Master Thesis

The influence of airport terminal service quality on passengers' airport choice behaviour

Gerben de Jong

Date: 2014-02-10

VU University Amsterdam
Faculty of Economics and Business Administration
MSc Business Administration – Transport and Supply Chain Management

Author:	Gerben de Jong
Student nr.:	2500306
E-mail:	jong.gerben.de@gmail.com
1st Supervisor:	Dr. C.L. Behrens
2nd Supervisor:	Dr. T. de Graaff

"The airport runway is the most important main street in any town."
(Norman J. Crabtree, former aviation director for the state of Ohio, 1925 – 2006)

Abstract

This thesis explores the relationship between (perceived) airport terminal service quality and passengers' choice for a departure airport.

While substantial research efforts have been dedicated at understanding passenger airport choice behaviour, the influence of qualitative airport terminal service dimensions received scant attention so far. Given increased airport choice and –competition, insights on this matter are highly relevant for today's air transport market.

The current study contributes to extant literature by combining insights from studies on passengers' airport choice behaviour and studies on airport terminal service quality. The analysis involves the estimation of a discrete choice model of airport choice in relationship to varying levels of ticket price access time and various airport terminal service quality attributes (availability of restaurants and shops, crowdedness, cleanliness and friendliness of staff), making use of stated preference data obtained from a sample of Dutch leisure passengers.

Assuming all attributes not included in the model are equal, the results show that ticket price and access time are the attributes with the highest influence on the departure airport choices of the leisure passengers. Interestingly, it was not possible to retrieve meaningful effects for differences in availability of restaurants and shops and airport terminal crowdedness. In contrast, varying levels of airport terminal cleanliness and friendliness of staff substantially influenced the leisure passengers' departure airport choices.

These findings reconfirm that today's leisure passengers are highly price-sensitive and highlights the great importance of access times in airport choice behaviour. Leisure passengers are generally unwilling to pay higher ticket prices or travel further to access airports that offer more opportunity to engage in shopping and eating out as well as a less congested airport terminal, suggesting that they are primarily in search of a no-frills airport service. Nevertheless, the results also indicate a great concern for airport terminal service quality dimensions such as airport terminal cleanliness and friendliness of staff, presumably because passengers perceive these airport terminal service quality dimensions as elementary facets of airport services.

Table of contents

1. Introduction ... 6
2. Passengers' airport choice behaviour ... 9
   2.1. Theory of airport choice behaviour ... 9
   2.2. Prior research on airport choice behaviour ... 10
   2.3. Conclusion ... 13
3. Airport terminal service quality ... 14
   3.1. Airport terminal service quality assessments ... 14
   3.2. Prior research on airport terminal service quality ... 15
   3.3. Limitations of prior research on airport terminal service quality ... 18
   3.4. Conclusion ... 18
4. Stated choice experiment design ... 20
   4.1. Stated choice experiments ... 20
   4.2. Design process ... 21
   4.3. Conclusion ... 30
5. Empirical analysis and results ... 32
   5.1. Data ... 32
   5.2. Methodology ... 33
   5.3. Model specifications ... 35
   5.4. Results ... 39
   5.5. Conclusion ... 44
6. Conclusion ... 46
   6.1. Findings ... 46
   6.2. Main implications ... 47
   6.3. Limitations and direction for further research ... 48

References ... 51

Appendices ... 55

List of figures

Figure 1.1 Thesis structure 8
Figure 2.1 Possible decision hierarchy for air travel (inspired by fig. 1 in Harvey, 1987) 9
Figure 4.1 Example of a choice situation as presented in the survey (in Dutch) 29
Figure G.1 Histogram for age of the respondents 62
Figure G.2 Histogram for gender of the respondents 62
Figure G.2 Histogram for air travel frequency of the respondents 62
Figure G.4 Histogram for education level of the respondents 62

List of tables

Table 2.1 Selection of airport choice studies 10
Table 3.1 Selection of airport terminal service quality studies 15
Table 4.1 List of attributes 22
Table 4.2 Attribute level labels 23
Table 4.3 Attribute coding 24
Table 4.4 Design matrix of the experimental design 26
Table 4.5 Correlation matrix of the experimental design 26
Table 4.6 Difference in attribute levels within choice sets (1=yes; 0=no) 28
Table 4.7 Variety in differences of fare- and access time attributes 28
Table 5.1 Correlation matrix for socio-economic characteristics and version number 33
Table 5.2 Effects coding structure for education level 38
Table 5.3 Estimation results for binary logit (BL) models 39
Table 5.4 Willingness-to-pay and willingness to accept increases in access time 41
Table B.1 Car access times of Dutch airports (in minutes) 57
Table B.2 Air ticket prices from Dutch airports to five major European capitals 58
Table D.1 Correlation matrix of version A 59
Table D.2 Correlation matrix of version B 59
Table D.1 Final experimental design 60
Table F.1 Description of the variables in the dataset 61
Table G.1 Descriptive statistics of the socio-economic characteristics of the sample 62
Table J.1 Complete estimation results for base model 65
Table J.2 Complete estimation results for full model 65
Table J.3 Complete estimation results for final model 66

"The standard notion places the ownership of passengers with the airline. I challenge this idea and argue that the future may be one in which the airport owns the passenger, and the airline simply serves as the airport distribution system."

–David Gillen, footnote in a paper on the evolution of airport ownership and governance (Gillen, 2011, p. 1)

1. Introduction

After the deregulation of the air transport industry, airports transformed from publicly owned properties to customer-oriented service providers (Fodness & Murray, 2007; Seneviratne & Martel, 1991). Airport managers started initiating efforts into protecting and enlarging their passenger base by increasing the customer service provided (Lee-Mortimer, 1993). Nowadays airports invest a lot of money to improve their service quality, so that passengers will select the airport in question as a departure- or transfer airport (Heathrow, 2013; Schiphol, 2013). But does the service quality of the airport influences the passengers' decisions for a departure airport?

Airport service quality can be assessed from various perspectives, such as airlines, passengers, employees and tenants (Lemer, 1992). This thesis takes the perspective of the passengers, who are ultimately the end customers of the airport and the main source of airport revenues (Seneviratne & Martel, 1991; Zidarova & Zografos, 2011). Passengers' perceptions of airport service quality depend on a combination of factors, among others the flight services provided by the airlines operating from the airport (*e.g.* airlines, routes, schedules), airport terminal service quality and airport accessibility. The focus of this thesis is on airport terminal service quality. Airport terminal service quality comprises many dimensions, including the friendliness of staff, processing time, availability of concessions and amenities, cleanliness of the airport terminal building and many others. As air passengers commonly spent an ample amount of time in the airport terminal, the service quality offered may substantially influence their perceptions of the airport's overall quality.

Nonetheless, we lack a clear picture of how passengers' choice of a departure airport is affected by qualitative airport terminal service dimensions. Therefore, this thesis explores the relationship between passengers' perceptions of airport terminal service quality and their departure airport choice. From a theoretical perspective, we contribute to extant literature by connecting research on airport choice behaviour and research on airport terminal service quality. These research streams are typically not joined together. From a practical perspective, our analysis provides airport management and policy makers with an increased understanding of the role qualitative airport terminal service dimensions play in their customers' departure airport choice decision-making processes.

This thesis distinguishes two main relevant research streams. On the one hand, the research that addresses airport choice behaviour, which investigates the determinants influencing a passenger's airport choice (*e.g.* Başar & Bhat, 2004; Harvey, 1987; Hess, Ryley, Davison, & Adler, 2013; Pels, Nijkamp, & Rietveld, 2001). On the other hand, the research that assesses airport terminal performance, evaluating how passengers value and perceive the offered service quality (*e.g.* Fodness & Murray, 2007; Seneviratne & Martel, 1991; Yeh & Kuo, 2003). Research on passengers' airport choice behaviour typically does not take into account qualitative airport terminal service dimensions and therefore the relative influence of such factors on airport choice decisions remains unknown. On the contrary, research on airport terminal service quality succeeds in identifying passengers' needs and desires regarding airport terminal service quality, but does not link the passengers' beliefs to their choice for a departure airport. Hence, similar to airport choice studies, this field of research falls short in deriving the relative importance of these in passengers' airport choice decisions. The above suggests that the limitations of one research stream practically mirror the limitations of the other research stream. The research on this subject is clearly fragmented and hence the theoretical goal of this thesis is to combine insights from the two research streams, thereby addressing both their limitations.

Given increased airport competition, this topic has high practical relevance for today´s air transport markets. Over the past decades, new airports have developed to accommodate growing demand for air travel (Bonnefoy, de Neufville, & Hansman, 2010). Moreover, the emergence of low-cost carriers resulted in the rise of secondary regional airports (Francis, Fidato, & Humphreys, 2003). As a consequence, regions in which one airport once served the whole market as a take-it-or-leave-it proposal, transitioned in multi-airport regions offering passengers a choice among airports (Bonnefoy et al., 2010). Simultaneously, passengers begun to increasingly orientate themselves in search for competitive advantage in the airport industry (Fodness & Murray, 2007).

Taking the Dutch context as an example, we observe that Schiphol used to dominate the air transport markets in the past. However, in today's short-haul market, Dutch passengers are regularly offered the choice between flights departing from multiple Dutch domestic airports and, in addition, are willing to travel further to access airports across the borders. The increased willingness of Dutch passengers to consider alternative airports is illustrated by Schiphol losing market share, especially on short-haul destinations that are also offered by other (regional) airports (Lieshout, 2012). This is further highlighted by CBS (2013) figures indicating that the percentage of intra-European passengers departing from Schiphol relative to other Dutch domestic airports decreased from 95% to 88% between 2003 and 2012. Both sources suggest that airport choice is an increasingly relevant topic for the Dutch short-haul air transport market[1].

In response to these developments, airports intensified their attempts to influence the passengers' choice for a departure airport, by differentiating themselves from the competition. Billion dollar investments are announced by airports to offer passengers high levels of service quality in airport

[1] In this context it should be noted that, departure airport choice is less relevant for the medium- and long-haul markets (at least in the Netherlands). The main airport, Schiphol, still dominates these markets. In 2012, 99,8% of the inter-continental passengers departed from Schiphol (CBS, 2013).

terminals (see for example Heathrow, 2013; Schiphol, 2013). Given these developments, one would expect a large body of literature dedicated on investigating the relationship between offered airport terminal service quality and airport choices of passengers. However, as argued above, there is remarkably little research that addresses this issue. This leads to the practical goal of this thesis, wherein we aim to provide airport management and policy makers with an increased understanding of the role qualitative airport terminal service dimensions play in their customers' departure airport choice decision-making processes.

Figure 1.1 summarises the structure of this thesis. Chapters 2 and 3 review the two relevant fields of literature. The field that addresses departure airport choice behaviour is presented in Chapter 2. The field that evaluates airport terminal service quality is subject of Chapter 3. In particular, we assess how previous studies within these fields have investigated the influence of airport terminal service quality on departure airport choice and further highlight the research gap we seek to address in this thesis. Chapter 4 and 5 represent the two-stage empirical part of this thesis. In Chapter 4 we design a stated choice experiment to generate stated preference data on departure airport choice in relation to varying levels of airport terminal service quality. In Chapter 5 we use the generated data to estimate discrete choice models of airport choice (binary logit models) and derive willingness-to-pay patterns for a number of airport terminal service quality dimensions. Chapter 6 presents the main conclusions, discusses the implications and suggests further research directions.

Figure 1.1 Thesis structure

Chapter 1: Introduction
(Global introduction of the research area, research goals, theoretical- and practical relevance and thesis structure)

Chapter 2: Passengers' airport choice behaviour
(Current understanding of passengers' departure airport choice behaviour)

Chapter 3: Airport terminal service quality
(Current understanding of airport terminal service quality from a passengers' perspective)

Chapter 4: Stated choice experiment design
(Design of a SP experiment to generate data on departure airport choice in relation to varying levels of airport terminal service quality)

Chapter 5: Empirical analysis and results
(Estimation of discrete choice models to derive willingness-to-pay patterns for airport terminal service quality dimensions)

Chapter 6: Conclusion
(Findings, conclusions, implications and future research directions)

"For once you have tasted flight you will walk the earth with your eyes turned skywards, for there you have been and there you will long to return."
– Leonardo da Vinci, one of the earlier pioneers of flight, 1452 – 1519

2. Passengers' airport choice behaviour

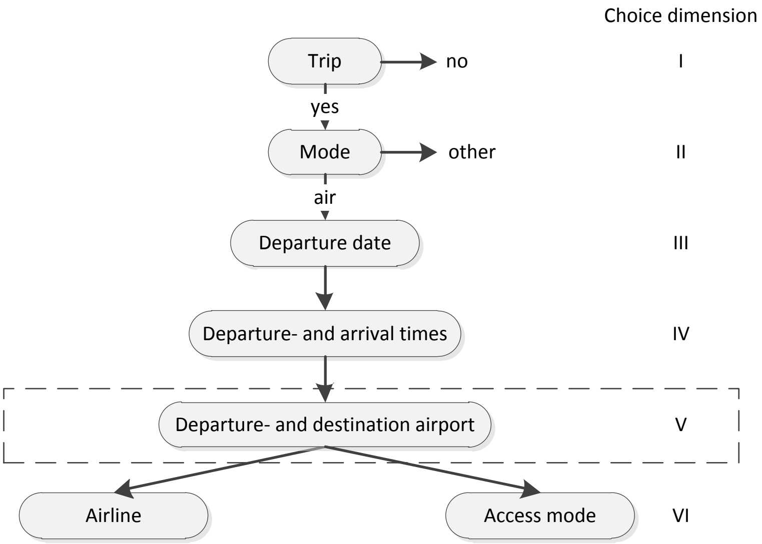
Here we study the state of the art in research on passengers' departure airport choice behaviour. Understanding of passengers' airport choice behaviour is of fundamental importance to airport management, airline carriers and policy makers. Hence a lot of intellectual resources have been devoted to the topic. In particular, we assess how prior research investigated the influence of airport terminal service quality dimensions on departure airport choice.

This chapter is structured as follows. Section 2.1 introduces the underlying theory of airport choice behaviour. Section 2.2 reviews previous research efforts, with particular attention on how prior research addressed the influence of qualitative airport terminal service dimensions on airport choice. Section 2.3 concludes and outlines our aimed contribution to this field.

