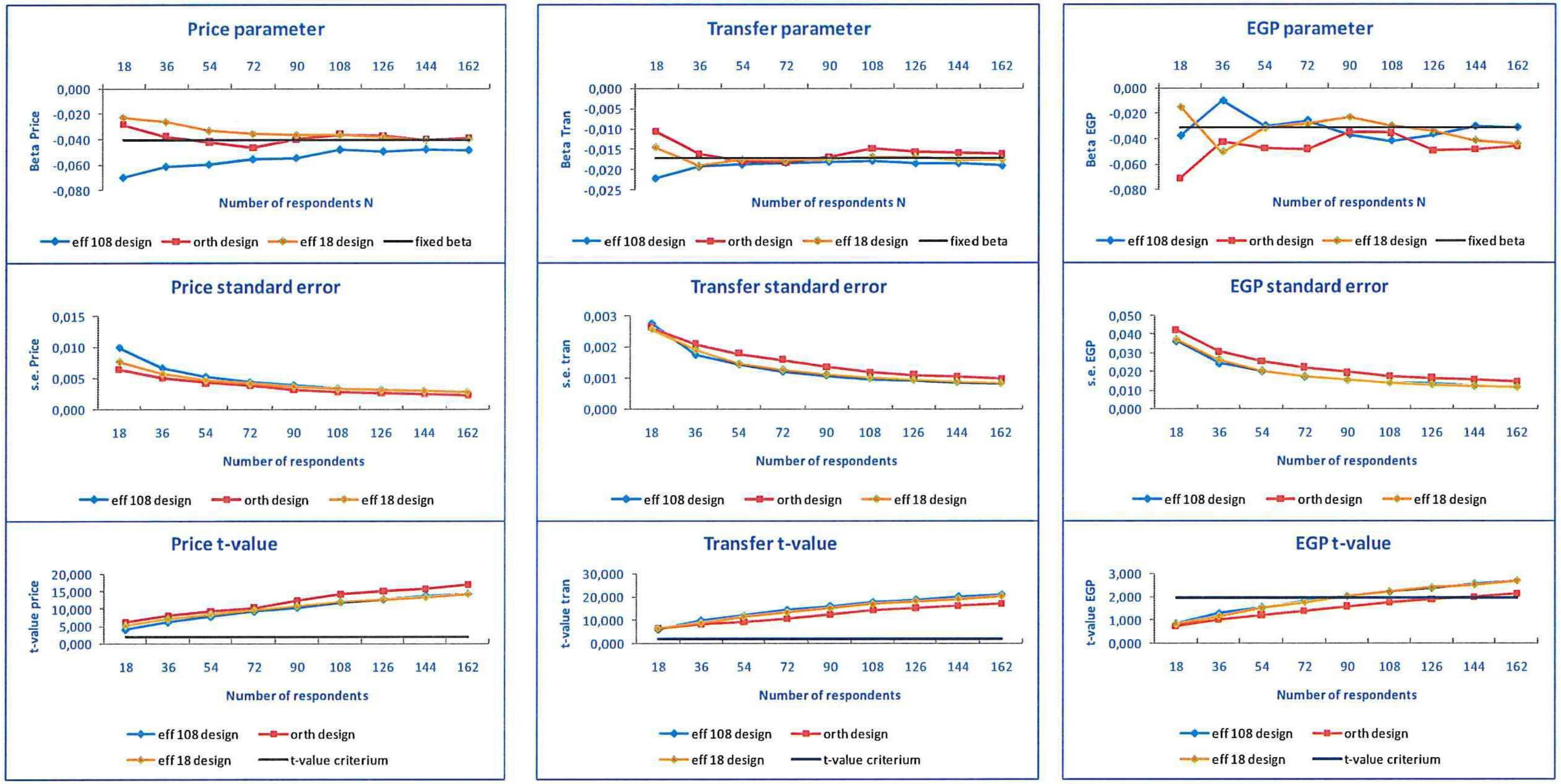


Figure 10-1 Estimation results for MNL model for each experimental design (part 1)

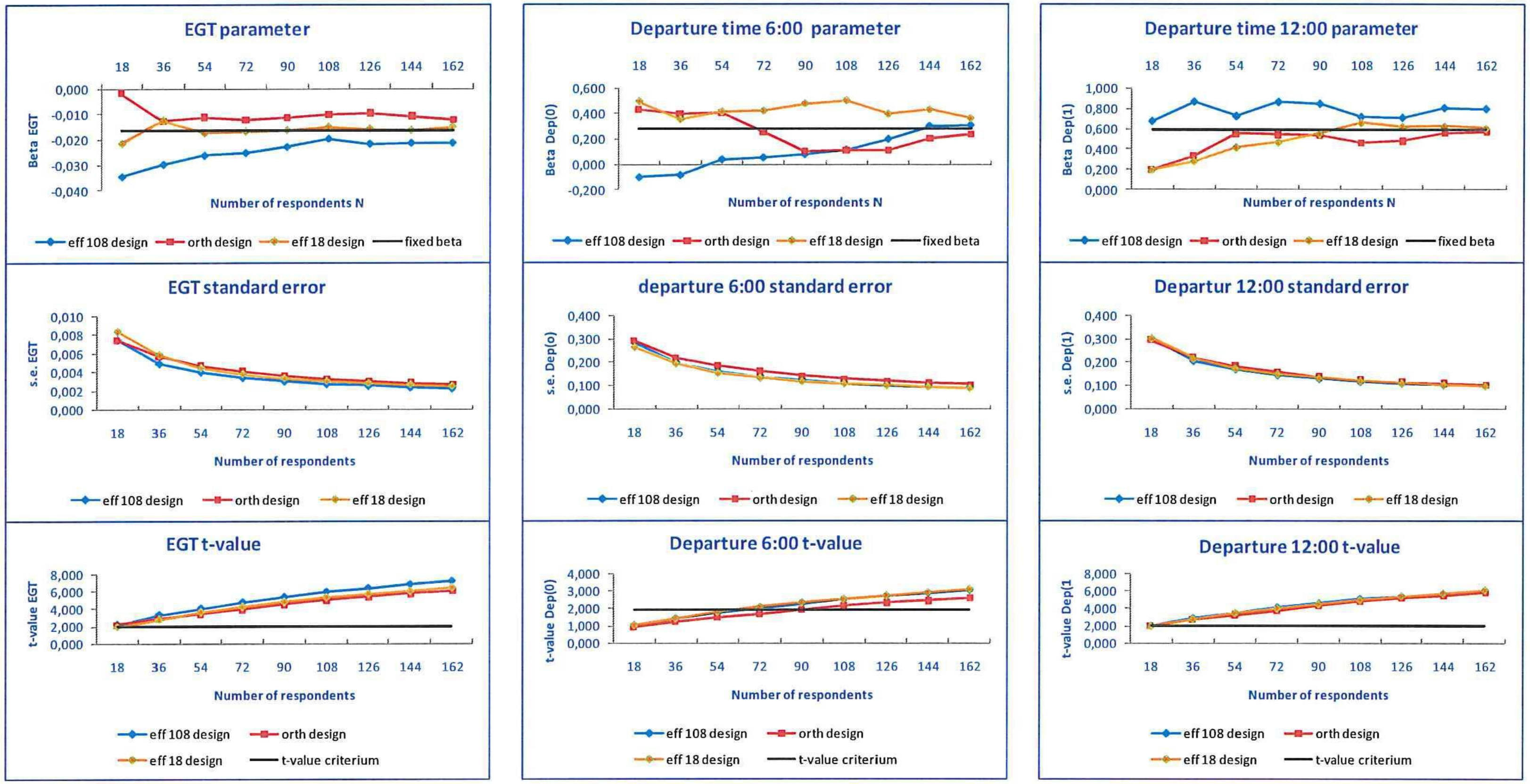


Figure 10-2 Estimation results for MNL model for each experimental design (part 2)

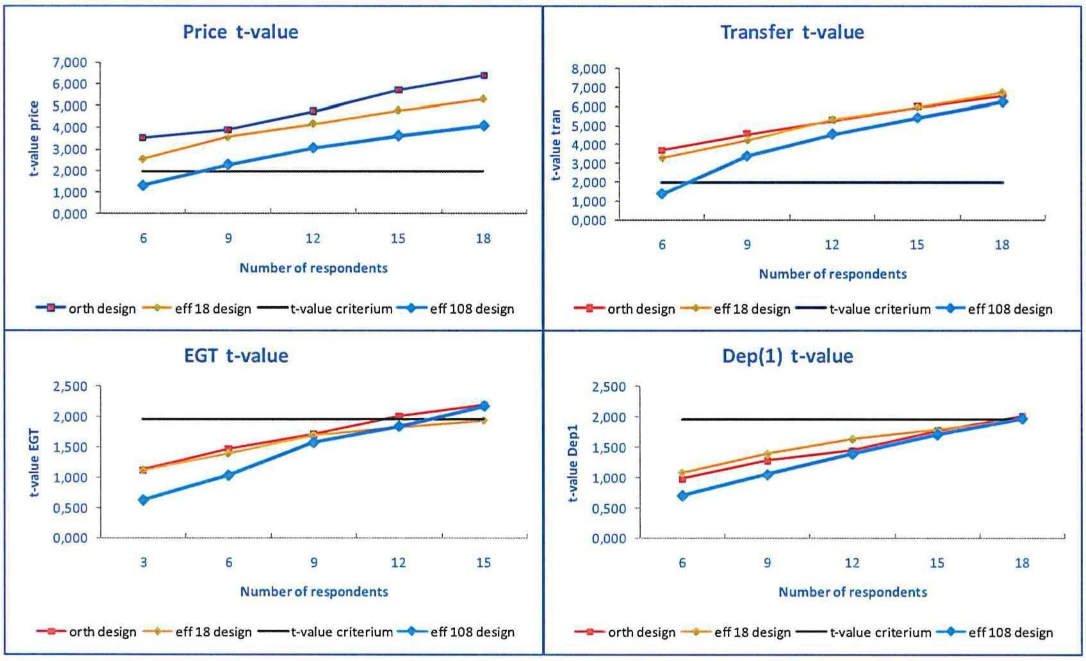


Figure 10-3 Significance of easy estimating parameters in function of N

Furthermore, the final t-value reveals the significance value of each parameter estimates. Figure 10-1 and Figure 10-2 show that parameter estimates of the efficient designs have a higher significance level than parameter estimates of the orthogonal design. Hence, it seems obvious that efficient designs yield better parameter estimates compared to orthogonal designs, where the efficient designs seem to optimise particularly for parameters that are more difficult to estimate. However, interpreting both efficient designs' comparison seems to be more difficult. Beside that both perform better than the orthogonal design, they both result in approximately the same estimates reliability. Although none of both are especially better regarding their performances, it is easier to construct an efficient design with 18 choice situations rather than 108. Since there is not much difference between sing 18 and 108 choice situations in the design, it suffices to generate a design with 18 choice situations for a comparable efficiency. In other words, designs can be kept limited in size.