2.1. Theory of airport choice behaviour

Prior to an air travel trip passengers engage in a process of decisions making. This process includes several choice dimensions, such as whether or not to make a trip, whether to travel by air or by another mode, departure date, desired arrival- and departure times, departure- and destination airports, airline and mode of access (Harvey, 1987). To model this complex decision making process, researchers often assume a simplified decision hierarchy, whereby passengers sequentially decide on these choice dimensions (Başar & Bhat, 2004). Figure 2.1 presents an illustration of a possible decision hierarchy for air travel.

Figure 2.1 Possible decision hierarchy for air travel (inspired by fig. 1 in Harvey, 1987)



The main goal of airport choice behaviour modelling is to assess the relative influence of determinants (*e.g.* price, service factors) on the airport choice of the passengers. Most of the studies focus on departure airport choice (little is known about transfer- and destination airport choice) and the departure airport choice dimension is often considered in isolation, under the assumption that any linkages with other choice dimensions can be neglected. Some studies consider multiple choice dimensions jointly, such as airport and airline (Pels et al., 2001), airport and access mode (Pels, Nijkamp, & Rietveld, 2003) or airport, airline and access mode (Hess & Polak, 2006b; Hess et al., 2013).

2.2. Prior research on airport choice behaviour

Table 2.1 summarises several previous research efforts in this field. There exists a tremendous body of literature on airport choice behaviour. The selected studies are assumed to reasonably represent the full range of the literature, but do not constitute all the available literature on this subject. For other reviews of the airport choice literature please refer to Hess and Polak (2006a) and De Luca (2012). Also, note that this chapter presents a concise overview of the existing research on airport choice. Hence we describe the findings of the studies in general ways only. We discuss some of the results of these studies in more detail in chapter 5, where we compare our findings to previous findings in related studies.

Table 2.1 Selection of airport choice studies

Author(s)	Data	Model	Location	Dimensions considered
Harvey (1987)	RP	MNL	San Francisco Bay Area	Departure airport
Pels et al. (2001)	RP	NL	San Francisco Bay Area	Departure airport and airline
Pels et al. (2003)	RP	NL	San Francisco Bay Area	Departure airport and access mode
Başar and Bhat (2004)	RP	PCMNL	San Francisco Bay Area	Departure airport
Hess and Polak (2005)	RP	ML	San Francisco Bay Area	Departure airport
Adler, Falzarano, and Spitz (2005)	SP	ML	US Domestic market	Departure airport and airline
Hess and Polak (2006a)	RP	NL	San Francisco Bay Area	Departure airport, airline and access mode
Hess and Polak (2006b)	RP	CNL	Greater London Area	Departure airport, airline and access mode
Hess, Adler, and Polak (2007)	SP	MNL	US Domestic market	Departure airport and airline
Loo (2008)	SP	MNL	Hong Kong-Pearl River Delta	Departure airport
Hess (2010)	SP	MNL	US Domestic market	Departure airport and airline
Marcucci and Gatta (2011)	SP	MNL, ML	Marche and Emilia-Romagna region (Italy)	Departure airport
Luken and Garrow (2011)	RP	MNL	New York Metropolitan Area	Departure airport
De Luca (2012)	SP	MNL, HL, CNL and ML	Campania, Southern Italy	Departure airport
Hess et al. (2013)	SP	MNL,NL and CNL	East Coast of the United States	Departure airport, airline and access mode

The majority of the earlier airport choice studies use revealed preference (RP) data, while the more recent publications use stated preference (SP) data. RP data consists of *actual* choices made in real world markets, while SP data consists of *stated* choices in hypothetical markets (*i.e.* experiments) (Hensher, Rose, & Greene, 2005). The main advantage of SP data is full information and control on the choices that respondents are faced with, which helps avoiding issues with multicollinearity, small data variation for factors of interest and uncertainty with regard to flight availability (Hess et al., 2007; Loo, 2008; Marcucci & Gatta, 2011).

The field advanced from simple multinomial logit models (MNL) (Harvey, 1987), to more advanced nested logit (NL) (Hess & Polak, 2006a; Pels et al., 2001, 2003), probabilistic choice set multinomial logit (PCMNL) (Başar & Bhat, 2004), mixed logit (Hess & Polak, 2005) and cross-nested logit (CNL) (Hess & Polak, 2006b). More complex models offer advantages in recognising that choice sets differ among respondents (PCMNL), allowing random taste variation (ML) or taking into account the relationship between multiple choice dimensions (NL and CNL). More advanced models often statistically outperform the basic MNL model, however some researchers concede that the MNL is still a very effective modelling solution, due to its ease of use and straightforward interpretation (De Luca, 2012). Basically all the studies start with a MNL model before moving to more advanced models.

Regarding the determinants of departure airport choice, basically all revealed preference studies find significant effects for *airport access time* and *flight frequency* across all population groups considered (Başar & Bhat, 2004; Harvey, 1987; Hess & Polak, 2005, 2006a, 2006b; Pels et al., 2001, 2003). Additionally, some studies find significant effects for *access costs* (Hess & Polak, 2006a, 2006b; Pels et al., 2003), *flight time* (Hess & Polak, 2006a, 2006b), *non-stop flight services* (Luken & Garrow, 2011), *aircraft type, waiting for or walking time to access modes* and *inertia variables* (Hess & Polak, 2006a). Surprisingly, none of the revealed preference studies is able to capture a significant effect for air fare. Most studies ascribe this problem to the lack of detailed fare data, pointing out that this does not necessarily mean that air fares do not significantly affect airport choices (Hess & Polak, 2005, 2006a; Pels et al., 2003).

Stated preference studies typically find significant effects for *air fare, access time, frequency or schedule delay, flight time, aircraft type, inertia effects, on-time performance* and *number of connections* (Adler et al., 2005; De Luca, 2012; Hess, 2010; Hess et al., 2007; Hess et al., 2013; Loo, 2008; Marcucci & Gatta, 2011). Additionally, positive effects are found for the *number of airlines operating from an airport* (Loo, 2008), *large free parking areas* and *the presence of low-cost carriers* (Marcucci & Gatta, 2011), *larger airports* (Hess, 2010) and *airport dwelling time* (De Luca, 2012).

As to the influence of socio-economic and trip characteristics on airport choice, several commonly investigated socio-economic characteristics of passengers include *age, gender, income, number of times flown* and *memberships of frequent flier programmes*. Examples of trip characteristics include *trip purpose, trip duration, length of haul* and *days of booking in advance*. However the reviewed studies show contradictory results regarding the effects of socio-economic and trip characteristics. Put simply, some studies find signification effects of these variables (e.g. Başar & Bhat, 2004; De Luca, 2012; Loo, 2008), while others do not or barely (*e.g.* Hess et al., 2013; Pels et al., 2001). Moreover, while the effects of socio-demographic characteristics are sometimes *statistically* significant, a solid theory that is able to explain these effects is often lacking. The above does not count for the influence of trip purpose on airport choice. The majority of studies suggest that airport choice behaviour significantly differs among passengers with different trip purposes, with a common finding being that business travellers have a higher value of time and that leisure travellers are more cost-conscious.

2.2.1. Airport terminal service quality in airport choice behaviour studies

To the best of our knowledge, the airport choice behaviour literature offers remarkably few insights on the influence of airport terminal service quality on airport choice. None of the reviewed RP studies investigated the influence of distinctive qualitative airport terminal service dimensions on airport choice. However, *nonzero statistically significant airport-specific constants* are found by many of the RP studies (Başar & Bhat, 2004; Harvey, 1987; Hess & Polak, 2005; Pels et al., 2001, 2003), indicating that other airport-specific factors, possibly the level of service quality offered in the airport terminal, may influence passengers' choice behaviour. Harvey (1987), for example, notes that one of the possible reasons for the relative size of airport specific constants in his models, is that passengers might perceive the airport terminal of one of the modelled airports as crowded and associate some kind of disutility with crowding.

Moving from RP studies to SP studies, we find that SP studies offer several additional insights regarding the influence of airport terminal service quality on airport choice. Nonetheless the insights are still not fully satisfactory. Loo (2008) represents the only of the reviewed SP studies that investigates the relative influence of two distinctive airport terminal service quality dimensions on airport choice, by including the *size of the airport shopping area* and *queue time at check-in* counters as variables in the model specification. The estimated coefficients for these two attributes are hardly significant and only for subsets of the population (*i.e.* queue time factor is only significant for males and the shopping area factor only for medium-haul passengers). Furthermore, there is no theory available in the paper to support these somewhat curious findings.

Adler et al. (2005), Hess et al. (2007) and Hess et al. (2013) investigate the influence of overall airport quality on airport choice, by including airport rankings (as specified by the respondents beforehand the SP experiment) in their model specifications. The studies factor out the access time differences among airports, so that the airport rankings can be truly regarded as representing the quality of the airports. Estimation results in these studies show substantial willingness-to-pay for flying out of the top-ranked airports, especially for business travellers. In the research of Adler et al. (2005) business travellers assign a value of $145 and non-business travellers a value of $87, for flying out of their most-preferred airport compared to their least-preferred airport. Results from Hess, Adler, and Polak (2007) indicate willingness-to-pay ranging from $83 for business travellers to $55 for non-business travellers. Lastly, the model of Hess et al. (2013) reveals a willingness-to-pay of $127 for the top-ranked airport. However, the method of using airport rankings as a proxy for all the airport service quality dimensions does not reveal any information on the relative importance of *distinctive* service quality dimensions. In other words, the answer to the question of what service quality dimensions exactly contribute to passengers' preference for one airport over another remains subject of speculation.

Lastly, Hess (2010) includes measures of airport size in his model specifications. The results show that when asked about their preference, passengers state that they dislike large airports. However, when having to choose among airports, passengers contradictorily are more likely to choose larger (in terms of air traffic share) and main airports. The author suggests that while passengers might expect higher congestion levels, they also potentially seem to expect higher levels of service at larger airports.

2.3. Conclusion

In this chapter we reviewed the literature addressing passengers' airport choice behaviour. Many research efforts have been dedicated on increasing the understanding of airport choice determinants and over the years the field has made significant progress.

Based on the reviewed studies, we conclude that a passenger's decision for a departure airport is for a large part determined by the following (groups of) factors: *Flight services offered from the airport*, among others, frequency, schedule delay, flight time, aircraft type and routing; *Airport accessibility*, which is a combination of access time, -costs and availability of transit modes; *Air fare*, defined as the amount of monetary value the passenger pays for an airport-airline combination; *Airport quality*, which nonetheless is underinvestigated by prior research.

During this literature review we specifically paid attention on how prior studies addressed the influence of airport terminal service quality dimensions on airport choice. Unfortunately the investigation of this topic is limited so far. Simply put, apart from Loo (2008), none of the reviewed studies considered the effect of distinctive airport terminal service quality dimensions. Hence the current literature on airport choice does not offer thorough insights on how the service quality offered in the airport terminal affect airport choice decisions. This further illustrates the research gap we want to address. In the next chapter we turn to the literature on airport terminal service quality to find out how this field have dealt with this topic.

"This is an extraordinary airport. . . it could be classed as one of the wonders of the modern world."

– John F. Kennedy, at the dedication ceremony of Chicago O'Hare airport, 23 March 1963

3. Airport terminal service quality

To gain insights into the influence of perceived airport terminal service quality on passengers' departure airport choices, we first need insights in what service dimensions passengers perceive as important. In this chapter we review prior research on airport terminal service quality from a passengers' perspective. Moreover, following the previous chapter, we assess what insights this field offers regarding the relationship between passengers' perceptions of airport terminal service quality and their choice for a departure airport.

This chapter is structured as follows. Section 3.1 introduces airport terminal service quality assessments. Section 3.2 discusses several academic research efforts, as well as industry practices within this field. Section 3.3 points on a number of limitations of the current literature. In section 3.4 we conclude on the content of this chapter and define a set of dimensions that comprise airport terminal service quality from a passengers' point of view.

3.1. Airport terminal service quality assessments

Starting from the early 1990s, increased customer and market pressures created urgency to investigate how passengers value and perceive quality of airport services. The general consensus was that passengers' views should be given higher priority when considering investments in airport terminals (Seneviratne & Martel, 1991) and that more research in this area was likely to yield substantial benefits to everyone who uses the airport (Lemer, 1992).

Earlier approaches typically evaluated airport terminal performance in terms of objective measures considering the flow of passengers and their luggage between the access modes and aircraft (*e.g.* space per passenger, walking distance, baggage delivery times) (Zidarova & Zografos, 2011). Acknowledgement that passengers consider a range of other more intangible factors, such as comfort, convenience and ambiance (Lemer, 1992), led to introducing the passengers' point of view and subjective measures (*e.g.* perceptions and beliefs).

At the airport terminal, passengers encounter numerous tangible and intangible service quality dimensions (Fodness & Murray, 2007). The way passengers evaluate the overall level of service quality is usually modelled as a multi-attribute value problem (see for example Kuo & Liang, 2011; Yeh & Kuo, 2003). That is, a passenger's overall impression of the service quality of an airport terminal depends on his/hers perception of the performance on various service dimensions weighted by the relative importance of the dimensions. Therefore, studies aimed at evaluating airport terminal service quality at a particular airport usually have two goals: (a) to identify relevant dimensions of airport terminal service and their relative importance as perceived by the passengers, (b) to measure the perceived performance of the concerning airport(s) on these service quality dimensions. Other studies solely focus on identifying relevant service dimensions at airport terminals, in an attempt to construct a universal model of airport terminal service quality. We briefly discuss several of these previous research efforts in the section below.

3.2. Prior research on airport terminal service quality

Table 3.1 summarises the reviewed studies and the airport terminal service dimensions they distinguish.