In terms of required sample size, the models' outcomes show that efficient designs require, as revealed by the theory, a lower sample size to produce a fixed level of reliability in the parameter estimates. This can be distinguished when observing the t-value of EGP and Dep(0), which are the most difficult parameters to estimate. In order to yield significant parameter estimates, both efficient designs require approximately 90 respondents to estimated EGP versus 144 for the orthogonal design, and 70 respondents to estimate Dep(0) against 90 for the orthogonal design. Notice that if the efficient designs would have been optimised according to the S-estimates instead of the D-error, the EGP parameter would have been the optimised criterion parameter, and even lower sample sizes can be expected.

10.4 Conclusions

By conducting such experiments, the aim of this chapter was to verify and validate the theory concerning the benefits and improvements of efficient design against orthogonal design. Since the literature on designing choice experiments and on design efficiency has been introduced relatively recently (see (Bliemer & Rose 2005, Rose & Bliemer 2006, Bliemer *et al.* 2007)), this research is the first that applies the theory in order to compare different design types on a practical case.

The outcomes of these analyses have shown that, as expected, efficient designs result in more reliable parameter estimates than orthogonal designs. In terms of sample size requirement, the results show that for parameters which are difficult to estimate efficient design require a lower sample size than orthogonal designs. The comparison of both efficient designs (with 18 choice situations or with 108 choice situations) has proved that designs with a limited number of choice situations can be as efficient as lager designs (which take more computation time to generate).

Chapter 11 Conclusions

11.1 Introduction

Since main airports are running out of capacity, a trade-off exists between building additional airport capacity at the existing airport and developing a secondary regional airport located elsewhere. In order to get some information on the attractiveness of these secondary airports, it is necessary to have knowledge on air travellers' preferences in selecting flight itineraries. Hence the purposes of the first objectives were to determine the factors influencing travellers' choice, and more especially to learn more about the effects of airport location on their itinerary choices. The second objectives had a total different aspect and concerned the methodology for constructing experimental design in SC experiments, which is a frequently used method to analyse travellers' behaviour. In this chapter, the conclusions of this study will be presented following by some recommendations.

11.2 Travellers' behaviour and decision attributes

When undertaking a trip, the traveller makes a series of decision. Some of these decisions are influenced by the traveller's preferences and some depend on the offered transport characteristics. Each traveller's choice is made from a specific choice set of alternatives, in which each observed alternatives is assumed as known and available to the traveller. Finally, the traveller chooses the alternative that maximizes his utility. It is necessary that the utility to undertake the trip has to be higher than the utility of not taking part of the trip.

In this study, the attributes influencing travellers' itinerary choice for leisure activity purpose have been identified. In the situation where two different airports are available, the choice of undertaking a trip depends on price, travel and transfer time, departure time, and airport location. The airport location has its importance since the egress travel time and the egress prices (prices to access the final destination from the airport) vary according to the airport location. The outcomes of the study show that the most important attributes influencing travellers' in their itinerary choice are respectively price, transfer time, egress time, and departure time.

In an average chosen itinerary, fare yields a relative large contribution to utility with 70%. In itineraries containing a transfer, the transfer yields in a disutility contribution of almost 30% and fare contribution approximately drops to 60%. Travellers prefer departing in the afternoon or rather than in the evening, as it contributes of 12% of utility. Departing in the morning is slightly preferred but has a negligible utility contribution (1%).

In the case itineraries have different airport locations, arriving at a secondary airport drops the fare contribution to utility to approximately 50%. This drop is the result of the additional egress time and egress price needed to access the final destination. Egress time and egress price contribute respectively to 36% and 12% of disutility. This study has shown that travellers prefer arriving at main airports rather than at secondary airports and avoid additional travel time to access their final destination.

In terms of willingness to pay, travellers are disposed to pay about €37 to avoid one hour of transfer time and €22 to avoid one hour of egress time. The willingness to pay for departing in the afternoon

(12:00) instead of in the evening is around €17, whereas it is estimated on €7,20 for departing in the morning instead of in the evening. Travellers have no significant preference for travelling with NCs or with LCCs. However, Dutch travellers prefer travelling with a Dutch airline company (KLM or Transavia). Travellers are prepared to pay €20 on their travel ticket in order to travel with KLM and approximately €13 to travel with Transavia.

11.3 Parameter estimates reliability towards experimental designs

Constructing an optimised experimental design is the base to perform a good SC experiment. Since the AVC matrix can be determined without conducting any survey, it is possible to determine in advance the efficiency of the design. In this study, three different experimental designs have been generated, namely an orthogonal design (type of design used in the literature) with 108 choice situations, an efficient design with 18 choice situations, and second efficient design with 108 choice situations.

The results of the design comparison have shown that, efficient designs are outperforming orthogonal designs as soon as parameters are difficult to be estimated. When estimating the collected data, the obtained parameter estimates seem to be more significant when the data relied on an efficient design rather than on an orthogonal design.

In terms of sample size requirement, data estimation relying on efficient designs requires a smaller sample size than data relying on orthogonal designs while reaching the same significance level. Because of the high value price and the difficulties of collecting data, this outcome is a great improvement in the methodology of how to perform SC experiments. No significant difference has been found between both efficient designs. It is suggested to generate an efficient experimental design with the smallest number of choice sets as possible, as an increase in choice sets will result in a higher required sample size.

11.4 Recommendations

While the government has to decide where to build new airport capacity, this study has permit to gain more insight into air travellers' preference, which can be used by airports and airlines to better adapt their market strategy in the case of the implementation of a secondary airport. In this section, some recommendations are made towards the main stakeholders affected by the problem. This section will end with some recommendations for further scientifically research.

Recommendation for airlines

The results have shown that airline type (NC or LCC) and image is not valued as very important by passenger (no significant preference between non Dutch airlines). Hence, all airlines can be attractive for travellers. In order to reduce their costs, airlines could take some new decisions. All measures as flight on off-peak hours, operating at secondary airports, abolishing labour costs, etc, will lead in cost difference. Obviously, secondary airports are not attractive for travellers because of their further location to main cities. However, this study has shown that price, transfer time and departure time are the other important factor influencing travellers' itinerary choice. Hence, offering tickets based on travellers' preferences (low prices, no transfer, preferred departure time...) would balance the unattractive airport location. This is why LCC, operating mostly at regional airports, are so successful the last 10 years.

Recommendations on the issue of developing a secondary airport

Given the capacity problems of the main airports, a solution would be to develop a secondary airport. Airports have two kinds of customers, namely the airlines and the travellers and it seems difficult to satisfy both while increasing their own profit. As every service, a secondary airport is only profitable if it is used by consumers. We have seen in this report that all stakeholders involved with this issue do not have the same perception on this topic. Despite that main airports are the main opponent regarding the development of secondary airports, the government has the final having the last word and has to take a decision. This research has shown that price and travel time are the most important factors influencing travellers in their itinerary choice. Hence, secondary airports are not the most preferred itinerary as their location is further way from the main cities. However, since the price plays an important role as well, it is easily possible for airlines to decrease their prices and still make the same profit as the airport taxes are lower at regional airports than at the main airports. This means that if the price advantages overrule the disadvantages of the airport location, regional airport still be attractive for travellers. In this case, it would be an option to separate airport strategies according to distinct consumer segmentation. For example, secondary airport would be used for domestic flights (inter Europe), and the main airport would be developed for network/hub operations and for international flights. Both airports would have individual strategies and a different market segmentation as a result of which regional airports will not be main airports' competitors.

However, this research focussed on the effect of the arriving airport location on air passenger´s choice. This means that the results apply for passengers travelling to the Netherlands, and no additional information is known on the local passengers´ preferences. If the prices are attractive, foreign travellers will use the airport but is this still the case for local travellers who have knowledge on the access facilities? Hence, before taking the decision to develop the regional airport of Lelystad, it would be necessary to get more insight in Dutch travellers´ behaviour and more precisely in the importance accorded to access time. Additionally, the research demarcation focussed on travellers with leisure purpose. Passengers travelling for business would certainly have other preference. Hence, this research partially contributes in the decision of developing a secondary airport and further research (on business traveller, on local travellers, etc) is required in order to get all travellers´ perception on regional secondary airports.

Recommendation for further scientifically research

While generating an efficient design, prior parameters are used to determine the efficiency of the design. However, even if coming out of a pilot survey, prior parameters always have some uncertainty. Using Bayesian prior parameters instead of fixed parameters permit to take into account a part of that uncertainty and yields more robust designs. However, estimating data with estimation models yields new parameter estimates. This new estimates should then be re-used as prior parameters for a secondary experiment where a new optimal efficient design would be generated. Hence some research should be done such that, each time a respondent observers the survey, the so far collected data is estimated which will results in new parameter estimates. Then the last outcomes should be used as prior parameters and a new design should be generated and presented to a new respondent. That way, after each observation, a new optimal efficient design is generated and yields each time on more reliable parameter estimate. However, including such kind of tools in the software should not be easy as the Ngene (software required to generate designs) is not linked to the survey outcomes.

While the experimental designs have been optimised for MNL models, some other estimation models have been estimated (ML). Despite that efficient designs constructed for MNL model may relatively be efficient for ML model estimations, it should be interesting, to compare two designs, one optimised for MNL model and the other one optimised for ML model. Due to the long running time for generating a ML design, no efficient ML design could be found that was more efficient for ML estimation than the design optimised for MNL model. Further investigation on this topic should be interesting.

Appendix A – Adopted Orthogonal Design

Ngene syntax

```
Design
;alts = alt1, alt2, alt3, alt4, alt5
;rows = 108
;block = 6
;orth = sim
;model:

U(alt1) = air*air[0,1,2,3,4,5] + pr*pr[0,1,2] + dep*dep[0,1,2]+ tr*tr[0,1,2] +
egp*egp[0,1,2] + egt*egt[0,1,2] + dest*dest[0,1]/

U(alt2) = air*air[0,1,2,3,4,5] + pr*pr[0,1,2] + dep*dep[0,1,2]+ tr*tr[0,1,2] +
egp*egp[0,1,2] + egt*egt[0,1,2] + dest*dest[0,1]/

U(alt3) = air*air[0,1,2,3,4,5] + pr*pr[0,1,2] + dep*dep[0,1,2]+ tr*tr[0,1,2] +
egp*egp[0,1,2] + egt*egt[0,1,2] + dest*dest[0,1]/

U(alt4) = air*air[0,1,2,3,4,5] + pr*pr[0,1,2] + dep*dep[0,1,2]+ tr*tr[0,1,2] +
egp*egp[0,1,2] + egt*egt[0,1,2] + dest*dest[0,1]/

U(alt5) = air*air[0,1,2,3,4,5] + pr*pr[0,1,2] + dep*dep[0,1,2]+ tr*tr[0,1,2] +
egp*egp[0,1,2] + egt*egt[0,1,2] + dest*dest[0,1]$
```