Table 3.1 Selection of airport terminal service quality studies

Author(s)	Input	Airport terminal service dimensions	
Seneviratne and Martel (1991)	Passengers	▪ Walking distance ▪ Information ▪ Availability of space ▪ Number of level changes ▪ Availability of seats	▪ Concessions ▪ Internal environment ▪ Waiting time ▪ Convenience
Lemer (1992)	Prior passenger surveys	▪ Compactness ▪ Delay ▪ Service reliability	▪ Service reasonableness ▪ Cost ▪ Comfort and diversion
Rhoades, Waguespack Jr, and Young (2000)	Airport operators and -consultants	▪ Restaurants and bars ▪ Rest-room facilities ▪ Retail- and duty free shops ▪ Special services ▪ Parking ▪ Car rental services	▪ Ground transportation ▪ Boarding areas ▪ Baggage claim ▪ Information display ▪ Inter-terminal transportation
Yeh and Kuo (2003) and Kuo and Liang (2011)	Airport- and travel experts, academic experts and airline executives	▪ Courtesy of staff ▪ Security ▪ Convenience ▪ Comfort	▪ Processing time ▪ Information visibility ▪ Reaction Capacity (only Kuo & Liang, 2011)
Fodness and Murray (2007)	Passengers	▪ Circulation effectiveness ▪ Movement efficiency ▪ Interaction	▪ Productivity facilities ▪ Shopping / eating ▪ Décor
ACI (2013)	-	34 key dimensions, including: ▪ Flight information screens ▪ Signposting ▪ Walking distance ▪ Restaurants and shopping ▪ Washrooms	▪ Helpfulness of staff ▪ Waiting times ▪ Cleanliness ▪ Ambience ▪ Arrival services
Skytrax (2013)	-	▪ Wayfinding ▪ Walking distance ▪ Comfort and seat availability ▪ Cleanliness ▪ WIFI and internet facilities ▪ Waiting times ▪ Friendliness of staff ▪ Flight information screens ▪ Airport shopping	▪ Availability of smoking areas ▪ Flight boarding announcements ▪ Choice of food and beverages ▪ Availability of luggage trolleys ▪ Rest areas ▪ Baggage delivery times

3.2.1. Academic research

Seneviratne and Martel (1991) were one of the first to explore which service dimensions passengers perceive as having the most significant influence on the performance of airport terminals. Taking a passenger perspective, they divided the airport terminal into three main elements: The *processing points*, which include all the primary activities encountered by the passengers (*e.g.* check-in, security). The *waiting areas*, which include all the places where passengers wait in between the primary activities (*e.g.* public waiting areas, departure lounges, restaurants and bars). And c*irculation places*, which include all the places where passengers circulate between the primary activities (*e.g.* corridors). Through a passenger survey the following dimensions were found to be the most important for the perception of the overall airport terminal service quality: *Walking distance, information provisioning, availability of space, number of level changes, availability of seats, concessions, internal environment, waiting time* and *convenience*. However, when solely circulation places are considered, passengers perceived information to be critical. While for waiting areas, availability of seats is perceived most important and at processing points, waiting time is crucial. Besides, significant differences exist among the perceptions of different passenger segments (*i.e.* age, gender and trip-purpose).

In another early approach, Lemer (1992) explored what constitutes adequate airport terminal performance as it may be viewed by the various stakeholders of the airport. Regarding the viewpoint of the passenger, he presents several, at that time actual surveys, which show that passengers have a variety of concerns, such as *compactness, delay, service reliability, service reasonableness, cost* and *comfort and diversion*. Moreover, he argues that passengers' concerns about airport terminal quality differ from one passenger segment to another. However, the author's main conclusion is that more research is needed, in order to arrive at a set of generally accepted measures that assess airport terminal performance.

Rhoades et al. (2000) took another approach and tried to identify the dimensions that comprise airport service quality based on input from industry experts and airport operators. They were asked to weigh the importance of a list of key factors of airport quality and rate the same factors from a passenger perspective. The most importance was given to parking, restrooms and baggage handling facilities. Factor analysis on the weighted factors identified four higher-level dimensions of airport terminal service quality: *Passenger services* (restaurants and bars, restroom facilities, retail- and duty free shops and special services), *airport access* (parking, car rental services and ground transportation), *airline-airport interface* (boarding areas, baggage claim and information display) and *intra-terminal transportation*.

Yeh and Kuo (2003) and Kuo and Liang (2011) both approached airport terminal service quality as a multi-attribute value problem. Yeh and Kuo (2003) consulted airport managers, governmental officials, expert academics and travel agents in Taiwan, to identify a number of distinctive airport service dimensions. Next, Taiwanese tour operators were asked to weigh the six airport service dimensions and assess the performance of 14 major Asian-Pacific international airports. According to the subjective assessment of the tour operators the six airport terminal service dimensions were ranked in the following order: *Courtesy of staff, security, convenience, comfort, processing time* and *information visibility*. Kuo and Liang (2011) called in academic experts and executives of airline industries to establish and weigh a set of airport terminal evaluation criteria. They identified the same dimensions as Yeh and Kuo (2003) and an additional dimension that they called *reaction capacity* (*i.e.* competence and reliability of staff and their response to complaints and delays). This set of evaluation criteria and respective weights were used in an empirical study for evaluating service quality of seven major Northeast-Asian international airports.

The research of Fodness and Murray (2007) was aimed to gain an understanding of the dimensionality of airport terminal service quality directly from the passengers' voice. Three different qualitative methodologies – in-depth interviews, focus groups and content analysis of submitted comments at an airport website – were used to identify 65 service quality themes that might contribute to passengers' preferences for one airport over another. Using survey data of 1.765 frequent travellers a conceptual model of airport service quality was constructed. It was concluded that passengers' expectation of airport service quality is a multidimensional, hierarchical construct that includes three key dimensions. The first dimension, *function*, includes how well passengers find their way to the different directions (effectiveness) and how much time is involved in their movements (efficiency). The second dimension, *interaction*, includes the interactions a passenger has with airport service personnel. The third dimension, *diversion*, relates to the passengers' opportunities to be engaged in job- or study-related activities at the airport (productivity), shopping and eating (maintenance) and enjoying the décor of an airport (décor).

3.2.2. Industry practices

When we move from academic efforts to industry practices, two major worldwide airport service quality studies can be identified. First, the Airport Council International Airport Service Quality programme (ACI, 2013). This programme includes both an airport service quality performance benchmark, which measures actual objective data about the levels of service delivered by an airport (*e.g.* length of queues, baggage delivery times) and a survey that measures passengers' perception of airport service quality. The latter is conducted at over 100 participating airports and covers 34 airport terminal service dimensions (a selection of these dimensions is named in table 3.1).

The second major industry practice, Skytrax (2013), is a non-profit worldwide airport survey based on measuring passengers' satisfaction on a number of airport terminal service dimensions. This online post-travel survey includes numerous airport terminal service quality dimensions (please refer to table 3.1). It is noteworthy that Skytrax (2013) also includes a dimension that relates to the WIFI and internet facilities that are offered in the airport terminal. This dimension is not included by any of the other reviewed studies.

3.3. Limitations of prior research on airport terminal service quality

The review of the literature on airport terminal service quality revealed a number of limitations.

First, the research in this field is highly ambiguous. Despite a number of attempts and calls (Fodness & Murray, 2007; Lemer, 1992; Rhoades et al., 2000), there is still no widely-accepted model for airport terminal service quality from a passengers' point of view. However, the question is whether such a model is feasible in practice. A list of all possible service quality dimensions is possibly infinite and, in the end, ambiguity is inherent to the concept of quality due to its subjective nature.

Moreover, the concept of airport terminal service quality is subject to change. Consider for example the introduction of self-service elements and internet facilities at airports. Given that those are recent developments, it is comprehensible that only Skytrax (2013) includes a dimension on WIFI and internet facilities. Nonetheless, an airport's WIFI facilities and the ease of use of self-service elements may substantially influence today's perceptions of overall airport terminal service quality and therefore further research should take these dimensions' influence into account.

Another striking limitation is that the existing research does not make any well-specified link between passengers' beliefs (relative importance and performance) about airport terminal service quality dimensions and behavioural intentions. As a consequence, the relative importance of airport terminal service quality on passengers' airport choice remains unknown, as also recognised by Fodness and Murray (2007).

Besides, the identification and weighing of the set of airport terminal service dimensions have largely relied on the subjective judgement of a limited panel of experts, or stated importance ratings obtained from passenger surveys. The first methodology derives importance ratings not directly from the passengers and thus remains an interpretation at best. The second methodology potentially suffers from customers' inability to discriminate between preferences, leading to everything being 'very important' (Garver, 2003).

Lastly, the effect of passenger characteristics on the perceived quality of service has largely been ignored in the literature. Apart from the earlier studies (Lemer, 1992; Seneviratne & Martel, 1991), the majority of the studies do not recognise differences among passengers' regarding desired quality of airport terminal service.

3.4. Conclusion

In this chapter we reviewed the field of literature that addresses airport terminal service quality. We found that the research in this field is highly ambiguous and that there is no widely-accepted consensus about which dimensions comprise airport terminal service quality from a passengers' perspective.

Based on which dimensions are mentioned most in the reviewed studies[2], we distinguish the following (groups of) dimensions: *Concessions and amenities*, such as restaurants and shops, washrooms and miscellaneous passenger services (*e.g.* money exchange, cash machines, rest areas); *Circulation*, which includes walking distance, signposting and congestion level; *Aesthetics*, which is a function of cleanliness, sound-, temperature- and humidity levels, lighting and ambiance; *Courtesy of staff*, consisting of friendliness, helpfulness, reliability and response to complaints; *Information provision*, including the flight information screen and auditory announcements; *Processing time*, which is the sum of the processing- and waiting times*; Comfort*, such as availability of seating; *Convenience*, meaning facilities or devices that facilitate processing activities; *Security*, implying the sense of security; and *WIFI and internet facilities*, which is only mentioned in one study but likely an important quality dimension for the future.

The main research gap of the state of the art in research on airport terminal service quality is that the present studies do not link passengers' beliefs of airport terminal service performance to their choice for a departure airport. Note that this limitation is in fact similar to the research gap we defined for the airport choice literature in the previous chapter. Again this leads to a lack of knowledge about the influence of airport terminal service quality on passengers' airport choice decisions.

[2] Please refer to Appendix A for a detailed overview of this assessment.

“Doubtless people are sometimes foolish, and buy things, as children do, to please a moment's fancy; but at least they think at the moment that there is a wish to be gratified.”
– Frank William Taussig, excerpt from a principles of economics textbook (Taussig, 1920, p. 116)

4. Stated choice experiment design

The previous chapters illustrated that both the field of literature on passengers’ airport choice behaviour, as well as the field of literature on airport terminal service quality offer only limited insight regarding the influence of airport terminal service quality on passengers’ airport choice decisions. We defined this as the main research gap this thesis aims to address.

As a first step in fulfilling this goal, we estimate a discrete choice model of airport choice to measure how passengers consider qualitative airport terminal service dimensions in choosing among airports. Stated choice experiments (Hensher et al., 2005) offer an appropriate tool to generate the stated preference data required for estimating such models. The current chapter’s main focus is to design this stated choice experiment.

The chapter is organised as follows. Section 4.1 introduces stated choice experiments. Section 4.2 reports on the design process, the decisions made in this process and the resulting final stated choice experiment design. Section 4.3 offers the concluding remarks on the content of this chapter.

4.1. Stated choice experiments

In a *stated choice experiment* respondents are shown a number of *choice sets*, each containing two or more hypothetical *alternatives* (*e.g.* products, goods, services, etc.). Respondents are asked to choose the alternative they like best, given a certain specified hypothetical *choice context*. The alternatives are each described by a number of *attributes* with varying *attribute levels*, along which the respondents can evaluate their choice. This generates a data set of “stated choices" made by the respondents, which can be analysed using discrete choice analysis methods to estimate the effects of the attribute levels on the choice outcome. The biggest appeal of stated choice experiments over other types of experiments (*e.g.* ranking-, rating experiments) is that respondents are asked to make a choice, which is more comparable to real life decision making than assigning ratings or rankings (Hensher et al., 2005).

Underlying each stated choice experiment is the *experimental design*. An experimental design is the specification of attributes, attribute levels, alternatives and choice sets. The experimental design is the foundation of the stated choice experiment and the way it is specified has important implications for the subsequent analysis and validity of the results. Therefore, basic scientific design methods are employed to ensure that the stated choice experiment will generate statistically powerful data that is able to reliably measure the effects of interest, while minimising the cognitive burden on respondents.

In the remainder of this chapter we generally follow the process for setting up stated choice experiments as described in Hensher et al. (2005):

- Refinement of the problem;
- Identification of relevant attributes and attribute levels;
- Generation of the experimental design;
- Construction of the survey instrument.

Much of the problem refinement has already happened in the first three chapters. Hence we skip this part of the process in this chapter (see the first paragraph of this chapter for a quick summary).

4.2. Design process

4.2.1. Identification of attributes and attribute levels

One of the first steps in setting up a stated choice experiment is defining the attributes and attribute levels to be used in the experiment. Attributes are typically identified through focus groups, interviews with experts or based on literature reviews (Hensher et al., 2005). Because of time and budget constraints inherent to master thesis research, we limit ourselves to literature review as the methodology to identify attributes.

Which *attributes* to include primarily depends on what the research hopes to achieve. We are interested in whether passengers consider levels of service quality offered in airport terminals when selecting a departure airport. In order to test this empirically, our stated choice experiment should offer respondents with a set of choices among airports that are characterised by attributes that relate to airport terminal service quality dimensions. Moreover, we are interested in whether people are willing to trade-off levels of factors that are currently believed to largely determine a passenger's airport choice (*e.g.* airport accessibility, ticket prices), for improvements on qualitative airport terminal service dimensions. Therefore we also need to include some attributes describing such primary airport choice drivers.

Recall that we have defined a set of primary airport choice drivers in Chapter 2 and a set of dimensions that comprise airport terminal service quality in Chapter 3. Ideally, we would like to include all these factors in a stated choice experiment. However, including too many attributes in stated choice experiments might make the choice tasks too complex and this may jeopardise the reliability of the choice responses (Caussade, Ortúzar, Rizzi, & Hensher, 2005; DeShazo & Fermo, 2002). Therefore we include samples of both types of factors in the experiment.

In Chapter 2 we concluded that a passenger's airport choice is for a large part determined by three main groups of determinants: *Flights services offered from the airport, airport accessibility* and *air fare* (see section 2.4). To avoid burdening the respondents with highly complex choices, it is decided not to include any attributes that relate to the flight services an airport offers. Instead respondents may assume that these are the same for all airports in the experiment. To represent the remaining two groups of airport choice determinants, we include an attribute describing the *total ticket price* and an attribute describing the *access time to the airport*.

In Chapter 3 we defined a set of dimensions that comprise airport terminal service quality from a passengers' perspective: *concessions and amenities, terminal circulation, courtesy of staff, aesthetics, information provisioning, processing time, comfort, convenience, security* and *WIFI and internet facilities* (see section 3.4). Clearly these are too many dimensions to consider including them all in the experiment. Furthermore, each dimension contains multiple variables. Courtesy of staff for example, can be described by the friendliness of the staff, by the problem-solving capabilities of the staff, the availability of staff, etc. However these variables are most likely *perceptually correlated*. That is, if a respondent sees that the friendliness of the staff is high, he often also cognitively perceives that their problem-solving capabilities are good and that staff is abundantly available throughout the airport terminal, etc. (Hensher et al., 2005).

Taken the above into account, we include one attribute for each of the following airport terminal service quality dimensions: *concessions and amenities, terminal circulation, courtesy of staff* and *aesthetics.* We choose those airport service quality dimensions, because they are the service dimensions most mentioned in the airport terminal service quality literature (see Appendix A). To represent the concessions and amenities dimension we include an attribute on the *amount of restaurants, bars and shops at the airport terminal*. For the circulation dimension an attribute on the *crowdedness of the airport terminal* is included. The aesthetics dimension is represented by an attribute on the *cleanliness of the airport terminal*. Lastly, an attribute on the *friendliness of staff* is included to represent the courtesy of staff dimension. While the selected attributes do not capture all the dimensions of airport terminal service quality, they are assumed to describe four of the most important airport terminal service quality dimensions to the majority of the passengers. Table 4.1 summarises the six attributes that will be used in the experiment.

Table 4.1 List of attributes

Attribute name	Attribute description
Fare	Total ticket price
Access Time	Access time to the airport
Concessions and amenities	Amount of restaurants, bars and shops at the airport terminal
Circulation	Crowdedness of the airport terminal
Aesthetics	Cleanliness of the airport terminal (*e.g.* corridors, waiting areas, rest rooms)
Courtesy of staff	Friendliness of staff

Having identified the attributes, the number and description of *attribute levels* must be determined. The attribute levels should be relevant, plausible and easy to understand (Hensher et al., 2005). To determine plausible levels for the *fare* and *access time* attributes, average short-haul fares and airport access times of the Dutch population are approximated (see Appendix B and C for details). These levels can be considered relevant and plausible for the target population of the experiment (*i.e.* Dutch residents). We focussed on short-haul fares, because in the Dutch context airport choice is only relevant for short-haul flights (Schiphol is the only airport offering long-haul flights). Ticket prices from an online booking system (Cheaptickets, 2013) are used to calculate average ticket prices for flights from the five major domestic Dutch airports to five European capitals. Route planning software (GoogleMaps, 2013) is used to calculate the access times (by car) from the 12 Dutch provincial capitals to the five major Dutch airports. According to these methods, the average access time for the Dutch population is established at ≈ 80 minutes and the average short-haul ticket fares at ≈ EURO 95,-. Next, the attribute levels are varied around these averages to represent the typical ranges in fares and access times. The numbers are rounded for the convenience of the respondents.

The *airport terminal service quality attributes* will each have two attribute levels, both at the two extremes of the attribute level range, called an ‘end-point design’ (Hensher et al., 2005). It appeared to be very difficult to formulate meaningful average attribute levels for qualitative attributes. We tried to use more attribute levels (*e.g.* not clean – clean – very clean), but this resulted in highly difficult choice tasks. The downside of using two attribute levels is that only linear effects can be estimated. However since this is the first study that estimates the influence of qualitative terminal service quality attributes it is deemed sufficient. It is up to further research to investigate non-linear relationships between airport terminal service quality dimensions and airport choice. The final selection of attributes and associated attribute levels is summarised in table 4.2.

Table 4.2 Attribute level labels

Attribute name	Attribute levels
Fare	80 EURO 90 EURO 100 EURO 110 EURO
Access Time	60 minutes 75 minutes 90 minutes 105 minutes
Concessions and amenities	Great availability of restaurants, bars and shops Poor availability of restaurants, bars and shops
Circulation	Uncrowded Crowded
Aesthetics	Clean Not clean
Courtesy of staff	Friendly Unfriendly

4.2.2. Generation of the experimental design

Now that the number of attributes and levels are specified, the next step involves the generation of the *experimental design*. The experimental design is a list of all the combinations of attribute levels (*i.e.* alternative airports) that will be presented to the respondents during the questionnaire.

The number of all possible different alternatives, called the *full factorial design,* is equal to L^A, where L is the number of levels; and A is the number of attributes. Thus, from the attributes and levels as determined in the last paragraph (table 4.2), a total of 256 (*i.e.* $4^2 \times 2^4$) different alternative airports can be generated. However, it is possible to work with only a sample of all possible alternatives, called a *fractional factorial design*. To construct such a fractional factorial design there exist multiple different design methods, that help selecting alternatives from the full factorial in a structured way, ensuring that statistically powerful data will be generated from the stated choice experiment (Rose & Bliemer, 2009). In this research a so-called *orthogonal main effect only design* will be used.

Orthogonal main effect only designs ensure that the main attributes are orthogonal, meaning that the correlations between attribute levels used in the choice sets are minimised. Hence the resulting data exhibits zero multicollinearity between the main attributes. Unfortunately these designs produce data that has multicollinearity between interaction effects. Therefore these designs are capable of estimating the main effects (*i.e.* effect of an attribute on choice) independently, but are not capable of reliably estimating any interaction effects (*i.e.* combined effect of two attributes on choice) (Hensher et al., 2005; Rose & Bliemer, 2009). The first step in generating an orthogonal main effect only design is to assign orthogonal codes to the attribute levels, as is shown in table 4.3.

Table 4.3 Attribute coding

Nr.	Attribute name	Attribute code	Attribute levels	Orthogonal level code
A_1	**Fare**	FAR	80 EURO	-3
			90 EURO	-1
			100 EURO	1
			110 EURO	3
A_2	**Access Time**	ACC	60 minutes	-3
			75 minutes	-1
			90 minutes	1
			105 minutes	3
A_3	**Concessions and amenities**	CON	Great availability	-1
			Poor availability	1
A_4	**Circulation**	CIR	Uncrowded	-1
			Crowded	1
A_5	**Aesthetics**	AES	Clean	-1
			Not clean	1
A_6	**Courtesy of staff**	COU	Friendly	-1
			Unfriendly	1

Each choice set in the experiment contains two hypothetical alternative airports (called a *paired comparison* design). SPSS software (IBM, 2013) is used to create two different *orthogonal main effect plans*, one to represent the first alternatives in each choice set (the A alternatives), the other to create the second alternatives (the B alternatives). An orthogonal main effect plan is a matrix, where each row represents an alternative and each column a different attribute. The columns (attributes) of an orthogonal main effect plan are orthogonal (*i.e.* uncorrelated).

The smallest possible orthogonal main effect plan for two 4-level attributes and four 2-level attributes contains sixteen alternatives. Because two orthogonal main effect plans are used the total number of alternatives in the experimental design is equal to 32, meaning that each respondent has to complete sixteen choice sets to see all the possible alternative airports. This high number of choice sets may lead to *fatigue effects*, whereby respondents become tired or bored, leading them to increasingly making errors or becoming inattentive to choice sets at the end of the experiment (Arentze, Borgers, Timmermans, & DelMistro, 2003; Caussade et al., 2005). For this reason the design is divided in two different versions, using a *blocking attribute*. Version one incorporates the choice sets 1 to 8, while the second version contains the choice sets 9 to 16. Note that this halves the number of choice sets each respondent has to complete.

Table 4.4 displays the resulting experimental design in a so-called *design matrix*. The rows of the design matrix represent the choice sets, while the columns of the design matrix represent the attributes. The attribute levels are represented using the orthogonal codes reported in table 4.3. The A alternatives represent the first alternatives in each choice set and the B alternatives represent the second alternatives in each choice set. Furthermore, note that the design is blocked in two different versions, by sorting the blocking attributes (A_7 and A_{14}).

Note from table 4.4 that the sums of the attribute columns are equal to zero, which – when using orthogonal attribute level codes – indicates that each level of each attribute appears equally often over the whole experiment. Because blocking attributes are used to block the design, not only the columns of the full design are equal to zero, also the columns of the separate blocks equal zero, indicating that each level of each attribute appears equally often in each block. This is a major advantage of using blocking attributes versus blocking the design in a random manner.

Table 4.4 Design matrix of the experimental design

Choice set #	Alternative A FAR (A_1)	ACC (A_2)	CON (A_3)	CIR (A_4)	AES (A_5)	COU (A_6)	BLO (A_7)	Alternative B FAR (A_8)	ACC (A_9)	CON (A_{10})	CIR (A_{11})	AES (A_{12})	COU (A_{13})	BLO (A_{14})
1	3	-3	1	1	1	-1	1	-3	3	-1	-1	1	1	1
2	-1	-1	1	1	-1	-1	1	1	-3	1	-1	1	-1	1
3	3	3	1	-1	-1	1	1	-1	3	-1	1	1	-1	1
4	1	1	-1	1	-1	1	1	-3	-1	1	1	-1	-1	1
5	-1	1	1	-1	1	1	1	3	1	-1	-1	-1	-1	1
6	-3	-3	-1	-1	-1	-1	1	1	1	-1	1	-1	1	1
7	1	-1	-1	-1	1	-1	1	3	-3	1	1	1	1	1
8	-3	3	-1	1	1	1	1	-1	-1	1	-1	-1	1	1
SUM	**0**	**0**	**0**	**0**	**0**	**0**	**8**	**0**	**0**	**0**	**0**	**0**	**0**	**8**
9	-3	-1	1	-1	1	1	-1	3	3	1	1	-1	-1	-1
10	1	-3	1	-1	-1	1	-1	-1	1	1	-1	1	-1	-1
11	3	1	-1	-1	1	-1	-1	1	3	1	-1	-1	1	-1
12	3	-1	-1	1	-1	1	-1	-3	1	1	1	1	1	-1
13	1	3	1	1	1	-1	-1	-3	-3	-1	-1	-1	-1	-1
14	-3	1	1	1	-1	-1	-1	3	-1	-1	-1	1	1	-1
15	-1	-3	-1	1	1	1	-1	1	-1	-1	1	1	-1	-1
16	-1	3	-1	-1	-1	-1	-1	-1	-3	-1	1	-1	1	-1
SUM	**0**	**0**	**0**	**0**	**0**	**0**	**-8**	**0**	**0**	**0**	**0**	**0**	**0**	**-8**

Table 4.5 shows the correlation matrix of the final design. According to Hensher et al. (2005), for experimental designs of unlabelled experiments it is important that attributes are orthogonal within alternatives, while across-alternative orthogonality is of little concern. Hence, attributes of alternative A do not necessarily have to be orthogonal to attributes of alternative B. Note from table 4.5 that the design is *within-alternative orthogonal*.

Table 4.5 Correlation matrix of the experimental design

	FAR (A_1)	ACC (A_2)	CON (A_3)	CIR (A_4)	AES (A_5)	COU (A_6)	FAR (A_8)	ACC (A_9)	CON (A_{10})	CIR (A_{11})	AES (A_{12})	COU (A_{13})
FAR (A_1)	1											
ACC (A_2)	0	1										
CON (A_3)	0	0	1									
CIR (A_4)	0	0	0	1								
AES (A_5)	0	0	0	0	1							
COU (A_6)	0	0	0	0	0	1						
FAR (A_8)	-0,55	-0,15	0,11	-0,45	0,22	-0,11	1					
ACC (A_9)	0,3	-0,25	0,22	-0,34	0,11	0,34	0	1				
CON (A_{10})	0,11	-0,11	-0,25	0	0	0,25	0	0	1			
CIR (A_{11})	0	-0,11	-0,5	-0,25	-0,25	0,25	0	0	0	1		
AES (A_{12})	0,34	-0,45	0,25	0,25	-0,25	0	0	0	0	0	1	
COU (A_{13})	0	0	-0,5	0	0	-0,5	0	0	0	0	0	1

Within-alternative orthogonality between, for example, the fare and access time attribute means that the lowest fare attribute level (80 EURO) does not consistently appear together with the lowest access time attribute level (80 minutes) in the same alternative. Instead the lowest fare level appears equally often together with all four of the access time levels. Thus, if a respondent would pick all the alternatives with the lowest fare level, this respondent will automatically pick all the four different access time attributes. The correlation matrix reported in table 4.5 shows that this applies to all attribute levels used in the experiment. For a more elaborate explanation of orthogonality please refer to Hensher et al. (2005).

It is important to understand that only the full design is orthogonal and that the separate blocks are not. If one would make a correlation matrix of the separate blocks, it would be noted that the blocks are not completely within-alternative orthogonal (see Appendix D). However, as a result of blocking the design using blocking variables, the within-alternative correlations between attributes in version A exactly mirror the within-alternative correlations in version B and vice versa. Merging the data from a respondent that completed version A with the data from a respondent that completed version B, neutralises the correlations and hence the resulting aggregate data exhibits no multicollinearity. In other words, two respondents complete one full experimental design. This implies that as long as the number of respondents that complete the first version of the experiment is equal to the number of respondents that complete the second version of the experiment, orthogonality in the *data* is maintained (the strategy to ensure this, is discussed in section 4.2.3).

Another important property of experimental designs from a statistical perspective, is the number of times attribute levels of alternative A differ with attribute levels of alternative B within the same choice sets. This can be refered to as the *statistical efficiency* of the design (Hensher et al., 2005). Because choices in utility theory only depend on the differences between the alternatives (Train, 2009), this statistical property relates to the ammount of information that can be obtained from the experiment. Hence a high statistical efficiency directly increases the statistical power of the choice models.

Table 4.6 shows the number of times the attribute levels of alternative A are different from those of alternative B within the same choice set. As is shown a total of 71 times the levels between alternative A and alternative B are different, which results in an efficiency of ≈ 74% ($\frac{71}{96}$). It is possible to create designs where the attributes between alternatives within the same choice set never take the same level (100% statistical efficiency), known as *optimal orthogonal designs* (Street, Burgess, & Louviere, 2005). However, the more attributes between alternatives have different levels, the harder it is for the respondents to make a choice. Thus, some researchers have argued that these optimal orthogonal designs might promote certain undesirable forms of choice behaviour (*e.g.* lexicographic) and are therefore not preferable (Rose & Bliemer, 2009).