Choice situation	alt1.air	alt1.pr	alt1.dt	alt1.tran	alt1.egp	alt1.egt	alt1.dest	alt2.air	alt2.pr	alt2.dt	alt2.tran	alt1.egp	alt1.egt	alt2.dest	alt3.air	alt3.pr	alt3.dt	alt3.tran	alt1.egp	alt1.egt	alt3.dest
1	4	100	2	0	15	60	1	4	100	0	120	5	20	0	4	100	0	0	3	20	0
2	5	75	2	0	3	40	0	2	75	2	60	12	80	1	0	100	2	60	5	20	0
3	0	50	0	60	1	30	0	3	100	0	0	1	30	0	2	75	2	0	15	100	1
4	0	75	1	120	3	40	0	2	50	1	60	9	60	1	5	50	0	120	3	30	0
5	5	100	0	60	15	60	1	1	50	1	120	3	40	0	1	75	1	120	9	100	1
6	1	50	1	120	9	80	1	3	75	2	0	15	100	1	3	50	1	60	9	80	1
7	2	75	2	60	1	20	0	0	100	1	120	9	80	1	5	100	2	120	9	80	1
8	4	100	2	60	12	80	1	5	75	0	0	12	100	1	1	100	1	0	3	30	0
9	3	100	0	120	5	30	0	2	100	1	0	5	20	0	3	75	0	120	15	60	1
10	3	50	1	0	9	100	1	4	75	0	60	5	30	0	0	50	2	0	1	40	0
11	1	75	1	0	12	60	1	0	50	2	120	3	20	0	2	50	0	60	5	40	0
12	2	50	0	120	5	40	0	4	50	2	60	9	100	1	4	75	1	60	12	60	1
13	4	75	0	60	1	40	0	3	75	2	0	9	100	1	0	50	2	120	5	20	0
14	0	100	1	120	3	20	0	5	75	2	60	3	20	0	1	100	1	120	9	100	1
15	5	75	1	120	15	100	1	0	100	0	0	5	30	0	5	100	0	60	12	100	1
16	2	50	0	60	5	30	0	4	50	1	120	15	60	1	3	50	1	0	1	20	0
17	3	50	2	0	9	80	1	1	100	0	120	12	100	1	2	75	2	60	15	80	1
18	1	100	2	0	12	60	1	2	50	1	60	1	30	0	4	75	0	0	3	30	0
19	0	100	1	0	5	30	0	3	100	1	60	3	30	0	2	75	1	60	5	20	0
20	5	50	2	60	12	100	1	2	100	1	120	15	100	1	3	50	0	60	9	100	1
21	1	75	1	0	9	60	1	1	75	0	0	9	100	1	1	75	0	120	1	20	0
22	4	100	2	60	3	30	0	3	50	2	60	1	20	0	5	50	2	0	12	100	1
23	5	75	0	120	1	20	0	1	50	2	0	5	30	0	4	100	1	0	5	30	0
24	0	50	0	120	15	100	1	5	75	0	120	12	60	1	0	100	2	120	12	80	1
25	5	100	0	0	3	40	0	4	50	0	0	12	60	1	4	50	1	60	15	80	1
26	1	50	0	0	15	60	1	0	100	2	60	5	30	0	0	50	0	120	9	80	1
27	0	50	1	60	9	60	1	3	50	0	60	1	40	0	5	100	2	60	3	20	0
28	0	75	2	120	5	30	0	2	100	2	120	9	60	1	1	75	1	120	5	40	0
29	4	100	2	120	1	40	0	1	75	1	0	5	40	0	3	75	2	0	3	40	0
30	5	75	1	60	12	80	1	5	75	1	120	12	80	1	2	100	0	0	9	60	1
31	0	75	1	60	12	80	1	4	100	2	60	5	40	0	3	75	0	0	5	30	0
32	2	100	1	60	9	100	1	0	75	1	120	9	60	1	4	75	2	60	9	80	1
33	4	50	0	0	5	20	0	2	75	1	0	12	100	1	2	100	0	60	5	40	0
34	4	100	2	120	1	40	0	3	100	2	120	3	30	0	1	50	1	0	12	60	1
35	5	75	0	0	15	80	1	4	50	0	60	9	80	1	5	100	1	120	3	40	0
36	0	50	2	120	3	20	0	2	50	0	0	5	20	0	0	50	2	120	9	60	1
37	5	50	1	120	15	80	1	3	50	1	0	3	30	0	4	50	2	120	5	30	0
38	3	100	1	120	3	20	0	1	100	0	120	9	100	1	0	50	1	0	1	30	0
39	1	75	2	0	1	40	0	5	50	1	60	5	40	0	2	100	1	120	9	60	1
40	1	100	0	60	3	20	0	1	75	2	120	15	80	1	5	75	2	60	5	40	0
41	3	50	2	0	15	80	1	0	75	2	0	1	20	0	1	100	0	60	12	60	1
42	2	75	0	60	9	100	1	5	100	0	60	12	60	1	3	75	0	0	12	100	1
43	0	100	1	0	1	30	0	2	50	2	0	15	100	1	5	50	1	60	12	100	1
44	5	50	1	0	12	100	1	4	100	1	60	9	60	1	1	50	0	120	5	40	0
45	4	50	2	60	5	40	0	1	50	2	60	3	30	0	3	100	2	60	9	80	1
46	4	75	0	120	9	60	1	3	100	1	120	3	40	0	0	75	1	120	3	20	0
47	2	100	0	120	12	80	1	2	75	0	0	1	30	0	2	75	2	0	1	20	0
48	0	75	2	60	5	20	0	3	75	0	120	15	60	1	4	100	0	0	15	80	1
49	5	100	2	0	1	20	0	5	100	0	60	15	60	1	0	75	1	60	1	30	0
50	1	50	0	60	3	30	0	4	100	0	120	1	30	0	1	50	0	60	12	60	1

Choice situation	alt4.air	alt4.pr	alt4.dt	alt4.tran	alt1.egp	alt1.egt	alt4.dest	alt5.air	alt5.pr	alt5.dt	alt5.tran	alt1.egp	alt1.egt	alt5.dest	Block
1	1	100	1	60	9	60	1	1	100	2	0	1	30	0	1
2	5	75	2	120	3	40	0	4	50	0	120	5	40	0	1
3	2	50	1	60	3	20	0	5	75	0	60	3	20	0	1
4	1	50	0	0	9	80	1	0	75	2	60	15	60	1	1
5	3	75	0	0	15	80	1	2	50	1	120	9	100	1	1
6	3	100	2	120	5	40	0	3	100	1	0	12	80	1	1
7	4	100	2	60	15	100	1	0	50	2	60	3	30	0	2
8	2	50	0	120	1	30	0	3	100	0	120	5	20	0	2
9	0	75	2	60	12	100	1	4	75	1	0	5	40	0	2
10	4	75	0	120	15	80	1	2	75	2	0	9	100	1	2
11	0	50	1	0	3	20	0	2	100	1	120	12	60	1	2
12	5	100	1	0	1	20	0	4	50	0	60	9	80	1	2
13	1	100	0	60	1	20	0	0	75	1	120	1	40	0	3
14	5	50	0	60	12	60	1	4	50	2	0	3	30	0	3
15	0	75	2	0	5	30	0	4	100	0	60	1	20	0	3
16	2	75	1	120	12	100	1	0	100	2	60	15	60	1	3
17	4	50	2	0	9	80	1	2	50	1	0	15	80	1	3
18	3	100	1	120	5	40	0	5	75	0	120	12	100	1	3
19	3	100	0	120	1	20	0	0	100	2	60	1	40	0	4
20	1	50	0	120	3	20	0	1	75	0	120	3	30	0	4
21	4	75	1	0	12	100	1	3	50	0	0	5	20	0	4
22	5	75	2	60	5	30	0	1	50	1	0	9	60	1	4
23	0	50	2	60	9	80	1	5	75	2	120	15	80	1	4
24	2	100	1	0	15	100	1	5	100	1	60	12	100	1	4
25	4	100	0	60	15	100	1	1	75	0	0	3	20	0	5
26	5	50	1	120	1	30	0	1	50	1	60	5	40	0	5
27	0	75	0	60	12	100	1	5	100	2	120	5	30	0	5
28	1	75	1	120	5	30	0	0	100	0	120	9	80	1	5
29	3	50	2	0	3	20	0	3	50	2	60	12	100	1	5
30	2	100	2	0	9	60	1	5	75	1	0	9	60	1	5
31	5	100	0	0	15	100	1	1	50	2	120	3	20	0	6
32	0	50	1	60	1	30	0	1	100	0	0	5	40	0	6
33	5	75	1	60	15	80	1	3	75	2	60	1	30	0	6
34	1	75	0	0	3	40	0	2	75	1	60	15	60	1	6
35	4	50	2	120	3	20	0	3	100	1	0	12	100	1	6
36	0	100	2	120	9	60	1	5	50	0	120	9	60	1	6
37	2	75	0	0	12	80	1	0	50	0	120	3	40	0	7
38	3	50	1	60	5	20	0	3	75	1	60	1	20	0	7
39	0	100	0	0	5	30	0	4	100	1	0	5	30	0	7
40	2	100	2	120	12	100	1	2	100	0	0	9	80	1	7
41	4	50	2	120	9	100	1	1	75	2	60	12	60	1	7
42	4	75	1	60	1	20	0	5	50	2	120	15	100	1	7
43	5	75	1	0	9	60	1	2	75	0	0	5	40	0	8
44	0	100	2	60	3	40	0	5	50	1	60	1	30	0	8
45	1	50	1	0	15	60	1	3	100	2	120	1	20	0	8
46	5	50	2	60	9	100	1	1	100	0	120	12	60	1	8
47	1	100	0	120	5	30	0	1	50	2	60	15	80	1	8
48	3	75	0	120	3	30	0	3	75	1	0	12	100	1	8
49	2	75	2	0	3	30	0	2	100	2	60	3	20	0	9
50	3	100	2	0	15	80	1	3	75	0	120	5	40	0	9