Table 4.6 Difference in attribute levels within choice sets (1=yes; 0=no)

Choice set #	FAR	ACC	CON	CIR	AES	COU	**SUM**
1	1	1	1	1	0	1	**5**
2	1	1	0	1	1	0	**4**
3	1	0	1	1	1	1	**5**
4	1	1	1	0	0	1	**4**
5	1	0	1	0	1	1	**4**
6	1	1	0	1	0	1	**4**
7	1	1	1	1	0	1	**5**
8	1	1	1	1	1	0	**5**
9	1	1	0	1	1	1	**5**
10	1	1	0	0	1	1	**4**
11	1	1	1	0	1	1	**5**
12	1	1	1	0	1	0	**4**
13	1	1	1	1	1	0	**5**
14	1	1	1	1	1	1	**6**
15	1	1	0	0	0	1	**3**
16	0	1	0	1	0	1	**3**
SUM	**15**	**14**	**10**	**10**	**10**	**12**	**71**

The last statistical property that needs attention is the *variety in differences*. For 2-level attributes this is not important, because the difference is either 1 unit or 0 units (no difference). However, the experimental design also incorporates two 4-level attributes (fare and access time). The difference for attributes with four levels can take several forms as is shown in table 4.7.

Table 4.7 Variety in differences of fare- and access time attributes

Unit difference	FAR	ACC
0 (no difference)	1	2
1(10 euros for the fare attribute and 15 minutes for the access time attribute)	6	7
2 (20 euros for the fare attribute and 30 minutes for the access time attribute)	5	4
3 (30 euros for the fare attribute and 45 minutes for the access time attribute)	4	3

Table 4.7 reports how many times the possible differences occur. Note that the differences in the fare and access time attributes between alternatives are not orthogonal. Some differences (*e.g.* 0 unit fare difference) are clearly under-represented, while others (*e.g.* 1 unit access time difference) are over-represented. These issues may be solved by using *difference type designs* (Rose & Bliemer, 2009). However these are difficult to construct and clearly beyond the scope of this thesis, thus we assume that the current variety in the differences is sufficient.

The full experimental design, with the orthogonal codes substituted by the real attribute levels, is shown in appendix E.

4.2.3. Construction of the survey instrument

Sawtooth (2013) is used to construct an online survey instrument, including the stated choice experiment, that can be distributed to potential respondents.

First, the generated experimental design (table 4.4) is imported in the software. As stated in the previous section, each respondent completes eight choice sets during the course of the survey. Some respondents are shown the first 8 choice sets (version A), others the last 8 choice sets (version B). If we want the resulting data to be orthogonal, we need to ensure that an equal number of respondents complete the two versions. Moreover, the group of respondents that complete one version must not substantially differ from the group of respondents that complete the other version. That is, we do not want only males completing version A, with only females completing version B, as this has the potential to provoke a *blocking effect bias*. To ensure that the versions are distributed evenly and randomly, we program the software in such a way that the first respondents will be shown version A, the second respondent version B, the third respondent version A again, and so on.

To counter any *order effect bias*, the choice sets are randomised within the versions. This means that respondents completing the same version, complete the same choice sets, however, the order in which they view the choice sets is randomised (Hensher et al., 2005). In each choice set the respondents are forced to choose one of the alternatives (*i.e.* there is no “non-choice” option). This forces the respondents to make trade-offs between the attributes, which reveals information of the influence of the attribute levels upon choice. Furthermore, in order to assure that the respondents treat every choice as independent to previous decisions, the survey instrument is set up in such a way that it is not possible to return to previously completed choice sets. An example of a choice set as presented during the survey is shown in figure 4.8.

Figure 4.1 Example of a choice situation as presented in the survey (in Dutch)

Ticket prijs	**90 EUR**	**100 EUR**
Reistijd naar vliegveld	**90 minuten**	**60 minuten**
Winkels en restaurants	**Veel restaurants en winkels**	**Veel restaurants en winkels**
Drukte op het vliegveld	**Druk**	**Rustig**
Netheid van het vliegveld	**Schoon**	**Niet schoon**
Vriendelijkheid personeel	**Onvriendelijk**	**Onvriendelijk**

The survey contains one *select question*, asking the respondents to state what they consider the most important determinant(s) of their departure airport choice. The response option for this question includes a long list of possible determinants and an open-end text field in which the respondents can type responses that are not listed in the pre-specified response options (*i.e.* an “other specify” option). This question is helpful in checking whether important attributes are excluded from the choice experiment and helps generating a pool of airport choice drivers that can be used for further research.

Four questions about socio-economic characteristics of the respondents are included, to derive information about age, gender, travel frequency and education level. These questions were formulated as follows:

- Please specify your age;
- Please specify your gender;
- Please name the number of times you have flown in the past 12 months;
- Please specify the highest level of education you have completed.

A question about trip purpose, which is a common classification of passengers within the air industry, is not included. In this thesis we consider leisure trips only. This narrows the focus and avoids uncertainty at respondents whether they are making their decisions in the context of a leisure or a business trip. The *situation description*, which precedes every choice situation, specifies to the respondents that they have to imagine making this choice in the context of a leisure trip:

"Imagine a situation where you are taking a leisure trip to a European destination. You have decided to travel by air and after some research found two air travel options that are exactly similar (same airline, same departure- / arrival time, same travel time, etc.). ***The only differences between the two options are the departure airport and the ticket price.*** *Please consider the following situations and choose the airport you would take for your leisure trip."*

As the final step in constructing the survey instrument, the preliminary survey was presented to four potential respondents in *face-to-face pilot tests*. The participants could state any ambiguities and defects they encountered. This helped avoiding any misunderstandings that respondents might have during the survey. The results of the pilot tests were satisfactory; the participants easily understood how to complete the survey and it did not take them an excessive amount of time to do so. After several final refinements of the survey instrument it was ready to be distributed to potential respondents.

4.3. Conclusion

In this chapter we designed a stated choice experiment that is able to generate stated preference data on passengers' airport choices in relationship to levels of airport terminal service quality. This data can be used to measure the influence of qualitative airport terminal service dimensions on passengers' airport choice as well as estimating to which extent passengers are willing to pay or willing to accept increases in access time in exchange for higher levels of terminal service quality.

In order to achieve this, both attributes that relate to main determinants of departure airport choice (*fare* and *access time*) as well as attributes that relate to important airport terminal service quality dimensions (*availability of restaurants and shops*, *crowdedness*, *cleanliness* and *friendliness of staff*) are included in the experiment.

The underlying experimental design belongs to the family of orthogonal main effects only designs. This means that the influence of the design attributes on departure airport choice can be estimated more or less independently. However, it will not be possible to reliably estimate any interaction effects from the data that will be generated by the stated choice experiment[3].

Furthermore, the experimental design has a fairly high statistical efficiency and presents the respondents with a decent variety in choices. Both of these statistical properties are assumed to sufficiently meet the standards needed for estimating the models that we are interested in. Further research may deploy optimal orthogonal- and difference type designs to optimise both these statistical properties, but one should be aware that these typically carry other drawbacks and are highly complex to construct (Rose & Bliemer, 2009).

Throughout the whole design process, attention was paid to how the respondents would experience the experiment, in terms of complexity and length of the experiment. The result is a stated choice experiment that is able to reliably measure the effects of interest, while the cognitive burdens on respondents are minimised.

[3] Note that this refers to interaction effects for the main design attributes. Reliable estimation of interaction effects between main design attributes and socio-economic respondent characteristics are perfectly possible.

"Everything is worth what its purchaser will pay for it"
– Publilius Syrus, Latin writer, ~100BC

5. Empirical analysis and results

Here we estimate discrete choice models of choice between two alternative airports, characterised by varying levels of price, access time and various qualitative airport terminal service dimensions. The models allow testing the relative influence of these quality dimensions in passengers' airport choice decision making as well as enable the estimation of willingness-to-pay and willingness to accept increase in access time for higher levels of airport terminal service quality.

This chapter is organised as follows. Section 5.1 discusses the data used to estimate the models. Section 5.2 elaborates on the methodology of discrete choice models and random utility theory. Section 5.3 explains the model specifications. Section 5.4 analyses the model results, goodness of fit and willingness-to-pay implications. Finally, section 5.5 concludes on the content of this chapter.

5.1. Data

The data used to estimate the choice models is obtained by the online stated choice experiment discussed in the previous chapter. Respondents were selected using a convenient sampling method called *snowballing* (Hensher et al., 2005). Friends, family and colleagues were asked to complete the survey and to pass the survey on to other potential respondents that they might know. While this nonprobability sampling method places limitations on the generalisability of the results, it is a quick and convenient way to gather a large amount of data. The survey was targeted at Dutch residents of the age 18 or higher. To ensure that respondents were able to relate to the choice situations in the experiment, only those that ever had experienced an air travel trip were eligible to participate.

During the survey period from the 18th of October to the 25th of November 2013, 212 completed responses were collected. A total of 104 respondents completed version A of the experiment, while 108 respondents completed version B. To maintain orthogonality, the number of respondents that completed the first version should be equal to the number of respondents that completed the second version (as explained in section 4.2.3). Hence the last four respondents that completed the second version were discarded from the dataset, resulting in a final sample of 208 respondents.

Each of the 208 respondents completed eight choice sets, resulting in a total of 1664 choice observations in the data set. Each row in the data set corresponds to a choice observation. For example, the first row in the dataset corresponds to the first choice observation from the first respondent. Each row contains information about the respondent's socio-economic characteristics, the available alternatives in the choice set and the actual choice the respondent made. Appendix F gives an overview of the variables in the dataset.

Appendix G describes the socio-economic characteristics of our sample. The final sample is well balanced in terms of age (AGE) and gender (GEN). Air travel frequency (ATF) is positively skewed, with 0 and 1 air travel trips in the past 12 months constituting 57% of the total sample and several outliers on the right side of the distribution (40, 24, 14 and 12 air travel trips in the past 12 months). Education level (EDU) is slightly negatively skewed. The four lowest education levels (Primary, LBO, MAVO and VMBO) take account for only 12% of the total sample, the three middle education levels (MBO, HAVO and HBO) constitute 61% of the sample and the highest two education levels (HBO and WO) take account of the remaining 27%.

Figure 5.1 presents the correlation coefficients among the socio-economic characteristics and the version number (VER). The low correlations indicate that there is not sufficient multicollinearity between the socio-economic characteristics to cause problems for the model estimation. Moreover, the low correlations between the version number and the socio-economic characteristics suggest that there are no systematic differences between the group of respondents that completed version A and the group of respondents that completed version B (at least in terms of age, air travel frequency, gender and education level).

Table 5.1 Correlation matrix for socio-economic characteristics and version number

	VER	AGE	ATF	GEN	EDU
VER	1				
AGE	0,04	1			
ATF	0,10	-0,15	1		
GEN	-0,01	0,09	-0,15	1	
EDU	0,03	-0,34	0,08	-0,08	1

5.2. Methodology

We model the choice of an individual, labelled q, facing two alternative airports i captured in choice set C. From each airport i the individual q derives a certain level of utility, denoted U_{iq} $(i = 1, 2; q = 1, .., Q)$. Assuming utility maximising behaviour, individuals choose the alternative from which they obtain the higher level of utility, thus each individual chooses alternative i only if: $U_{iq} > U_{jq}\ \forall\, j \neq i$.

The methodology we use to estimate the discrete choice models is based on random utility theory (McFadden, 1974). The essence of random utility theory is that the total utility an individual derives from an alternative is unknown to the researcher, however the researcher observes certain attributes of the available alternatives and characteristics of the individual that can be related to the total utility. Therefore the total utility U_{iq} can be decomposed into a systematic (*i.e.* observed) element V_{iq} and a random (*i.e.* unobserved) element ε_{iq}, resulting in:

$$U_{iq} = V_{iq} + \varepsilon_{iq} \tag{5.1}$$

The *systematic element* relates to the observed factors, which in this thesis are the underlying attributes of the airports and the socio-economic characteristics of the individuals, as described in the following utility function:

$$V_{iq} = \beta_n X_{ni} + \gamma_n Z_{nq} \tag{5.2}$$

where X_{ni} represent the levels of the n underlying attributes of alternative i; Z_{nq} are the n characteristics of individual q; and β_n and γ_n are the parameters to be estimated.

The *random element* captures all factors that contribute to utility but are not included in the systematic element (*i.e.* not specified in the utility function). Assuming that this random element has a certain distribution, enables calculating the probability P that individual q will choose the i-th alternative from choice set C, which can be described as follows:

$$Pq(i|C) = P\left[U_{iq} > U_{jq}\right] = P\left[\left(U_{iq} = V_{iq} + \varepsilon_{iq}\right) > \left(U_{jq} = V_{jq} + \varepsilon_{jq}\right) \forall\, j \neq i\right] \tag{5.3}$$

where all terms are defined earlier.

Once all parameters of the utility function are estimated, it is possible to calculate the marginal rates of substitution between the attributes. Marginal rates of substitution reveal what respondents are willing to give up from one attribute, for an improvement in another attribute. Given that one of the attributes used in the stated choice experiment represent the price level, the willingness-to-pay for improvements in other attributes is calculated as follows:

$$W_n = \beta_n \,/\, \beta_{price} \;\forall\, n \neq price \tag{5.4}$$

where W_n represents the willingness-to-pay for a unit change in attribute n; β_n is the parameter of attribute n; and β_{price} is the parameter of the price attribute. By expressing all the attributes in a single unit of measurement (*i.e.* monetary value), it is possible to calculate trade-offs between any of the attributes used in the experiment.

Above we explained the conventional way of first estimating the model in *preference space* as in Eq. (5.2) and thereafter derive willingness-to-pay by calculating the ratio of an attribute's coefficient to the price coefficient as in Eq. (5.4). Another way of deriving willingness-to-pay patterns is to estimate the model directly into *willingness-to-pay space*. This method has recently gained in popularity (Scarpa, Thiene, & Train, 2008; Train & Weeks, 2005). One of the advantages is that the coefficient estimates can be directly interpreted as willingness-to-pays and that the model provides standard errors for the willingness-to-pay distribution (Scarpa et al., 2008).

Following Train and Weeks (2005) and Scarpa et al. (2008) we now show how to derive the utility function in willingness-to-pay space. First, decouple the price attribute from Eq. (5.2) and denote the separated price attribute for alternative i as P_i with coefficient δ, as is shown:

$$V_{iq} = \beta_n X_{ni} + \gamma_n Z_{nq} - \delta P_i \tag{5.5}$$

Next utility in willingness-to-pay space is specified as follows:

$$V_{iq} = \delta\,(\beta_n/\delta)\,X_{ni} + \delta\,(\gamma_n/\delta)\,Z_{nq} - \delta P_i \tag{5.6}$$

Note that (β_n/δ) and (γ_n/δ) are essentially equal to willingness-to-pay calculations in Eq. (5.4), hence we denote those W_n. Dividing all terms by δ gives the model specification we use for estimating models in willingness-to-pay space, as follows:

$$V_{iq} = \delta\,(W_n X_{ni} + W_n\,Z_{nq} - P_i) \tag{5.7}$$

where all terms are specified earlier.

After calculation, the W_n coefficients can be interpreted as willingness-to-pay for a unit change in attribute n; δ can be interpreted as the marginal utility of income, *i.e.* the change in utility given a unit change in income. For a more elaborate overview of random utility based discrete choice models please refer to Train (2009).

5.3. Model specifications

The models we estimate are called binary logit (BL) models. Binary logit is essentially a multinomial logit (MNL) model with only two possible choice outcomes (*i.e.* each choice situation offers two alternatives). These models are derived by taking Eq. (5.3) and assuming that all random elements ε_{iq} are independent identically extreme value (Gumbel) distributed (Train, 2009). Moreover, we assume that all 1664 choice observations in the dataset are independent. Note that this is unlikely since the choices made by the same individual are probably correlated with each other. However, Louviere, Hensher, and Swait (2000) note that this assumption is valid, as long as the choice experiment encourages respondents to view each choice situation as independent. We tried to ensure this by preventing respondents to return to previous choice situations (see section 4.2.3).

5.3.1. Base model specification

The models analyse the airport choices that respondents made during the experiment. For each choice we observed the alternatives available in the choice set, the attributes of the alternatives and certain socio-economic characteristics of the respondents. Given the observed factors and assuming that each respondent maximises utility, the systematic utility function (in willingness-to-pay space) for each airport alternative i and individual q is then specified as follows:

$$V_{iq} = \delta_1 \left(-FAR_i + W_1 ACC_i + W_2 CON_i + W_3 CIR_i + W_4 AES_i + W_5 COU_i\right) \tag{5.8}$$

where the choice for an airport depends on the price level FAR; the access time to the airport ACC; the availability of restaurants and shops in the airport terminal CON; the level of crowdedness in the airport terminal CIR; the cleanliness of the airport terminal AES; the friendliness of the airport staff COU; δ_1 is the marginal utility of income; and the Ws are the willingness-to-pay coefficients to be estimated. We refer to this model as the *base model specification* (BL_{base}).

Fare level and *access time* are included as continuous variables. The actual numerical attribute level labels used in the experiment (recall that these are 80, 90, 100 or 110 EURO; and 60, 75, 90 and 105 minutes respectively) enter the model through FAR and ACC. *Availability of concessions*, *congestion level*, *cleanliness* and *courtesy of staff* enter the model as dummy variables. CON, CIR, AES and COU take the value of 1 if respectively the availability of concessions is great, the terminal is not crowded, the terminal is clean or the airport staff is friendly and the value of 0 every time the availability of concessions is poor, the terminal is crowded, the terminal is not clean or the staff is unfriendly. This coding structure should be taking into account when interpreting results in section 5.4.

5.3.2. Full model specification

Furthermore, we want to control for potential confounding of the main effects by differences in socio-economic characteristics among respondents. Hence we add interaction terms between all observed socio-economic characteristics and main effects to the base model specification:

$$\begin{aligned} V_{iq} = \delta_1 [&-FAR_i + W_1 ACC_i + W_2 CON_i + W_3 CIR_i + W_4 AES_i + W_5 COU_i + \gamma_6 (FAR_i \left(AGE_q - AGE'\right)) + \gamma_7 (FAR_i \left(ATF_q - ATF'\right)) + \gamma_8 (FAR_i GEN_q) + \gamma_9 (FAR_i EDUm_q) + \\ &\gamma_{10} (FAR_i EDUh_q) + \gamma_{11} (ACC_i \left(AGE_q - AGE'\right)) + \gamma_{12} (ACC_i \left(ATF_q - ATF'\right)) + \\ &\gamma_{13} \left(ACC_i GEN_q\right) + \gamma_{14} \left(ACC_i EDUm_q\right) + \gamma_{15} \left(ACC_i EDUh_q\right) + \gamma_{16} (CON_i \left(AGE_q - AGE'\right)) + \gamma_{17} (CON_i \left(ATF_q - ATF'\right)) + \gamma_{18} (CON_i GEN_q) + \gamma_{19} (CON_i EDUm_q) + \\ &\gamma_{20} (CON_i EDUh_q) + \gamma_{21} (CIR_i \left(AGE_q - AGE'\right)) + \gamma_{22} (CIR_i \left(ATF_q - ATF'\right)) + \\ &\gamma_{23} (CIR_i GEN_q) + \gamma_{24} (CIR_i EDUm_q) + \gamma_{25} (CIR_i EDUh_q) + \gamma_{26} (AES_i \left(AGE_q - AGE'\right)) + \gamma_{27} (AES_i \left(ATF_q - ATF'\right)) + \gamma_{28} (AES_i GEN_q) + \gamma_{29} (AES_i EDUm_q) + \\ &\gamma_{30} (AES_i EDUh_q) + \gamma_{31} (COU_i \left(AGE_q - AGE'\right)) + \gamma_{32} (COU_i \left(ATF_q - ATF'\right)) + \\ &\gamma_{33} \left(COU_i GEN_q\right) + \gamma_{34} \left(COU_i EDUm_q\right) + \gamma_{35} \left(COU_i EDUh_q\right)] \end{aligned} \tag{5.9}$$

where AGE represents the respondent's age; ATF the respondent's air travel frequency; GEN the respondent's gender; $EDUm$ and $EDUh$ the respondent's education level; AGE' is a centering variable for AGE; ATF' is a centering variable for ATF; the γs are the interaction coefficients to be estimated; and all other terms are defined earlier. We refer to this model as the *full model specification* (BL_{full}).

It is important to note that we only want to control for the effects of socio-economic characteristics and that we are not particularly interested in these effects (*e.g.* the difference between males and females, or elder and younger passengers) *per se*. Ultimately, we want to estimate *overall mean* willingness-to-pay coefficients that are irrespective of the age, gender, air travel frequency and education level of the respondent. To this end, we use *mean centering variables* for the socio-economic characteristics that are included as continuous variables (*i.e.* age and air travel frequency) and *effects coding structure* for the socio-economic characteristics that are included as categorical variables (*i.e.* gender and education level) . Note that if we would not mean center the age and air travel frequency variables, the willingness-to-pay coefficients W would represent the willingness-to-pay of a passenger that is zero years of age and has travelled zero times in the past 12 months. Similarly, if we would use the more conventional dummy coding structure instead of effects coding structure for the gender and education level variables, we would perfectly confound the willingness-to-pay coefficients W with one of the gender and education level categories instead of calculating an overall mean effect. See Appendix H for the mathematical rationale of using effects coding, or refer to Hensher et al. (2005, pp. 119-120) and Louviere et al. (2000, pp. 86-87) for further elaboration.

Age and *air travel frequency* are included as continuous variables. Integers representing the respondent's age and air travel frequency enter the model directly through AGE and ATF. The mean centering variable for age AGE' is fixed at the mean of age (= 42, rounded for convenience of interpretation). Since preliminary data exploration of our sample indicated that air travel frequency has a positively skewed distribution with several outliers (see Appendix G), we decided to fix the mean centering variable for air travel frequency ATF' at the *median* of air travel frequency (= 1) which is more robust to outliers.

Gender and *education level* are included as categorical variables, with two- (male, female) and three categories (lower, middle, higher) respectively. The lower education category represents primary school, LBO, MAVO and VMBO, the middle education category MBO, HAVO and HBO and the higher education category VWO and WO[4]. Following Hensher et al. (2005, pp. 119-120) we effect code the gender and education level variables as follows. We create a number of new variables equal to the number of categories minus one. That is, one new variable for gender (denoted GEN) and two new variables for education level (denoted $EDUh$ and $EDUm$). GEN takes the value of -1 every time the respondent is a male and the value of 1 every time the respondent is a female. $EDUh$ is associated

[4] We tried other categorisations of education level in our models. For example with a higher category representing HBO and WO, the middle category MBO, HAVO and VWO and the lower category primary school, LBO, MAVO and VMBO. This resulted in unidentifiable models with multiple highly insignificant parameter coefficients, leading us to the conclusion that the current categorisation is preferable.

with the higher education level such that every time the respondent belongs to the higher education category this variable takes the value of 1 and when the respondent belongs to the middle education category this variable equals 0. Similarly, $EDUm$ represents the variable for middle education level, taking the value of 1 if the respondent belongs to the middle education category and the value of 0 if the respondent belongs to the higher education category. When a respondent belongs to the lower education category, both variables take the value of -1. Table 5.2 shows this effects coding structure.

Table 5.2 Effects coding structure for education level

Variable	$EDUh$	$EDUm$
Education category		
Higher	1	0
Middle	0	1
Lower	-1	-1

5.3.3. Final model specification

The *final model specification* (BL_{final}) is specified by excluding the interaction parameters that are not significant at the 80% confidence level in the BL_{full} model estimation. In this context, the education level parameters deserve additional attention. Recall from the previous section that we created an effect coding structure which uses two variables to represent education level ($EDUh$ and $EDUm$). This also meant that for estimating the interaction between one of the main parameters, for example the price level, and the education level we needed to create two interaction parameters ($FAR \cdot EDUh$ and $FAR \cdot EDUm$). This raises the question of what to do when one of the interaction parameters is not significant at the 80% confidence level, while the other interaction parameter is significant. For example, $FAR \cdot EDUh$ may be significant at the 80% confidence level, while $FAR \cdot EDUm$ is not significant at the 80% confidence level. In this situation $FAR \cdot EDUm$ should still be included in the model specification (although not significant), because if we would exclude the interaction between middle education category and price and at the same time include the interaction with middle education category and price, we would perfectly confound the price coefficient with the high education category instead of calculating an *overall mean price effect*[5] (Hensher et al., 2005, p. 351).

5.3.4. Pre-estimation expectations

Fare level and *access time* are expected to have a negative willingness-to-pay coefficient, as money already spent on air tickets cannot be spent on other goods and time spent accessing the airport cannot be used anymore for other activities[6]. The coefficients of the airport terminal service quality attributes can logically take either sign. However, common sense indicates that the willingness-to-pay for *cleanliness* and *friendliness of staff* are most likely not negative. Regarding the *availability of restaurants and shops*, some passengers might like the opportunity to engage in shopping and eating

[5] See Appendix I for mathematical details.

[6] Neglecting the rare but plausible occasion of a passenger that enjoys the activity of driving to the airport

out at the airport and therefore they probably have a positive willingness-to-pay for an airport with a great availability of restaurants and shops. On the other hand passengers may associate the presence of a lot of shops and restaurants with a packed and confusing terminal design in which they struggle to find their ways. These passengers might have a negative willingness-to-pay for a great availability of restaurants and shops. The same can be argued for the effect of *crowdedness*, as some passengers might enjoy the presence of a lot of people, while others do not as they may perceive busy terminals with potentially longer queue times and less space for movements.

5.4. Results

Table 5.3 reports the coefficient estimates, robust standard errors and the final log-likelihood of the binary logit model estimations. All reported models are estimated with BIOGEME Version 1.8 (Bierlaire, 2003). The first two models (BL_{base} and BL_{full}) correspond to Eq. (5.8) and Eq. (5.9) respectively. After estimation of these models, the BL_{final} model was specified according to the procedure described in section 5.3.3. This procedure for removal of parameters led to the exclusion of 14 parameters from the BL_{full} model, resulting in the BL_{final} model comprising 22 parameters. The models are estimated in willingness-to-pay space, hence the coefficients can be directly interpreted as willingness-to-pays[7]. As we primarily control for interaction effects and are not so much interested in these effects *per se*, we do not report the coefficients of the interaction parameters here. Please refer to Appendix J for the complete estimation results.

Table 5.3 Estimation results for binary logit (BL) models

	BL_{base}		BL_{full}		BL_{final}	
	Coefficient	Rob. s.e.	Coefficient	Rob. s.e.	Coefficient	Rob. s.e.
Marg. utility of income (δ)	0.0816 ***	0.0049	0.0961 ***	0.0065	0.0950 ***	0.0061
Access time (min.)	-0.5600 ***	0.0350	-0.5440 ***	0.0355	-0.5450 ***	0.0344
Availability of restaurants and shops	2.39 **	1.0600	3.57 ***	1.0600	3.10 ***	1.0000
Crowdedness	2.17 *	1.1400	1.49	1.0800	1.83 *	1.0500
Cleanliness	13.70 ***	1.0400	12.50 ***	1.0200	12.20 ***	0.9660
Friendliness	11.40 ***	0.9710	9.44 ***	0.9680	9.67 ***	0.9150
Incl. socio-economics	NO		ALL[a]		SELECTED[b]	
Observations	1664		1664		1664	
Parameters	6		36		22	
Init LL	-1153.397		-1153.397		-1153.397	
Final LL	-766.524		-721.686		-725.868	

*, **, *** indicates significance at 90, 95 and 99% level, respectively.
[a] Included all two-way interactions between socio-economic characteristics and main effects in the model specification – see Eq. (5.9)
[b] Included selected two-way interactions: For *fare* and *access time* with age, air travel frequency, gender and education level; for *availability of restaurants and shops* with age and air travel frequency; for *crowdedness* with air travel frequency; for *cleanliness* with age; and for *friendliness of staff* with age and air travel frequency.

[7] Note that we also estimated the models in 'preference space' and derived willingness-to-pays using Eq. (5.4), which apart from some rounding differences yield similar willingness-to-pay estimates.

5.4.1. Goodness of fit

Table 5.3 reports the *init log-likelihoods* (Init LL) and the *final log-likelihoods* (Final LL). The init log-likelihood is the log-likelihood of a model where all parameters are equal to zero (*i.e.* no model). The *final log-likelihood* is the log-likelihood of the estimated models. Values of the final log-likelihood that are closer to zero indicate better model fit.

We deploy *likelihood ratio-tests* (Train, 2009, pp. 78-79) to test whether one model is a statistically significant improvement over another model. The null hypothesis of the likelihood ratio-test is that one model does not significantly improve the other model. Next, the test is computed as follows:

$$-2(LL_1 - LL_2) \sim X^2_{d.f. equal\ to\ difference\ in\ number\ of\ parameters} \tag{5.10}$$

where LL_1 is the log-likelihood of one model; LL_2 is the log-likelihood of the other model; and $X^2_{d.f.}$ is a critical chi-square statistic with degrees of freedom equal to the difference in number of parameters used in the two models. The value of the -2LL function is compared to the $X^2_{d.f.}$ statistic, to obtain the confidence level with which the null hypothesis can be rejected.