Figure A-1 Orthogonal design

51	3	100	0	60	15	60	1	2	50	1	60	3	40	0	5	50	2	0	15	60	1	1	50	1	120	1	40	0	3	50	1	0	3	30	0	9
52	0	75	2	0	5	40	0	3	75	2	0	12	80	1	3	75	0	120	3	30	0	0	50	0	60	15	60	1	2	50	0	0	9	80	1	9
53	4	75	1	120	9	100	1	0	50	1	0	9	60	1	2	100	1	0	9	100	1	5	100	1	120	12	100	1	1	75	2	120	9	100	1	9
54	2	50	1	120	12	80	1	1	75	2	120	5	40	0	4	100	2	120	5	40	0	4	75	0	60	1	20	0	4	100	1	60	15	60	1	9
55	1	50	0	120	5	40	0	5	50	2	120	1	40	0	2	100	0	0	1	30	0	4	75	2	60	3	30	0	2	50	0	0	3	20	0	10
56	3	75	1	0	12	60	1	1	50	2	0	12	60	1	3	75	2	0	12	60	1	2	100	2	60	5	30	0	0	100	1	60	5	40	0	10
57	2	100	0	120	9	80	1	0	100	1	60	15	60	1	1	100	2	60	3	30	0	5	50	0	120	15	60	1	5	75	1	120	1	30	0	10
58	5	50	1	0	3	40	0	5	75	0	120	5	30	0	5	75	1	120	15	60	1	3	50	1	0	1	40	0	0	75	2	120	12	80	1	10
59	3	100	2	60	1	30	0	0	75	0	60	3	40	0	4	50	0	120	1	40	0	1	100	1	0	12	100	1	4	100	0	60	9	100	1	10
60	1	75	2	60	15	60	1	4	100	1	0	9	80	1	0	50	1	60	15	100	1	0	75	0	120	9	60	1	4	50	2	0	15	60	1	10
61	3	50	2	120	3	20	0	3	75	1	60	9	80	1	4	75	0	0	9	100	1	5	75	2	0	9	60	1	0	100	1	120	5	30	0	11
62	2	75	2	120	15	80	1	2	50	0	120	3	40	0	0	75	2	60	12	100	1	3	100	0	60	3	40	0	0	75	2	0	1	20	0	11
63	1	75	0	0	9	80	1	5	75	1	120	5	20	0	5	50	1	0	5	30	0	1	50	2	0	15	60	1	4	50	0	60	1	40	0	11
64	1	100	1	60	5	40	0	1	50	0	0	15	80	1	1	100	0	60	1	20	0	2	50	0	60	1	40	0	2	50	1	60	12	100	1	11
65	5	50	1	60	1	20	0	0	100	2	60	3	20	0	3	100	1	120	5	20	0	4	100	1	120	5	30	0	5	75	0	0	15	60	1	11
66	3	100	0	0	12	100	1	4	100	2	0	9	100	1	2	50	2	120	12	80	1	0	75	1	120	12	80	1	4	100	2	120	12	80	1	11
67	1	100	0	0	12	100	1	3	50	0	120	3	20	0	3	100	2	120	1	40	0	3	75	2	120	9	60	1	0	75	0	60	5	30	0	12
68	0	50	0	0	9	60	1	2	100	2	0	15	80	1	4	100	1	0	12	100	1	1	100	0	0	3	40	0	0	50	1	120	1	20	0	12
69	5	75	2	120	5	30	0	1	100	2	60	9	60	1	2	50	2	0	1	20	0	3	50	0	0	9	100	1	5	100	0	0	3	40	0	12
70	5	50	1	60	1	20	0	5	50	0	0	1	40	0	1	75	0	120	15	80	1	2	50	2	120	5	20	0	1	100	2	0	9	100	1	12
71	3	100	2	120	15	100	1	3	75	1	120	15	100	1	5	50	0	60	5	20	0	5	100	1	60	5	30	0	5	50	2	120	15	60	1	12
72	1	75	1	60	3	30	0	1	75	1	60	3	30	0	0	75	1	60	12	80	1	1	75	1	60	12	80	1	4	75	1	60	12	80	1	12
73	3	75	0	60	15	100	1	5	75	2	60	1	40	0	4	75	1	60	1	40	0	0	50	2	120	15	100	1	2	75	1	60	5	20	0	13
74	4	50	0	60	3	30	0	0	50	1	0	15	60	1	0	75	0	120	3	40	0	4	100	0	0	1	30	0	5	100	2	0	3	30	0	13
75	2	100	1	120	1	20	0	4	75	2	120	3	20	0	2	50	0	60	12	80	1	1	75	2	120	1	40	0	3	50	2	120	1	40	0	13
76	2	50	2	0	3	30	0	0	100	0	0	12	100	1	5	100	1	0	1	20	0	0	75	1	60	15	60	1	1	50	1	120	12	100	1	13
77	4	75	1	120	15	100	1	2	100	0	60	5	30	0	1	50	2	0	15	80	1	5	100	1	60	12	60	1	0	100	0	0	15	80	1	13
78	0	100	2	0	9	60	1	4	50	1	120	9	80	1	3	100	2	120	15	60	1	5	50	0	0	3	30	0	4	75	0	60	9	60	1	13
79	1	50	0	120	1	40	0	1	75	0	60	12	60	1	5	75	0	0	15	60	1	3	50	0	120	12	80	1	1	100	1	120	1	20	0	14
80	3	75	0	120	12	60	1	3	50	2	120	15	80	1	1	75	2	60	1	20	0	1	75	1	0	5	20	0	4	75	2	0	3	40	0	14
81	5	75	1	0	5	20	0	0	75	0	120	1	40	0	3	50	1	0	12	100	1	2	100	0	120	9	80	1	5	50	0	60	3	30	0	14
82	5	100	2	60	9	80	1	5	50	2	0	1	20	0	0	100	0	60	5	30	0	3	100	1	0	12	60	1	0	50	1	60	15	80	1	14
83	0	50	2	60	12	100	1	1	100	1	60	5	40	0	2	100	1	120	3	30	0	2	75	2	60	1	40	0	0	75	0	0	9	100	1	14
84	1	100	1	0	5	30	0	5	100	1	0	12	80	1	4	50	2	120	9	100	1	4	50	2	60	5	40	0	5	100	2	120	15	60	1	14
85	3	50	1	120	1	30	0	4	50	1	120	12	80	1	0	100	0	0	3	40	0	0	50	1	120	5	40	0	1	50	0	0	5	30	0	15
86	2	75	2	0	3	40	0	3	50	1	0	5	40	0	1	75	2	0	15	80	1	4	75	1	120	9	100	1	5	100	1	60	1	20	0	15
87	4	50	2	0	15	80	1	1	75	2	120	1	20	0	5	75	1	120	9	80	1	2	100	0	60	3	20	0	5	75	2	120	5	40	0	15
88	1	100	1	120	5	20	0	5	100	0	60	9	100	1	3	100	2	60	5	40	0	1	100	2	0	9	80	1	1	75	1	120	12	100	1	15
89	5	100	0	60	9	60	1	2	75	2	60	15	80	1	2	50	0	120	12	60	1	3	75	0	60	15	60	1	0	100	0	60	12	60	1	15
90	0	75	0	60	12	100	1	0	100	0	0	3	20	0	4	50	1	60	1	20	0	5	50	2	0	3	30	0	3	50	2	0	9	80	1	15
91	2	75	2	60	5	20	0	4	75	0	0	5	20	0	2	50	2	120	3	40	0	5	50	1	0	5	40	0	1	75	1	120	5	30	0	16
92	4	100	0	120	12	80	1	0	75	0	60	9	80	1	3	100	1	120	15	80	1	0	75	1	0	1	40	0	2	50	2	0	1	20	0	16
93	0	50	2	60	9	100	1	2	50	2	120	12	80	1	1	50	1	0	5	40	0	3	100	2	60	9	80	1	4	100	2	60	3	40	0	16
94	3	75	0	120	3	20	0	4	100	1	0	3	40	0	5	100	0	60	9	80	1	4	100	0	120	3	20	0	2	100	0	60	15	100	1	16
95	4	50	1	0	1	40	0	2	100	1	120	1	20	0	4	75	2	60	3	20	0	2	75	0	120	15	60	1	3	50	1	0	12	60	1	16
96	2	100	1	0	15	80	1	3	50	2	60	15	100	1	0	75	0	0	9	60	1	1	50	2	60	12	80	1	3	75	0	120	9	80	1	16
97	4	75	1	60	3	30	0	5	100	2	120	15	100	1	4	100	2	120	12	60	1	3	50	1	120	12	80	1	2	50	2	60	1	40	0	17
98	0	100	1	60	15	100	1	1	75	1	0	1	20	0	0	100	1	0	15	60	1	4	75	2	0	5	20	0	2	100	0	120	3	30	0	17
99	2	100	2	120	9	100	1	4	100	2	0	3	30	0	5	75	0	120	1	40	0	2	100	1	120	9	80	1	3	75	1	0	3	20	0	17
100	2	50	0	0	5	20	0	0	75	1	60	12	100	1	1	50	2	0	3	30	0	0	100	2	0	3	20	0	1	75	2	0	15	60	1	17
101	3	75	0	0	1	30	0	2	50	0	120	1	30	0	3	50	0	60	1	30	0	5	75	0	60	1	40	0	4	100	1	120	9	80	1	17
102	4	50	2	120	12	60	1	3	50	0	60	15	60	1	2	75	1	60	15	100	1	1	50	0	60	15	100	1	3	50	0	60	15	100	1	17
103	2	50	2	120	12	60	1	5	75	1	0	1	30	0	3	50	1	60	3	20	0	4	50	1	60	12	80	1	2	100	1	0	1	40	0	18
104	1	75	2	120	9	80	1	1	50	0	60	12	100	1	4	50	0	120	15	60	1	2	75	2	120	5	20	0	2	75	2	60	3	30	0	18
105	3	100	1	60	5	40	0	0	50	0	120	15	80	1	2	75	1	120	3	30	0	4	100	2	120	12	60	1	4	50	1	120	5	20	0	18
106	3	75	0	0	1	30	0	4	75	1	60	5	20	0	1	100	2	60	9	100	1	0	100	1	60	1	30	0	0	50	0	120	12	60	1	18
107	4	50	1	60	15	60	1	5	100	2	0	12	60	1	5	75	2	0	1	30	0	3	75	0	0	1	40	0	4	75	0	60	9	80	1	18
108	2	100	0	0	3	40	0	0	100	2	120	1	40	0	0	100	0	0	15	100	1	2	50	0	0	15	100	1	3	100	2	0	15	100	1	18

Figure A-2 Orthogonal design

Appendix B – Adopted Efficient Designs

Ngene syntax efficient design – 18 choice situations

```
Design
;alts = alt1, alt2, alt3, alt4, alt5
;rows = 18
;bdraws = halton(1000)
;eff = (mnl,d,mean)
;block=3

;cond:

if (alt1.dest = 0,alt1.Egp = [1,3,5]),
if (alt1.dest = 1,alt1.Egp = [9,12,15]),
if (alt1.dest = 0,alt1.Egt = [20,30,40]),
if (alt1.dest = 1,alt1.Egt = [60,80,100]),
if (alt2.dest = 0,alt2.Egp = [1,3,5]),
if (alt2.dest = 1,alt2.Egp = [9,12,15]),
if (alt2.dest = 0,alt2.Egt = [20,30,40]),
if (alt2.dest = 1,alt2.Egt = [60,80,100]),
if (alt3.dest = 0,alt3.Egp = [1,3,5]),
if (alt3.dest = 1,alt3.Egp = [9,12,15]),
if (alt3.dest = 0,alt3.Egt = [20,30,40]),
if (alt3.dest = 1,alt3.Egt = [60,80,100]),
if (alt4.dest = 0,alt4.Egp = [1,3,5]),
if (alt4.dest = 1,alt4.Egp = [9,12,15]),
if (alt4.dest = 0,alt4.Egt = [20,30,40]),
if (alt4.dest = 1,alt4.Egt = [60,80,100]),
if (alt5.dest = 0,alt5.Egp = [1,3,5]),
if (alt5.dest = 1,alt5.Egp = [9,12,15]),
if (alt5.dest = 0,alt5.Egt = [20,30,40]),
if (alt5.dest = 1,alt5.Egt = [60,80,100])

;model:

U(alt1)=Air.Dummy[0|0|0|0|0] * Air
+ Pr[(n,-0.040,0.0048)] * Pr[50,75,100]
+ Dep.Dummy[(n,0,0.22) | (n,0.614,.21)] * Dep
+ Tr[(n,-.360,.11)]*Tr(0,1,2]
+ dest[0]*dest[0,1]
+ Egp[(n,0,0.03)]*Egp[1,3,5,9,12,15]
+ Egt[(n,-.018,.0057)]*Egt[20,30,40,60,80,100]/

U(alt2)=Air.Dummy * Air
+ Pr * Pr[50,75,100]
+ Dep.Dummy * Dep
+ Tr * Tr[0,1,2]
+ dest * dest[0,1]
+ Egp * Egp[1,3,5,9,12,15]
+ Egt * Egt[20,30,40,60,80,100]/

U(alt3)=Air.Dummy *Air
+ Pr * Pr[50,75,100]
+ Dep.Dummy * Dep
+ Tr * Tr[0,1,2]
+ dest * dest[0,1]
+ Egp * Egp[1,3,5,9,12,15]
+ Egt * Egt[20,30,40,60,80,100]/

U(alt4)=Air.Dummy *Air
+ Pr * Pr[50,75,100]
+ Dep.Dummy * Dep
+ Tr * Tr[0,1,2]
+ dest * dest[0,1]
+ Egp * Egp[1,3,5,9,12,15]
+ Egt * Egt[20,30,40,60,80,100]/

U(alt5)=Air.Dummy *Air
```

```
+ Pr * Pr[50,75,100]
+ Dep.Dummy * Dep
+ Tr * Tr[0,1,2]
+ dest * dest[0,1]
+ Egp * Egp[1,3,5,9,12,15]
+ Egt * Egt[20,30,40,60,80,100]$
```

Ngene syntax efficient design – 108 choice situations

```
Design
;alts = alt1, alt2, alt3, alt4, alt5
;rows = 108
;bdraws = halton(1000)
;eff = (mnl,d,mean)
;block=18

(The rest is the same as for the efficient design with 18 choice situations)
```

Choice situation	alt1.air	alt1.pr	alt1.dpt	alt1.tran	alt1.dest	alt1.egp	alt1.egt	alt2.air	alt2.pr	alt2.dpt	alt2.tran	alt2.dest	alt2.egp	alt2.egt	alt3.air	alt3.pr	alt3.dpt	alt3.tran	alt3.dest	alt3.egp	alt3.egt
1	2	100	1	60	0	3	30	1	50	2	0	1	15	60	2	100	2	60	0	1	40
2	5	75	1	60	0	5	20	2	100	2	60	0	3	40	4	75	2	0	1	12	60
3	0	75	1	0	1	9	100	4	75	0	0	0	5	40	3	75	1	120	1	9	60
4	2	100	2	60	1	12	100	3	50	2	120	0	5	20	1	50	0	0	1	9	100
5	1	50	0	0	0	1	30	4	50	1	120	1	15	100	5	50	2	60	1	15	60
6	3	50	0	60	1	15	80	0	50	1	60	1	9	60	5	75	1	120	0	3	30
7	4	50	2	120	0	1	40	0	100	0	60	0	3	30	1	75	0	60	1	15	80
8	1	100	2	0	0	3	20	2	50	1	120	1	9	100	3	100	0	120	0	5	20
9	5	50	2	120	1	12	80	4	100	0	60	0	1	20	0	75	1	120	0	5	40
10	0	100	0	120	1	9	80	3	100	2	0	0	3	20	1	50	2	60	1	12	80
11	1	50	2	120	0	1	20	0	100	1	120	0	5	30	5	100	0	0	0	5	30
12	4	50	0	120	1	15	60	2	100	0	60	1	12	60	0	100	1	0	0	1	30
13	3	75	0	0	1	12	60	5	75	2	120	0	1	30	0	50	1	60	1	12	100
14	2	75	1	120	0	5	40	5	50	0	0	1	15	80	4	100	1	120	0	3	20
15	5	100	1	60	1	9	60	3	75	0	0	1	9	80	4	50	2	0	1	15	100
16	4	75	0	0	0	3	40	1	75	2	0	1	12	80	2	75	2	60	0	1	20
17	3	100	1	60	1	15	100	1	75	1	120	0	1	40	2	50	0	0	1	9	80
18	0	75	2	0	0	5	30	5	75	1	60	1	12	100	3	100	0	120	0	3	40

Choice situation	alt4.air	alt4.pr	alt4.dpt	alt4.tran	alt4.dest	alt4.egp	alt4.egt	alt5.air	alt5.pr	alt5.dpt	alt5.tran	alt5.dest	alt5.egp	alt5.egt	Block
1	4	75	1	120	0	1	30	0	50	0	60	0	1	30	1
2	3	100	0	60	0	3	40	3	50	0	120	1	9	100	1
3	5	75	0	60	1	12	80	2	50	2	120	1	12	60	1
4	5	50	0	120	1	9	80	4	75	1	0	1	15	60	1
5	0	50	2	120	1	15	80	4	100	2	0	0	1	20	1
6	2	100	2	0	0	5	30	2	50	0	120	0	5	40	1
7	2	100	1	0	0	3	20	5	100	1	60	1	12	100	2
8	0	100	1	0	1	15	60	3	100	2	120	0	5	20	2
9	4	75	2	0	0	1	40	1	75	1	0	1	15	80	2
10	0	50	0	120	1	9	60	5	75	1	0	0	3	40	2
11	2	75	0	60	0	5	40	3	50	1	0	1	9	80	2
12	5	50	2	0	1	12	100	1	100	2	60	0	3	30	2
13	1	100	0	60	0	3	30	2	100	2	0	0	3	20	3
14	3	50	2	0	1	9	100	1	100	2	60	1	12	80	3
15	1	100	1	60	0	1	20	0	75	0	120	0	5	30	3
16	3	50	1	120	0	5	20	5	75	0	60	0	1	40	3
17	4	75	2	60	1	12	60	0	75	1	60	1	15	100	3
18	1	75	1	120	1	15	100	4	50	0	120	1	9	60	3

Figure B-1 Efficient design - 18 choice situations

Choice situation	alt1.air	alt1.pr	alt1.dpt	alt1.tran	alt1.dest	alt1.egp	alt1.egt	alt2.air	alt2.pr	alt2.dpt	alt2.tran	alt2.dest	alt2.egp	alt2.egt	alt3.air	alt3.pr	alt3.dpt	alt3.tran	alt3.dest	alt3.egp	alt3.egt	alt4.air	alt4.pr	alt4.dpt	alt4.tran	alt4.dest	alt4.egp	alt4.egt	alt5.air	alt5.pr	alt5.dpt	alt5.tran	alt5.dest	alt5.egp	alt5.egt	Block
1	5	50	0	0	1	15	60	0	100	1	60	0	1	30	1	75	2	120	0	1	30	4	50	1	120	1	9	100	3	100	2	60	0	3	40	1
2	4	75	0	60	1	12	60	5	75	1	60	1	12	100	1	50	1	120	1	9	60	3	100	1	60	0	3	30	2	50	2	0	0	5	40	1
3	2	75	1	120	0	1	20	1	75	1	60	1	12	60	3	50	2	0	0	5	40	4	50	0	120	1	9	80	0	75	0	120	0	1	20	1
4	1	50	0	120	1	12	60	0	50	2	0	1	12	100	5	100	0	120	0	1	20	3	75	2	60	1	12	100	4	100	1	0	0	3	20	1
5	4	75	0	120	0	1	30	3	75	0	60	1	12	60	2	75	2	0	0	1	20	1	50	1	60	1	15	100	5	100	0	60	0	3	30	1
6	0	100	1	0	0	1	40	1	100	2	0	0	1	30	3	50	1	120	1	15	60	5	75	2	60	1	9	80	2	75	0	0	0	1	40	1
7	3	75	0	60	1	9	80	0	75	2	120	0	1	40	4	50	1	120	1	15	60	1	75	2	0	0	1	20	2	50	0	0	1	15	100	2
8	2	100	0	60	0	3	30	2	100	1	60	0	3	40	0	50	2	0	1	9	60	4	50	0	0	1	15	100	5	75	1	120	0	5	30	2
9	3	75	0	60	1	12	80	2	50	2	120	0	1	40	1	100	0	60	0	3	40	0	100	0	60	0	3	30	5	75	1	0	1	15	60	2
10	4	100	2	120	0	1	30	0	100	1	0	0	5	20	5	50	0	0	1	9	100	2	100	2	60	1	12	80	3	50	2	120	1	15	80	2
11	4	100	0	60	0	3	40	0	100	0	60	1	12	60	3	100	1	60	1	12	80	2	50	1	120	1	9	100	1	75	2	0	0	5	20	2
12	4	50	0	0	1	9	100	3	75	0	120	1	12	60	5	50	1	120	1	12	100	1	100	1	120	0	5	20	0	75	2	0	0	5	20	2
13	1	100	1	60	1	12	100	5	75	2	60	1	15	80	2	100	1	0	1	12	60	0	100	2	0	0	5	20	4	50	0	120	0	1	40	3
14	1	100	1	60	0	3	30	3	100	0	60	0	5	30	0	50	1	0	1	9	100	5	50	2	0	1	15	60	2	50	0	120	0	1	30	3
15	4	75	2	60	0	5	30	1	50	1	0	1	9	60	5	75	0	120	0	5	30	2	75	2	60	0	3	40	0	50	0	120	0	5	40	3
16	0	100	0	60	1	12	80	4	50	2	120	1	9	60	2	75	0	0	1	15	100	3	100	1	0	0	5	40	5	100	0	60	1	12	100	3
17	1	50	2	0	1	9	100	5	50	0	120	1	15	60	3	100	1	0	0	5	20	0	100	0	60	0	3	40	4	75	2	120	0	3	40	3
18	0	50	2	120	0	1	30	2	75	1	0	1	15	60	3	75	0	120	1	12	80	4	50	0	60	1	12	100	5	100	0	60	0	1	30	3
19	5	50	2	120	1	9	60	1	100	1	0	0	5	20	0	75	2	60	1	12	60	4	50	0	120	0	1	40	2	50	1	0	1	15	100	4
20	4	75	2	120	1	12	80	0	100	2	60	0	3	40	2	100	1	0	0	5	20	3	50	0	120	0	5	40	5	75	2	0	1	9	60	4
21	4	50	2	120	0	5	30	5	75	1	0	1	9	60	3	100	0	60	0	5	40	1	75	0	0	1	12	100	2	75	0	60	1	9	80	4
22	1	50	2	120	0	5	40	4	100	0	60	0	5	30	5	75	1	0	1	9	60	3	100	0	60	0	3	40	2	75	0	60	1	12	100	4
23	3	100	1	60	0	1	30	1	50	2	120	0	1	20	4	50	1	0	1	15	100	5	75	0	120	1	15	60	2	100	0	60	0	1	30	4
24	4	100	0	60	0	1	30	3	100	0	60	0	3	40	1	100	0	60	0	1	30	0	75	1	0	1	15	60	2	50	2	120	0	1	40	4
25	1	100	1	0	1	9	80	4	50	0	60	1	9	80	5	50	2	120	0	5	40	3	100	1	0	0	5	20	2	75	0	120	1	12	80	5
26	2	100	2	0	0	5	20	5	50	1	120	0	1	20	4	75	2	60	1	12	80	1	50	0	0	1	15	80	3	50	2	0	1	12	100	5
27	2	50	2	0	1	15	100	3	75	0	0	1	12	60	1	100	1	120	0	3	40	5	75	2	120	0	1	20	0	100	1	120	0	1	30	5
28	2	75	0	120	0	5	30	5	100	1	60	0	3	40	3	100	0	60	0	3	30	4	50	2	0	1	9	60	0	50	1	120	1	15	100	5
29	0	50	2	60	1	15	60	2	100	1	60	0	3	30	5	50	0	0	0	1	40	3	50	1	120	1	15	80	1	100	1	60	1	15	80	5
30	2	50	1	120	1	15	100	1	100	2	60	0	3	30	5	75	2	120	1	9	80	0	75	0	0	0	1	20	3	100	1	60	1	12	80	5
31	5	50	0	120	1	9	60	4	50	2	120	1	15	100	0	100	2	60	0	3	40	3	100	1	0	0	5	20	1	75	2	60	1	9	100	6
32	5	50	1	120	1	15	60	1	50	0	0	0	1	40	0	50	2	120	0	1	30	3	75	1	120	1	15	80	2	100	1	0	0	1	20	6
33	1	50	2	60	1	9	100	5	75	0	120	1	9	80	2	100	2	60	1	12	80	3	75	0	120	1	12	60	4	100	1	0	0	5	20	6
34	0	75	1	120	1	15	80	3	50	1	120	0	5	40	1	50	0	0	1	9	60	2	75	2	0	0	5	30	4	100	2	60	0	3	30	6
35	4	75	2	120	0	3	40	3	75	2	120	1	12	80	0	50	1	120	1	15	100	1	75	0	0	0	1	20	5	100	1	60	1	12	80	6
36	5	50	2	60	1	15	100	4	100	0	120	0	3	40	2	100	0	60	0	1	30	3	100	1	0	1	9	60	0	100	0	120	0	3	30	6
37	4	100	2	0	0	5	20	3	75	1	0	1	9	100	5	75	0	120	0	1	40	1	50	0	120	1	15	60	0	100	1	60	1	9	80	7
38	5	75	0	0	0	5	20	1	75	2	60	1	9	80	3	75	2	60	1	9	80	4	75	1	120	1	15	80	2	50	1	120	1	9	100	7
39	4	100	1	120	0	5	40	5	100	2	60	0	1	40	1	50	0	120	1	15	60	0	75	2	0	1	9	80	3	100	1	0	0	1	40	7
40	4	100	1	0	0	3	20	1	50	0	120	0	5	40	3	100	1	60	0	3	40	0	50	2	0	1	9	60	2	50	2	60	1	12	100	7
41	0	100	2	60	0	5	30	2	50	0	0	1	15	100	3	100	2	60	0	3	30	5	100	1	60	0	1	30	4	75	1	120	1	9	60	7
42	5	75	2	120	0	1	20	4	75	0	0	0	3	20	2	50	1	0	1	15	100	0	75	1	60	1	9	100	3	50	2	120	1	9	60	7
43	2	50	0	120	1	9	60	4	50	1	0	1	15	100	1	100	0	60	0	3	20	5	100	2	60	0	1	20	3	100	2	60	0	3	20	8
44	2	50	1	0	1	9	80	4	100	0	60	0	3	30	3	50	2	60	1	15	80	5	75	2	120	0	5	30	0	50	0	120	0	5	30	8
45	1	75	0	0	0	5	40	3	75	1	60	1	12	100	0	50	1	120	1	9	60	2	100	0	60	0	3	30	4	75	2	0	0	5	30	8
46	2	100	0	60	0	3	30	3	75	0	60	1	9	80	0	50	1	120	1	15	100	1	50	2	120	1	9	80	4	100	1	0	0	5	20	8
47	5	75	2	120	0	5	40	1	100	1	60	1	15	80	2	100	1	60	1	12	60	0	75	0	0	0	1	40	4	75	1	120	1	12	60	8
48	5	75	2	0	0	5	30	3	50	0	0	1	9	100	4	100	0	60	0	3	30	0	50	1	120	1	12	60	1	75	0	60	1	15	100	8
49	5	75	1	120	1	9	80	2	100	2	60	0	3	40	0	50	2	120	0	1	40	1	100	1	60	0	1	40	3	50	0	0	1	15	60	9