Comparing the BL_{base} model to the BL_{full} model gives an improvement in the log-likelihood of 44.838 units at the cost of 30 additional parameters, which according to the likelihood ratio-test is significant at the 99% level of confidence. This suggests that controlling for interactions with socio-economic characteristics significantly improves the model performance and should not be neglected. The BL_{full} model also offers an improvement in the log-likelihood of the BL_{final} model by 4.182 units. However, at the cost of 14 additional parameters this improvement is not significant at the 90% level. Thus the parsimony benefits offered by the BL_{final} model (*i.e.* the model uses less parameters) are not offset by the small decrease in model performance compared to the BL_{full} model, leading to the conclusion that the BL_{final} model is the preferred model.

5.4.2. Coefficient estimates

Without trivialising the above discussion about which model performs best, it is noteworthy that the differences in coefficient estimates between the different model specifications are modest. All coefficients have the expected and same signs in each of models. Moreover, except for the coefficient relating to crowdedness, all reported coefficients are significant at the 95% confidence level, with most at the 99% level. The coefficient relating to crowdedness is significant only at 90% level in the BL_{base} model and the BL_{final} model. Moreover, after controlling for all socio-economic characteristics in the BL_{full} model the crowdedness coefficient does not remain significant at all.

Following our conclusion that the BL_{final} model is the preferred model, we focus our interpretation of the coefficient estimates on the estimates obtained from that model. Before discussing the willingness-to-pay implications, we interpret the *Robust Wald-statistic*s. Wald-statistics relate to the explanatory power of the attributes and can be interpreted as an indication of their influence on the airport choices made by the leisure passengers during the experiment (Hensher et al., 2005, pp. 343-345). Robust Wald-statistics are computed as follows:

$$Robust\ Wald\ = \frac{Coefficient_i}{Robust\ std\ err_i} \tag{5.11}$$

where $Coefficient_i$ is the coefficient estimate of parameter i; and $Robust\ std\ err_i$ is the robust standard error estimate of parameter i.

The Robust Wald-statistics are largest for the marginal utility of income (= 15.55) and access time attributes (= -15.82), suggesting that the price and access time levels have the largest influence on the airport choices. This might be a reflection of the cost-consciousness of leisure passengers and the high importance they attach to airport access times. Besides, we find high Robust Wald-statistics for the cleanliness (= 12.67) and friendliness (= 10.57) attributes, indicating that leisure passengers have a great concern for the cleanliness of the airport terminal and the friendliness of airport staff. In contrast, the Robust Wald-statistics of the attributes relating to the availability of restaurants and shops (= 3.1) and airport terminal crowdedness (= 1.74) are much lower. Apparently, passengers' perception of these airport terminal service quality attributes only have a small influence on their departure airport choice. Interestingly, the Robust Wald-statistic of airport terminal crowdedness is lower than the critical Wald-value of 1.96 (assuming a 95% confidence level). This suggests that airport terminal crowdedness does not significantly influence leisure passengers' airport choice behaviour.

5.4.3. Willingness-to-pay implications

Table 5.3 summarises the willingness-to-pay and willingness to accept increases in access time[8] estimates obtained from the BL_{final} model.

Table 5.4 Willingness-to-pay and willingness to accept increases in access time

Parameter	EURO	Minutes
Access time reductions (1h)	32.70 EURO	-
Great availability of restaurants and shops relative to poor availability of restaurants and shops	3.10 EURO	5.69 Min.
Uncrowded relative to crowded	*1.83 EURO	*3.36 Min.
Clean relative to not clean	12.20 EURO	22.39 Min.
Friendly staff relative to unfriendly staff	9.67 EURO	17.74 Min.

* Only significant at the 90% level

[8] The willingness-to-pay estimates are equal to the coefficients reported in table 5.3. The willingness to accept increases in access time are obtained by calculating the ratio of the willingness-to-pay for the airport terminal service quality dimensions with the willingness-to-pay for a minute reduction in access times.

Looking first at the willingness-to-pay estimate for access time, the coefficient for access time in the BL_{final} model is highly significant with the expected negative sign (see Table 5.3). Translating this coefficient into a commonly used measure in transportation studies, called the *value of travel time savings* (VTTS), suggests that the Dutch leisure passengers in our sample are willing to pay EURO 32.70 for each hour reduction of access time[9]. This high value is broadly consistent with values of travel time savings for leisure passengers, reported by Adler et al. (2005) ($26.1/h ≈ EURO 25.25/h), Hess et al. (2007) ($35.8/h ≈ EURO 34.63/h) and Hess (2010) ($22.43/h ≈ EURO 21.27/h)[10]. Hence, in line with prior research, the value of access time savings we found reflect that access time is an important determinant of passengers' departure airport choice.

Turning to the willingness-to-pay estimates for the airport terminal service quality dimensions, our results reveal a positive willingness-to-pay for each of the included airport terminal service quality attributes. The respondents displayed a preference for departing from airports that were characterised by higher levels of service quality. Furthermore, the expected (positive) signs indicate that none of the airport terminal service quality attributes was completely ignored during the experiment.

However, the magnitudes of the estimates differ strongly depending on which airport service quality attribute is considered. The willingness-to-pay estimates suggest that the leisure passengers in our sample are willing to pay a surplus of EURO 12.20 to depart from an airport that they perceive as clean, relative to an airport that they perceive as not clean, *ceteris paribus*. Relative to unfriendly, the passengers are willing to pay an additional EURO 9.67 for an airport where they perceive the staff as friendly, *ceteris paribus*. Translating these willingness-to-pay estimates into willingness to accept increases in access time, suggests that the leisure passengers in our sample are willing to drive up to ≈ 22 and ≈ 18 minutes further to access an airport that is respectively perceived as clean or as having a friendly staff, if the closer airport is perceived as being not clean or having unfriendly staff, *ceteris paribus*. On the other hand, the willingness of leisure passengers to pay EURO 3.10 for departing from an airport where the availability of restaurants and shops is perceived as great, relative to an airport where the availability of restaurants and shops is perceived as poor, is very small. Similarly, the leisure passengers' willingness-to-pay for departing from an uncrowded airport terminal relative to a crowded airport terminal is only EURO 1.83 and moreover hardly significant.

Our finding regarding the availability of restaurants and shops is especially interesting, considering the common belief that passengers have a great concern for activities (*e.g.* shopping, eating out) to occupy themselves during their time spent waiting in the airport terminal (see for example Fodness & Murray, 2007). And that this is especially the case for leisure passengers, who typically arrive at the

[9] Hess et al. (2007) rightfully notes that these VTTS values cannot be compared to standard measures of VTTS (for example the values found by a recent comprehensive study in the Netherlands of KiM, 2013), for the reason that we used an air fare coefficient instead of an access time coefficient in calculating the VTTS.

[10] Values in EURO are calculated by first correcting for inflation using the CPI Inflation Calculator (U.S. Bureau of Labor Statistics, 2013) and next using the exchange rate of 1 U.S. Dollar = 0.7351 Euro.

airport terminal well in advance and thus face high amounts of airport dwelling time (Lemer, 1992). A possible explanation for the contrast between these premises and our findings is the time period in which the referred studies were conducted. During the time that passed, the emergence of low-cost carriers opened the air travel markets for the “masses”, increasing the price sensitivity of demand for (leisure) air travel. This price sensitivity also appears to have its influence on airport choice behaviour, potentially leading to today’s leisure passengers’ reluctance to accept higher ticket prices in exchange for greater availability of restaurants and shops. In this context, our results can be interpreted as further proof of the changed expectations of leisure passengers and coincide with the movement of leisure passengers to smaller, regional airports, that typically offer less opportunity to engage in shopping and eating out, in return for lower cost flights.

Similarly, our findings suggest that leisure passenger do not regard an uncrowded airport terminal as a necessary element of airport services and readily accept higher crowding levels for lower ticket prices. These findings are more or less in line with those of Hess (2010), who found that, even though passengers might expect more congestion at larger airports, they still prefer these larger airports. Apparently, the expectations passengers have about airport terminal crowdedness have little influence on their departure airport choice decisions. In conjunction with the findings regarding availability of restaurants and shops, we hypothesise that airport terminal crowdedness as well as the availability of restaurants and shops may be regarded as *nice-to-haves* rather than *need-to-haves* in the opinion of the cost-conscious leisure passenger.

Despite the (supposedly) increased price sensitivity, our results show that leisure passengers still have great concern for airport terminal service quality dimensions such as airport terminal cleanliness and friendliness of airport staff. One reason for the large influence and willingness-to-pay estimates we found for these airport terminal service quality dimensions might possibly be that leisure passengers regard them as elementary components of airport service. In this context, one might raise the question whether passengers’ perception of airport terminal cleanliness and friendliness of staff truly play a role in real-world (*i.e.* non-experimental) airport choice decisions. Ultimately, the actual influence of a service quality dimension on airport choice behaviour depends on the *difference* in (perceived) performance on the dimensions between the airports that are considered (Garver, 2003). This implies that before cleanliness and friendliness can influence airport choice behaviour, passengers must first consider them as factors on which airports can differentiate and not as given elements of airport service.

Another reason for the substantial willingness-to-pay we found for airport terminal cleanliness and friendliness of airport staff may be that passengers perceptually correlate these service dimensions to other facets of airport service quality. Poor cleanliness of an airport terminal might lead passengers to believe that other aspects of the airport are carelessly organised as well. Similarly, the friendliness of staff may directly influence passengers’ expectations of the staffs’ problem-solving behaviour if things go wrong (*e.g.* wrong booking information, delays, flight cancellation). This however is mostly speculative. Nevertheless, our results stress the importance for airports to maintain a good reputation when it comes to cleanliness and friendliness.

5.5. Conclusion

In this chapter we empirically explored the relationship between passengers' perception of airport terminal service quality and their choice for a departure airport. The analysis was carried out by estimating a discrete choice model of airport choice using stated preference data obtained from a sample of Dutch leisure passengers.

In our model, the airports are characterised by ticket price, access time, availability of restaurants and shops, crowdedness, cleanliness and friendliness of the airport staff. The estimated willingness-to-pay and willingness to accept increases in access time coefficients reveal the relative importance that the respondents in our sample assign to each of the included airport terminal service quality dimensions. Furthermore, we controlled for potential confounding of the main coefficients by estimating additional models that include all or selected two-way interactions between the main parameters and socio-economic passenger characteristics. Based on measures of model fit, we conclude that the model specification with the selected two-way interactions is the preferred model.

The model results show that the price- and access time levels had the largest influence on the airport choices made during the experiment. This underlines the common finding of prior research that ticket prices and airport access times are important determinants of departure airport choice. Interestingly, it was not possible to retrieve meaningful effects for greater availability of restaurants and shops and lower airport terminal crowdedness, with the effect of the latter being hardly statistically significant. This suggests that leisure passengers' perceptions of competing airport's performance on these airport terminal service quality dimensions barely influence their choice among these airports. In contrast, higher levels of cleanliness of the airport terminal and friendliness of the airport staff considerably improved the chances that an airport was chosen during the experiment, indicating a great concern of leisure passengers for clean airport terminals and friendly airport staff.

These findings reconfirm that today's leisure passengers are highly price-sensitive and highlight the great importance of access times in airport choice behaviour. Regarding airport terminal service quality, the results suggests the presence of a dichotomy between *nice-to-have* and *need-to-haves*. The cost-conscious leisure passengers' air travel choice behaviour is in all likelihood primarily governed by ticket prices, leading them to generally choose the departure airport that offers the flight with the lowest ticket price (assuming all other flight services are perceived as equal) that they can reach within a reasonable access time. Differences in perceived performance of alternative airports on service quality dimensions that are regarded as *nice-to-haves*, such as the opportunity to engage in shopping and eating out at the airport terminal or lower airport terminal crowding levels, can be offset by very small differences in ticket prices. Hence the influence from such factors on departure airport choice is likely very small. Nevertheless, our results also indicate that the cost-conscious leisure passengers still have a great concern for factors such as airport terminal cleanliness and friendliness of staff, presumably because they regard these airport terminal service quality dimensions as elementary facets of airport services (*i.e. need-to-haves*).

"There is hardly anything in the world that some man cannot make a little worse and sell a little cheaper, and the people who consider price only are this man's lawful prey."
(John Ruskin, 1819 – 1900)

6. Conclusion

This thesis focused on the influence of (perceived) airport terminal service quality on passengers' airport choice behaviour in an attempt to contribute to a better understanding of this scarcely investigated matter. To this end, insights from the field of literature on airport choice behaviour and the field addressing airport terminal service quality are combined. In addition, the analysis involved the estimation of discrete choice models of airport choice in relationship to varying levels of airport terminal service quality, making use of stated preference data collected from a sample of Dutch leisure passengers. In this final chapter we present the main findings, formulate conclusions and implications and discuss a number of limitations and directions for further research.

This chapter is structured as follows. Section 6.1 briefly summarises the content and main findings of the previous chapters. Section 6.2 discusses the main implications. Section 6.3 discusses limitations and directions for further research.

6.1. Findings

Chapter 2 studied the current understanding of passengers' airport choice behaviour by reviewing the airport choice literature. Although the literature in this field is well-advanced, the influence of airport terminal service quality on airport choice is scantily investigated. The few studies addressing this topic used overly generic measures of service quality, such as airport-specific constants (see for example Başar & Bhat, 2004; Hess & Polak, 2005) or airport rankings (see for example Adler et al., 2005; Hess et al., 2013), that conceal the influence of *distinctive* airport terminal service quality dimensions. Nonetheless, the estimated influence of such measures is often substantial and statistically significant, indicating that airport-specific attributes, possibly related to the service quality offered in the airport terminal, influences passengers' airport choice behaviour.

Chapter 3 examined several studies that address airport terminal service quality from a passengers' perspective. Typically, the main focus of such studies is to identify relevant dimensions of airport terminal service quality and subsequently assess passengers' beliefs (in terms of importance and performance) regarding these dimensions. Unfortunately, a well-specified link between passengers' perceptions of airport terminal service quality and their choice for a departure airport is lacking. Furthermore, the literature within this field is highly ambiguous. At present, there does not exist a widely-accepted model of airport terminal service quality from a passengers' point of view. Nevertheless, based on the reviewed studies we distinguished a categorisation of dimensions that comprise airport terminal service quality from a passengers' perspective. This categorisation includes concessions and amenities, circulation, aesthetics, courtesy of staff, information provisioning, comfort, waiting- and processing time, convenience, security and WIFI facilities.