Figure B-2 Efficient design - 108 choice situation (part 1)

50	5	100	2	60	0	5	30	3	75	1	120	0	1	40	1	75	2	0	0	5	40	2	50	0	0	1	9	60	0	75	1	120	1	15	80	9
51	1	50	0	120	1	15	80	2	100	1	120	0	3	20	0	100	0	0	0	5	30	5	100	1	120	0	3	20	3	50	2	0	1	9	80	9
52	3	75	0	60	0	1	40	4	50	2	0	1	15	60	1	50	1	120	1	9	100	2	100	0	60	0	5	30	0	75	1	120	0	1	40	9
53	3	50	1	60	1	9	100	2	75	0	0	0	5	20	5	75	2	120	1	12	60	1	75	2	120	1	12	80	4	100	2	60	0	3	20	9
54	4	100	0	0	0	3	20	5	75	1	60	1	9	100	2	50	2	120	0	1	20	4	75	0	60	1	12	80	0	50	1	0	1	15	100	9
55	3	75	2	0	1	12	60	2	50	2	0	1	12	60	1	100	0	0	0	1	30	0	50	1	120	0	5	40	4	100	0	60	0	3	30	10
56	1	100	1	60	1	12	60	3	50	2	120	0	5	20	4	100	0	0	0	3	20	2	75	0	120	0	5	30	5	50	1	0	1	9	100	10
57	0	75	1	0	1	12	100	3	75	2	120	0	5	40	5	75	1	120	1	15	80	2	75	0	0	0	1	20	1	50	2	120	1	12	60	10
58	0	50	0	0	1	9	60	5	50	1	120	1	15	100	2	100	1	0	0	5	30	4	100	2	60	0	3	30	1	50	2	120	0	3	40	10
59	0	50	2	0	1	15	60	4	75	0	120	0	3	30	5	75	0	120	0	3	40	3	100	1	0	0	1	30	1	50	1	120	1	9	100	10
60	3	100	2	60	0	3	30	5	50	1	0	1	9	100	1	50	0	120	1	15	60	4	100	2	60	0	1	30	2	75	2	120	0	3	20	10
61	5	100	1	0	0	1	20	1	100	0	60	0	3	20	4	50	0	120	0	5	40	2	75	2	60	1	12	60	3	50	2	0	1	9	100	11
62	5	100	1	0	1	12	60	2	50	2	0	1	15	100	1	75	2	60	1	15	80	4	100	1	60	0	3	40	3	50	0	120	0	1	20	11
63	1	100	0	60	0	3	40	5	100	1	0	0	5	30	0	75	0	0	1	9	80	4	50	1	120	1	9	100	3	50	2	120	1	15	60	11
64	3	50	0	120	1	15	60	1	50	2	0	1	9	100	2	100	1	120	0	3	20	5	100	2	60	0	1	30	0	100	1	60	0	5	30	11
65	0	75	2	60	0	1	30	3	50	0	120	0	1	30	1	100	2	60	0	3	30	2	100	0	0	0	3	20	4	50	1	0	1	15	80	11
66	4	50	2	120	1	9	80	5	50	1	120	1	15	80	0	100	1	0	0	1	20	3	75	2	120	0	3	30	1	75	0	0	0	5	30	11
67	1	75	2	60	1	15	80	0	75	2	60	1	12	100	4	50	0	0	1	9	100	2	75	1	120	0	5	20	5	100	1	60	1	15	80	12
68	2	75	1	120	1	12	80	0	50	1	0	1	15	100	5	50	0	120	0	1	20	3	75	2	0	0	3	20	1	75	2	60	1	12	60	12
69	0	100	1	60	1	15	80	1	100	2	0	0	3	20	4	75	1	120	1	15	80	2	100	1	120	0	3	20	5	50	0	0	1	9	100	12
70	2	50	2	120	1	12	60	5	100	2	0	0	3	20	0	50	1	0	1	9	100	1	100	1	120	0	5	30	4	50	0	120	0	5	20	12
71	1	100	2	60	0	3	20	4	75	0	0	0	1	20	2	50	1	60	1	15	100	3	75	2	120	1	12	60	0	75	1	120	1	12	100	12
72	5	50	1	0	1	15	100	2	50	2	120	0	1	30	4	100	0	60	0	5	30	3	100	0	60	0	3	20	1	75	0	0	1	12	60	12
73	2	50	0	0	1	9	100	0	100	0	60	1	12	80	4	75	2	0	1	15	80	1	100	1	120	0	3	20	3	100	1	120	0	5	20	13
74	2	100	1	0	0	1	20	0	50	0	0	1	9	100	4	75	0	60	1	12	100	1	100	2	60	1	15	80	5	50	2	120	1	15	60	13
75	2	75	1	120	0	5	20	0	75	0	120	1	15	80	1	50	2	0	1	9	100	5	75	0	0	1	15	60	3	100	2	60	0	1	30	13
76	0	100	0	120	0	3	20	3	75	1	120	0	5	40	5	100	1	0	0	5	20	4	50	2	0	1	9	80	2	75	0	60	1	12	80	13
77	4	75	0	0	1	12	100	0	75	0	120	1	12	100	3	75	1	120	1	12	60	5	100	2	60	0	3	40	1	100	2	0	0	5	40	13
78	4	100	1	60	0	3	40	2	100	1	60	0	1	40	1	50	0	0	1	15	60	3	50	2	120	0	1	30	0	100	0	60	0	1	30	13
79	3	50	1	0	1	9	60	4	75	2	120	0	5	30	2	75	2	60	0	5	30	5	50	0	120	0	5	30	0	75	2	60	0	3	40	14
80	3	50	1	120	0	1	40	4	75	0	0	1	12	80	5	75	2	0	0	5	20	2	50	0	0	1	15	60	1	75	1	60	1	12	100	14
81	3	50	2	0	0	1	40	0	75	0	120	0	1	20	2	50	0	60	1	12	100	5	75	1	60	1	12	100	1	50	1	120	1	15	60	14
82	3	75	0	0	0	5	20	0	100	2	60	0	5	40	5	50	2	0	1	12	100	4	50	1	120	1	9	60	1	100	1	60	0	3	40	14
83	0	100	2	60	0	3	40	2	75	2	0	1	15	80	3	75	2	0	1	15	80	1	100	1	0	1	12	60	5	50	0	120	0	1	40	14
84	2	75	0	60	1	12	100	0	75	2	120	0	1	20	1	100	1	0	0	1	20	3	50	0	0	1	15	100	4	100	1	120	0	1	20	14
85	1	75	0	0	1	12	100	2	100	1	0	0	5	40	3	100	0	120	0	3	30	0	50	2	120	1	9	60	4	100	0	60	1	15	80	15
86	5	75	2	120	0	1	20	4	100	1	60	0	1	30	0	100	2	60	0	3	20	1	50	1	0	1	15	100	3	50	0	120	0	3	20	15
87	4	75	0	60	1	12	80	1	50	1	120	0	5	30	3	75	0	0	0	1	30	0	75	0	60	1	12	80	5	50	2	0	1	9	60	15
88	3	100	1	120	0	3	20	5	50	1	120	1	15	80	0	75	0	0	1	12	60	1	50	2	0	0	1	40	2	100	0	60	0	3	30	15
89	1	50	1	120	1	15	100	2	100	1	60	1	9	80	0	75	2	60	1	9	80	4	75	2	60	1	15	80	5	75	0	0	0	1	20	15
90	3	75	0	0	1	15	100	2	100	1	0	1	12	80	5	50	2	120	0	5	40	0	50	0	120	1	9	80	4	100	1	0	0	1	20	15
91	3	75	1	0	0	1	40	2	50	0	120	0	1	20	5	100	0	60	0	3	30	0	100	1	60	0	3	30	4	50	2	0	1	15	60	16
92	0	50	1	120	1	15	80	1	75	0	60	1	12	80	2	100	1	60	0	3	20	4	50	2	0	0	5	40	5	50	0	0	1	9	60	16
93	3	100	0	60	1	15	80	4	100	1	60	1	9	60	0	75	1	120	1	9	80	1	100	0	60	1	12	80	5	75	2	0	0	5	40	16
94	1	75	1	0	0	3	40	4	100	0	60	0	3	30	3	75	0	120	0	5	30	5	75	2	120	1	12	80	2	50	2	120	1	9	60	16
95	2	75	1	60	1	9	80	0	50	0	120	0	5	20	4	75	2	60	0	1	40	5	50	1	0	1	15	100	3	75	2	0	1	12	60	16
96	0	75	2	0	1	12	100	5	100	2	60	0	3	30	4	50	1	120	1	15	60	1	75	0	0	0	1	40	5	100	2	60	0	3	40	16
97	0	50	0	60	0	5	40	5	50	2	60	1	9	80	4	100	1	0	0	5	20	2	100	1	60	1	12	60	3	75	2	60	1	9	80	17
98	2	50	2	0	1	15	60	3	50	0	0	1	12	100	1	50	1	120	0	1	40	4	100	0	60	0	1	30	0	75	0	60	1	12	80	17
99	2	50	0	120	0	3	20	0	100	1	0	0	1	30	4	75	2	0	0	3	40	5	75	1	0	1	12	100	1	75	1	120	1	12	60	17
100	0	100	0	0	0	1	20	3	50	2	0	1	15	60	4	100	1	60	0	1	40	2	50	1	120	1	9	100	5	100	1	60	0	5	40	17
101	1	50	1	120	0	1	40	0	50	0	0	1	15	60	2	50	2	60	1	12	100	5	75	2	60	1	12	100	3	75	2	0	1	9	80	17
102	3	50	1	0	1	9	100	4	75	0	120	0	5	20	2	75	2	60	1	12	80	0	75	2	0	0	5	20	1	75	0	120	0	3	30	17
103	5	100	0	60	1	12	80	1	100	2	60	0	5	30	3	75	2	60	1	12	100	0	50	0	120	0	5	40	2	75	1	0	1	9	60	18
104	1	75	1	120	1	9	60	2	100	2	120	0	3	20	0	100	2	60	0	3	20	5	100	2	60	0	5	40	4	50	0	0	1	15	100	18
105	5	100	2	60	0	3	30	1	50	1	120	1	9	60	4	75	2	60	1	12	80	2	50	2	0	1	15	100	0	75	0	0	0	5	40	18
106	3	100	2	60	0	3	30	1	75	2	0	1	15	60	4	100	1	0	1	9	60	2	50	0	120	0	1	40	0	100	1	60	1	12	80	18
107	0	100	2	0	0	5	20	5	75	2	120	1	12	80	3	75	0	120	0	5	20	4	50	1	60	1	9	100	1	100	2	60	0	3	20	18
108	5	75	1	120	0	5	40	4	75	2	0	1	9	60	2	75	0	60	1	9	80	0	100	0	60	0	3	40	1	100	0	60	0	5	30	18

Figure B-3 Efficient design - 108 choice situation (part 2)

Appendix C – Estimation Methodology

Several statistical approaches for estimating parameters of choice models exist. The following paragraph discusses the method of maximised likelihood estimation, which is the most commonly used estimation method (Louviere *et al.* 2000). The maximum likelihood estimates are that set of population parameters that generate the observed sample most often i.e. that predicts the observations with the largest probability. The application of such a procedure can be helpful in the estimation of the parameter estimates and their asymptotic t-value, or in measuring the goodness of fit for the model as a whole (Louviere *et al.* 2000).

According to the multinomial probability distribution, the probability $P=[P_{iqs}]$ of person n choosing the alternative that was actually chosen in choice situation s can be expressed as:

$$\prod_{j=1}^{J}(P_{iqs})^{y_{iqs}} , \tag{12.1}$$

where $y=[y_{iqs}]=1$ if individual n chooses alternative j and zero otherwise. Note that since $y_{iqs}=0$ for all non-chosen alternatives and P_{iqs} raised to the power of zero equals 1, this term is simply the probability of the chosen alternative.

The probability of observing a sequence of choices over all choice situations s , is equal to

$$\prod_{q}^{Q}\prod_{s}^{S}(P_{iqs})^{y_{iqs}}. \tag{12.2}$$

Finally observing a certain sequence over the whole population has a probability of

$$\prod_{i}^{I}\prod_{q}^{Q}\prod_{s}^{S}(P_{iqs})^{y_{iqs}}, \tag{12.3}$$

We would like to maximize this probability, which is called the likelihood. Instead of maximizing the likelihood, typically the logarithm of the likelihood is maximized, giving the same outcome (this in order to deal with large numbers instead of very small numbers). The log-likelihood can be defined as

$$LL(\beta)=\log\prod_{i=1}^{I}\prod_{q=1}^{Q}\prod_{s=1}^{S}(P_{iqs}(\beta))^{y_{iqs}} , \tag{12.4}$$

where Q denotes the total number of respondents, S is the total number of choice situations faced by respondent n, I is the total number of alternatives in each choice set, and $P=[P_{iqs}]$ (for $q=1,\ldots,Q$, $s=1,\ldots,S$, and $i=1,\ldots,I$) is the probability of choosing alternative $i\in S$, and β is a vector of parameter to be estimated.

However, some models (i.e. mixed logit models) are estimating the parameters of assumed distributions of the parameters instead of fixed parameter estimates. This is because the parameter estimates across the poplation may vary due to the heterogeneity of respondents. Therefore, the parameters are assumed random following a certain probability distribution, and the expected log-likelihood will be maximized. This leads to the following (expected) log-likelihood function to be maximized (Revelt & Train 1998, Bliemer & Rose 2008):

$$LL(\theta) = \log\left[E\left(\prod_{i=1}^{I}\prod_{q=1}^{Q}\prod_{s=1}^{S}(P_{iqs}(\theta))^{y_{iqs}}\right)\right], \tag{12.5}$$

where θ is the vector of distributed parameters.

By assuming that all decisions makers are independent, and since $E[AB] = E[A]\cdot E[B]$, equation (12.5) can be rewritten as:

$$\begin{aligned} LL(\theta) &= \log\left[\prod_{q=1}^{Q} E\left(\prod_{i=1}^{I}\prod_{s=1}^{S}(P_{iqs}(\theta))^{y_{iqs}}\right)\right] \\ &= \sum_{q=1}^{Q}\log\left[E\left(\prod_{i=1}^{I}\prod_{s=1}^{S}(P_{iqs}(\theta))^{y_{iqs}}\right)\right] \\ &= \sum_{q=1}^{Q}\log\int_{\beta}\prod_{i=1}^{I}\prod_{s=1}^{S}(P_{iqs}(\beta))^{y_{iqs}} f(\beta\mid\theta)\,d\beta, \end{aligned} \tag{12.6}$$

where f is a multivariate distribution function over vector parameters β with distributional parameters θ to be estimated. Notice that in choice experiments, a single respondent faces multiple choice situations and therefore these observations are dependent (Revelt & Train 1998). Models taking into account this dependency are called 'panel mixed logit models' and equation (12.6) defines the maximum log-likelihood function of such models.

However, some estimation models (e.g. cross-sectional ML models) assume that choice observations from a single respondent over a series of choice situations are independent from each other. In that case, the maximisation of the log-likelihood can be written as:

$$\begin{aligned} LL(\theta) &= \sum_{q=1}^{Q}\log\left(\prod_{i=1}^{I}\prod_{s=1}^{S} E\left((P_{iqs}(\theta))^{y_{iqs}}\right)\right) \\ &= \sum_{i=1}^{I}\sum_{q=1}^{Q}\sum_{s=1}^{S} y_{iqs}\log E\left(P_{iqs}(\theta)\right) \\ &= \sum_{i=1}^{I}\sum_{q=1}^{Q}\sum_{s=1}^{S} y_{iqs}\log\int_{\beta} P_{iqs}(\beta) f(\beta\mid\theta)d\beta. \end{aligned} \tag{12.7}$$

Clearly, if parameters are assumed fixed instead of random, this likelihood function simplies to the well-known function for the MNL model:

$$LL(\beta) = \sum_{i=1}^{I} \sum_{q=1}^{Q} \sum_{s=1}^{S} y_{iqs} \log\left(P_{iqs}(\beta)\right). \quad (12.8)$$

References

Adler, N. and B. Golany (2001) Evaluation of deregulated airline networks using data envelopment analysis combined with principal component analysis with an application to Western Europe, *European Journal of Operational Research* **132** (2) 260-273.

Air Traffic Control the Netherlands - LVNL (2008) URL: http://www.lvnl.nl/html/en/aboutlvnl/responsibility/index.html, Accessed on 27 February, 2008.

AIS (2008) *Airspace structure and ATS airspace classification*, URL: http://www.ais-netherlands.nl/aim/080228-080410/eAIP/html/graphics/eAIP/ENR-6-2-1-A3p.pdf, Accessed on 29 February.

Arentze, T., A. Borgers, H. Timmermans and R. DelMistro (2003) Transport stated choice responses: effects of task complexity, presentation format and literacy, *Transportation Research* **39** 229-244.

Ashford, N. and M. Bencheman (1987) Passengers' choice of airport: an application of the Multinomial Logit model, *Transportation Research Records* **1147** 1-5.

Barrett, S. D. (2000) Airport competition in the deregulated European aviation market, *Journal of Air Transport Management* **6** (1) 13-27.

Basar, G. and C. R. Bhat (2004) A parametrized considersation set model for airport choice: an application to the San Fransisco Bay area, *Transportation Research* **38B** 889-904.

Ben-Akiva, M. and M. Bierlaire (1999) Discrete choice methods and their application to short term travel decisions, *Handbook of Transportation Science*, 5-35, Kluwer, Dordrecht.

Ben-Akiva, M. and S. R. Lerman (1985) *Discrete Choice Analysis: Theory and Application to Travel Demand*, MIT Press, Cambridge, Massachusetts.

Benckendorff, P. (2006) An exploratory analysis of traveler preferences for airline website content, *Information Technology & Tourism* **8** 149-159.

Bliemer, M. C. J. and J. M. Rose (2005) Efficiency and Sample size Requirements for Stated Choice Studies, '*Working paper*', Institute of Transportation and Logisitcs Studies, Sydney.

Bliemer, M. C. J. and J. M. Rose (2006) *Designing stated choice experiments: state-of-the-art*, Paper presented at the11th International Conference on Travel Behaviour Research, Kyoto, Japan.

Bliemer, M. C. J. and J. M. Rose (2008) *Construction of experimental designs for mixed logit models allowing for correlation across choice observations*, Proceedings of the 88th Transport and Research Board Annual Meeting, Washington DC.

Bliemer, M. C. J., J. M. Rose and S. Hess (2007) *Approximation of Bayesian Efficiency in Experimental Choice Designs*, Proceedings of the 87th Transport and Research Board Annual Meeting, Washington DC, USA.

Bondzio, L. (1996) Modelle für den Zugang von Passagieren zu Flughäfen, *PhD Thesis*, Ruhr-Universität Bochum, Bochum.

Bovy, P. H. L. (2008) On modeling route choice sets in transportation networks: a synthesis, *Transportation Reviews* **28**.

Bovy, P. H. L., M. C. J. Bliemer and R. van Nes (2006) Transportation Modelling, *Reader*, TU Delft, Delft.

Bovy, P. H. L. and E. Stern (1990) *Route choice and way finding in transport networks*, Kluwer Academic Publishers, Dordrecht.

Bunch, D. S., J. J. Louviere and D. Anderson (1994) A Comparison of Experimental design Strategies for Multinomial Logit models:The Case of Generic Attributes '*Working paper*', Graduate School of Management, California at Davis.

Button, K. (1991) *Airline deregulation: international experiences*, Fulton, London.

Caussade, S., J. d. D. Ortuzar, L. I. Rizzi and D. A. Hensher (2005) Assessing the influence of design dimensions on stated choice experiment estimates, *Transportation Research Part B: Methodological* **39** (7) 621-640.

Chorus, C. G. (2007) Traveler Response to Information, *TRAIL Thesis Series*, TRAIL Research School, Delft.

De Neufville, R. (1995) Management of multi-airport systems, *Journal of Air Transport Management* **2** 99-110.

De Neufville, R. (2000) Planning multi-airport systems in metropolitan regions in the 1990s, *Working paper*, Massachussets Institute of Technology, Cambridge USA.

De Neufville, R. and A. Odoni (2003) *Airport Systems: Planning, Design and Management*, McGraw-Hill, New York.

Dellaert, B. G. C., J. D. Brazell and J. J. Louviere (1999) The Effect of Attribute Variation on Consumer Choice Consistency, *Marketing Letters* **10** (2) 139-147.

DeShazo, J. R. and G. Fermo (2002) Designing choice sets for stated preference methods: the effects of complexity on choice consistency, *Journal of Environmental Economics and Management* **44** 123-143.

Dillman, D. A. (2000) *Mail and internet surveys: The Tailored design method*, Wiley, New York.

Dutch government (2007) Coalitieakkoord tussen de Tweede Kamerfracties van CDA, PvdA en ChristenUnie, Den Hague.

Enserink, B., J. F. M. Koppenjan and W. A. H. Thissen (2002) Analysis of Complex Environments (Analyse van complexe omgevingen), *Reader*, TU Delft, Delft.

Foerster, J. F. (1979) Mode choice decision process models: A comparison of compensatory and non-compensatory structures, *Transportation Research Part A: General* **13** (1) 17-28.

Franke, M. (2003) *Building blocks of future business models—refocusing airline strategy*, SIB conference, Bremen.

Franke, M. (2004) Competition between network carriers and low-cost carriers--retreat battle or breakthrough to a new level of efficiency?, *Journal of Air Transport Management* **10** (1) 15-21.

Garrow, L. A., S. P. Jones and R. A. Parker (2007) How much airline customers are willing to pay: An analysis of price sensitivity in online distribution channels, *Journal of Revenue and Pricing Management* **5** (4) 271-290.

Gillen, D. and W. G. Morrison (2005) Regulation, competition and network evolution in aviation, *Journal of Air Transport Management* **11** (3) 161-174.

Gosling, G. D. (2000) Aviation system performance measures for state transportation planning, *Transportation Research Record* **1703** 7-15.

Graham, A. (2001) *Managing Airports - An International Perspective*, Butterworth-Heinemann, Oxford.

Graham, B. and C. Guyer (2000) The role of regional airports and air services in the United Kingdom, *Journal of Transport Geography* **8** (4) 249-262.

Grammig, J., R. Hujer and M. Scheidler (2005) Discrete choice modelling in airline network management, *Journal of Applied Econometrics* **20** (4) 467-486.

Hansson, T., J. Ringbeck and M. Franke (2003) Flight for survival: a new business model for the airline industry, *strategy+business* 78-85.

Harvey, G. (1987) Airport choice in a multiple airport region, *Transportation Research* **21A** 439-449.

Heerkens Thijssen, F. B. L. (2007) A multi airport system in the Netherlands? Developing a framework for analysing the consequences of diverting air traffic to secondary airports in the Netherlands, *'Working paper'*, Delft University of Technology, Delft.

Heiner, R. A. (1983) The origin of predictable behavior, *American Economic Review* **73** 560-595.

Hensher, D. A. (2003) Revealing differences in willingness to pay due to the dimensionality of stated choice designs: an initial assessment, *Working paper*, Institue of Transport Studies, Sydney.

Hensher, D. A. (2004a) How do respondents handle stated choice experiments? - Information processing strategies under varying information load, *Working paper*, Institute of Transport Studies, Sydneey.

Hensher, D. A. (2004b) Identifying the influence of stated choice design dimensionality on willingness to pay for travel time savings, *Working paper*, Institute of Transport Studies, Sydney.

Hensher, D. A. and M. Bradley (1993) Using stated response choice data to enrich revealed preference discrete choice models, *Marketing Letters* **4** (2) 139-151.

Hensher, D. A. and J. J. Louviere (1982) On the design and analysis of simulated choice or allocation experiments in travel choice consistency, *Transportation Research Record* **890** 11-17.

Hensher, D. A., J. M. Rose and W. H. Greene (2005) *Applied choice analysis. A primer*, Cambridge University Press, Cambridge.

Hess, S. (2007) Posterior analysis of random taste coefficients in air travel behavior modelling, *Journal of Air Transport Management* **13** 203-212.

Hess, S. and J. W. Polak (2005) Mixed logit modelling of airport choice in multi-airport regions, *Journal of Air Transport Management* **11** (2) 59-68.

Hoogendoorn-Lanser, S. (2005) *Modelling travel behavior in multi-modal networks*, TRAIL Research School, Delft, The Netherlands.

Huber, J. and K. Zwerina (1996) The Importance of Utility Balance in Efficient Choice Designs, *Journal of Marketing Research* **33** (3) 307-317.

KLM (2008) URL: http://www.klm.com/travel/corporate_en/mission_people/mission/index.htm, Accessed on 3 March, 2008.

Loo, B. P. Y. (2007) Passengers' airport choice within multi-airport regions (MARs): some insights from a stated preference survey at Hong Kong International Airport, *Jounal of Transport Geography.*

Louviere, J. J., D. A. Hensher and J. Swait, D. (2000) *Stated Choice Methods: Analysis and Application*, Cambridge University Press, Cambridge, United Kingdom.

Louviere, J. J., H. Oppewal, H. Timmermans and T. Thomas (1997) Handling large numbers of attributes in conjoint applications: Who says existing techniques can't be applied? But if you want an alternative, how about hierarchical choice experiments?, *Mimeograph.*

Manski, C. F. (1977) The structure of random utility models, *Theory and Decision* **8** (3) 229-541.

Mason, K. J. (2000) The propensity of business travellers to use low cost airlines, *Journal of Transport Geography* **8** (2) 107-119.

McFadden, D. (1974) Conditional choice analysis of qualitative choice behavior, *Frontiers in Econometrics*, Edited by P. Zarembka, 105-142, Academic Press, New York.

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

Ministry of Transport (2006) URL: http://www.verkeerenwaterstaat.nl/onderwerpen/luchtvaart/regionale_luchthavens/index.asp x, Accessed on 25 February, 2008.

Park, J.-W. (2007) Passenger perception of service quality: Korean and Australian case studies, *Journal of Air Transport Management* **13** 238-242.

Pels, E., P. Nijkamp and P. Rietveld (2001) Airport and Airline Choice in a Multiple Airport Region: An Empirical Analysis for the San Francisco Bay Area, *Regional Studies* **35** (1) 1 - 9.

Pels, E., P. Nijkamp and P. Rietveld (2003) Access to and competition between airports: a case study for the San Francisco Bay area, *Transportation Research Part A: Policy and Practice* **37** (1) 71-83.

Polydoropoulou, A. and M. Ben-Akiva (2001) Combined Revealed and Stated Preference Nested Logit Access and Mode Choice Model for Multiple Mass Transit Technologies, *Transportation Research Record* **1771** 38-45.

Proussaloglou, K. and F. S. Koppelman (1999) The choice of air carrier, flight, and fare class, *Journal of Air Transport Management* **5** (4) 193-201.

Revelt, D. and K. Train (1998) Mixed Logit with Repeated Choices: Households' Choices of Appliance Efficiency Level, *Review of Economics and Statistics* **80** (4) 647-657.

Rose, J. M. and M. C. J. Bliemer (2006) *Designing efficient data for stated choice experiments.*, Paper presented at the 11th International Conference on Travel Behaviour Research Kyoto, Japan.

Rose, J. M. and D. A. Hensher (2005) Handling individual specific availability of alternatives in stated choice experiments, *Working paper*, Sydney.

Rose, J. M., D. A. Hensher, S. Caussade, J. de Dios Ortuzar and R. C. Jou (2007) *Identifying differences in preferences due to dimensionality in stated choice experiments: a cross cultural analysis*, European Transport Conference, Leeuwenhorst.

Sándor, Z. and M. Wedel (2001) Designing Conjoint Choice Experiments Using Managers' Prior Beliefs, *Journal of Marketing Research* **38** (4) 430-444.

Schaafsma, A. A. M., P. H. L. Bovy and A. J. van Binsbergen (2001) Het lagenmodel voor verkeer en vervoer: een model voor een functionele systeembeschrijving, *TRAIL Research School*, Delft.

Schiphol (2008) URL: http://www.schipholgroup.com/, Accessed on 26 February, 2008.

Schoemaker, T. J. H., K. Koolstra and P. H. L. Bovy (1999) *Traffic in the 21st century - a scenario analysis for the traffic market in 2030*, Delft.

Theis, G., T. Adler, J.-P. Clarke and M. Ben-Akiva (2006) Risk Aversion to Short Connections in Airline Itinerary Choice, *Transportation Research Record* **1951** 28-36.

Train, K. E. (2003) *Discrete Choice Models with Simulation*, Cambridge University Press, Cambridge.

Transavia (2007) URL: http://en.transavia.com/tra/images.nsf/files/annualreport_transavia.com_0607.pdf/$FILE/annualreport_transavia.com_0607.pdf, Accessed on 4 March, 2008.

Tron, A., L. McCoomb, M. Nowicki and P. Kowal (2007) *Modeling Air Traveler Choice Behavior in the Southern Ontario Market*, 86th Annual Meeting of the Transportation Research Board, Washington D.C.

Van De Riet, O. A. W. T. (2003) *Policy Analysis in Multi-Actor Policy Settings*, Eburon, Delft.

Van De Riet, O. A. W. T. and A. Turk (2006) When does infrastructure function well A multi-actor performance criteria typology, *International Journal of Technology, Policy and Management* **6** 154-167.

Van Eggermond, M. A. B. (2007) Consumer choice behaviour and strategies of air transportation service providers, Delft University of Technology, Delft.

Van Zuylen, H. J. (2005) *Lecture Notes CT5804, ITS and Dynamic Traffic Management*, TU Delft, Delft.

Veldhuis, J., I. Essers, D. Bakker, N. Cohn and E. Kroes (1999) The integrated airport competition model, 1998, *journal of Air Transportation World Wide* **4**.

Walker, W. E., S. A. Rahman and J. Cave (2001) Adaptive policies, policy analysis, and policy-making, *European Journal of Operational Research* **128** (2) 282-289.