Chapters 4 and 5 empirically analysed the influence of airport terminal service quality on airport choice behaviour. Chapter 4 designed a stated choice experiment, as a means to generate a data set of 'stated' airport choices of Dutch leisure passengers in relationship to varying levels of airport terminal service quality. Chapter 5 estimated a random utility based discrete choice model of airport choice using the generated data set. In this model, the choice for a departure airport depends on the ticket price, access time, availability of restaurants and shops, airport terminal crowdedness, cleanliness of the airport terminal and friendliness of airport staff. Assuming that all attributes not included in the experiment are equal, the model results suggest that price- and access time levels have the largest influence on leisure passengers' departure airport choice. Furthermore, the results show that leisure passengers have a great concern for airport terminal cleanliness and friendliness of airport staff, thus suggesting that perceived performance of airports on these two service quality dimensions can substantially influence airport choice behaviour. Interestingly, it was not possible to retrieve meaningful effects for availability of restaurants and shops and airport terminal crowdedness, with the latter being hardly statistically significant.

In line with prior research, the model results highlight the cost-consciousness of leisure air passengers and the important role of access time in airport choice decisions. Regarding the influence of airport terminal service quality on departure airport choice, the results potentially hint at the presence of a dichotomy between service quality dimensions that are perceived as elementary facets of airport services (*i.e.* airport terminal cleanliness and friendliness of airport staff), and service quality dimensions that are perceived as non-essential facets of airport services (*i.e.* availability of restaurants and shops and airport terminal crowdedness).

6.2. Main implications

The most important implication of this thesis is that today's leisure passengers appear to be primarily in search of a *no-frills airport service*. Their airport choice behaviour is for a large part dictated by ticket prices and ground access times and they are generally unwilling to accept higher ticket prices and increased access times for higher quality of non-essential airport terminal service dimensions. This distinctly coincides with the commercial success of no-frills airlines and the attraction that these airlines exert on the leisure segment in particular. In this context, our research reconfirms that the leisure air transport segment is highly price-sensitive and that this price-sensitivity governs both the airline- as well as the airport choices made by the passengers in this segment. Nevertheless, our research suggests that the price-sensitive leisure passengers still attach great value to airport terminal service quality dimensions such as the cleanliness of the airport terminal and the friendliness of the airport staff. One reason may be that passengers perceive such airport service quality dimensions as elementary facets of airport services.

If the findings of this thesis can be replicated by other studies, an important practical implication for airport management and policy makers is that airport's perceived performances on non-essential dimensions of airport terminal service quality do not have a large influence on leisure passengers' decisions for a departure airport. Hence when prioritising investments in airport terminals with the aim of attracting additional passengers, sufficient efforts should be dedicated on evaluating which facets of the airport terminal are capable of influencing passengers' airport choice decisions. On the other hand, airport management and policy makers should be careful with promptly labelling all service quality dimensions as "frills", as passengers might actually perceive specific airport service quality dimensions as elementary facets of airport services and hence airport's performance on these service quality dimensions plays an important role in departure airport choice decisions.

In relation to Dutch practice, the findings are in line with the growing numbers of leisure passengers departing from regional no-frills airports that typically offer less airport terminal service quality in exchange for lower cost flights. The inferences of this thesis may also contribute to the ongoing debate over shifting leisure flights and -passengers from Amsterdam Schiphol Airport to Lelystad Airport (DutchNews, 2013). The much smaller Lelystad Airport will likely not be able to offer the same levels of airport terminal service quality as Amsterdam Schiphol Airport. However, as long as Lelystad Airport performs well on the elementary facets of airport services and given that the airport is adequately accessible, the effect on the customer satisfaction level of leisure passengers might be very small.

Moving on to the theoretical perspective, the findings of this thesis show that the magnitude of the influence can strongly differ among the various airport terminal service quality dimensions. This essentially confirms the importance of investigating the influence of *distinctive* airport terminal service quality on airport choice behaviour. This thesis presented stated choice experiments and -modelling as appropriate methods for such analyses. Some main benefits of stated choice experiments are the approximation of real-life decision making and their capability of forcing respondents to make trade-offs between airport terminal service quality dimensions, hence stimulating them to reveal what they consider truly important.

6.3. Limitations and direction for further research

To the best of our knowledge, our study represents the first airport choice study explicitly focused on the influence of qualitative airport terminal service dimensions, thereby valuably contributing to the scarce understanding of this matter. Nonetheless, our study has a number of drawbacks.

Firstly, our study deals with perceptions. Respondents likely attach different meanings to perceptual attributes, leading to so-called *attribute ambiguity* (Hensher et al., 2005). To understand this, imagine a plausible situation in which one respondent might perceive a certain airport terminal as clean, while another respondent might perceive the exact same terminal as not clean. Nonetheless, our research is aimed at understanding the ways along which the passengers evaluate their airport choice decisions. Unfortunately, at least from a researcher's perspective, qualitative dimensions are evaluated by passengers in terms of perceptions and beliefs and therefore we argue that, when truly seeking the passengers' perspective one cannot avoid having to deal with the ambiguities and uncertainties that come along[11].

Secondly, the use of stated choice experiments constrained the number of attributes we were able to investigate (Caussade et al., 2005). Hence we could not include any attributes relating to the flight services offered from the airports and respondents had to assume that all the airports offered exactly the same flight (*i.e.* same airline, same departure- / arrival time, same travel time, routings, etc.). It is not hard to imagine that this assumption never holds in practice. As an additional consequence we are not able to estimate trade-offs between levels of airport terminal service quality and levels of flight services from the data we collected. For the same reason, we could only investigate four airport terminal service quality dimensions. A more exhaustive list of airport terminal service quality dimensions (among others information provisioning, processing time and security) may be investigated in further research.

Thirdly, the stated choice experiment in the present study uses dichotomous attribute levels (*e.g.* an airport terminal is either clean or not clean). Dichotomous attribute levels represent a major simplification of reality. Instead of either clean or not clean, passengers may perceive one airport as being 60% clean, with the other airport as being 80% clean. Moreover, by using only two attributes, any potential non-linear relationships remain hidden[12]. On the other hand, using only two dichotomous attribute levels greatly enhanced the ease with which the respondents could grasp the choice situations during the experiment, thereby contributing to the quality of the provided responses. It is left for further research to explore whether non-linear relationships exist between perceptions of airport terminal service quality and airport choice.

Further, two remarks should be made regarding the quality of the data. First, the use of a nonprobability convenient sampling technique (*i.e.* snowballing) potentially limits the generalisability of the results, as it might cause certain subgroups of the population to be overrepresented, with others being underrepresented. However, this issue is partly countered by having a relative large sample size (n = 208) and, furthermore, data exploration showed that our sample is at least well balanced in terms of age and gender. Secondly, our sample is limited to leisure passengers only. Literature addressing airport choice continuously stresses that travellers with different trip purposes (*e.g.* leisure, business, visiting friends and family) have different behaviour and, therefore, further research addressing airport terminal service quality should take these differences into account.

[11] In this context, we support the argument of Louviere et al. (2000, p. 257), that consumers are always right about the way they think about products as they ultimately are the buyers of the products.

[12] Some passengers might be willing to pay substantially for departing from an airport that they perceive as 'clean enough' relative to an airport that they perceive as 'not clean', but are not willing to pay any surplus for departing from an airport that is 'very clean' relative to an airport that is 'clean enough'.

Transfer passengers, for example, might be more concerned about the availability of activities to occupy themselves in airport terminals, due to having lengthy layovers. Similarly, business travellers might be more concerned about airport terminal crowdedness, due to the greater desire for quietness in order to engage in business activities during their airport dwelling time.

At this point we should also stress that our results and associated inferences are purely based on a passengers' perspective. Consider for example the availability of restaurants and shops. Even supposing that passengers' airport choice behaviour is barely influenced by this service quality dimension, airport management may still have a very strong incentive for offering a great variety of restaurants and shops, for instance because the revenues of such concessions constitute a major source of the airport's profits.

Furthermore, the method presented in this thesis can be improved by the use of more advanced modelling solutions. In this context, the use of panel mixed logit models is especially promising, as these models are capable of capturing the (unobserved) taste heterogeneity that is likely present in the influence of qualitative airport terminal service dimensions on airport choice behaviour due to the subjective nature of service quality.

Finally, throughout this thesis it is recurrently argued that the role of airport terminal service quality on airport choice behaviour is highly relevant for today's air transport market, while at the same time this topic received scant attention in prior research efforts. By drawing a connection between the studies on airport choice behaviour and the studies on airport terminal service quality, this thesis contributed to both fields of literature. At this moment we should however stress that the empirical findings of this thesis are limited to one data set and subject to a number of limitations, hence the inferences of this thesis should be studied more elaborately by further research. In the end, this thesis hopes to be the first step towards to a better understanding of the relationship between (perceived) airport terminal service quality dimensions and airport choice behaviour.

References

ACI. (2013). Airport Service Quality - The World Leader in Airport Service Benchmarking. Retrieved 7 november, 2013, from http://www.airportservicequality.aero/

Adler, T., Falzarano, C. S., & Spitz, G. (2005). Modeling service trade-offs in air itinerary choices. *Transportation Research Record: Journal of the Transportation Research Board, 1915*(-1), 20-26.

Arentze, T., Borgers, A., Timmermans, H., & DelMistro, R. (2003). Transport stated choice responses: effects of task complexity, presentation format and literacy. *Transportation Research Part E: Logistics and Transportation Review, 39*(3), 229-244.

Başar, G., & Bhat, C. (2004). A parameterized consideration set model for airport choice: an application to the San Francisco Bay area. *Transportation Research Part B: Methodological, 38*(10), 889-904.

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

Bierlaire, M. (2008). Estimation of discrete choice models with BIOGEME 1.8. biogeme.ep.ch.

Bonnefoy, P. A., de Neufville, R., & Hansman, R. J. (2010). Evolution and Development of Multiairport Systems: Worldwide Perspective. *Journal of Transportation Engineering, 136*(11), 1021-1029.

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

CBS. (2013). Luchtvaart; passagiersvervoer op Nederlandse luchthavens. from Centraal Bureau voor de Statistiek http://statline.cbs.nl/StatWeb/publication/?DM=SLNL&PA=37570hvv&D1=12-17&D2=1&D3=0&D4=5-14&HDR=T,G1&STB=G2,G3&VW=T

Cheaptickets. (2013). Online booking system. Retrieved 31 August, 2013, from http://www.cheaptickets.nl/

De Luca, S. (2012). Modelling airport choice behaviour for direct flights, connecting flights and different travel plans. *Journal of Transport Geography, 22*, 148-163.

DeShazo, J., & Fermo, G. (2002). Designing choice sets for stated preference methods: the effects of complexity on choice consistency. *Journal of environmental economics and management, 44*(1), 123-143.

DutchNews. (2013). Schiphol airport's €1bn plans include shifting charters to Lelystad. Retrieved 16 january, 2014, from http://www.dutchnews.nl/news/archives/2013/05/schiphol_airports_1bn_plans_in.php

Fodness, D., & Murray, B. (2007). Passengers' expectations of airport service quality. *Journal of Services Marketing, 21*(7), 492-506.

Francis, G., Fidato, A., & Humphreys, I. (2003). Airport–airline interaction: the impact of low-cost carriers on two European airports. *Journal of Air Transport Management, 9*(4), 267-273.

Garver, M. S. (2003). Best practices in identifying customer-driven improvement opportunities. *Industrial Marketing Management, 32*(6), 455-466.

Gillen, D. (2011). The evolution of airport ownership and governance. *Journal of Air Transport Management, 17*(1), 3-13.

GoogleMaps. (2013). Web Mapping Service Application. Retrieved 28 August, 2013, from https://maps.google.com/

Harvey, G. (1987). Airport choice in a multiple airport region. *Transportation Research Part A: General, 21*(6), 439-449.

Heathrow. (2013). Heathrow announces £3bn of new private-sector investment in UK infrastructure. Retrieved 24 October, 2013, from http://mediacentre.heathrowairport.com/Press-releases/Heathrow-announces-3bn-of-new-private-sector-investment-in-UK-infrastructure-45b.aspx

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

Hess, S. (2010). Evidence of passenger preferences for specific types of airports. *Journal of Air Transport Management, 16*(4), 191-195.

Hess, S., Adler, T., & Polak, J. W. (2007). Modelling airport and airline choice behaviour with the use of stated preference survey data. *Transportation Research Part E: Logistics and Transportation Review, 43*(3), 221-233.

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

Hess, S., & Polak, J. W. (2006a). Airport, airline and access mode choice in the San Francisco Bay area. *Papers in regional Science, 85*(4), 543-567.

Hess, S., & Polak, J. W. (2006b). Exploring the potential for cross-nesting structures in airport-choice analysis: a case-study of the Greater London area. *Transportation Research Part E: Logistics and Transportation Review, 42*(2), 63-81.

Hess, S., Ryley, T., Davison, L., & Adler, T. (2013). Improving the quality of demand forecasts through cross nested logit: a stated choice case study of airport, airline and access mode choice. *Transportmetrica A: Transport Science, 9*(4), 358-384.

IBM. (2013). SPSS Software: http://www-01.ibm.com/software/nl/analytics/spss/.

KiM. (2013). The social value of shorter and more reliable travel times: Netherlands Institute for Transport Policy Analysis for the Ministry of Infrastructure and the Environment.

Kuo, M. S., & Liang, G. S. (2011). Combining VIKOR with GRA techniques to evaluate service quality of airports under fuzzy environment. *Expert Systems with Applications, 38*(3), 1304-1312.

Lee-Mortimer, A. (1993). Customer focus takes off. *The TQM Magazine, 5*(3).

Lemer, A. C. (1992). Measuring performance of airport passenger terminals. *Transportation Research Part A: Policy and Practice, 26*(1), 37-45.

Lieshout, R. (2012). Measuring the size of an airport's catchment area. *Journal of Transport Geography, 25*, 27-34.

Loo, B. P. (2008). Passengers' airport choice within multi-airport regions (MARs): some insights from a stated preference survey at Hong Kong International Airport. *Journal of Transport Geography, 16*(2), 117-125.

Louviere, J. J., Hensher, D. A., & Swait, J. D. (2000). *Stated choice methods: analysis and applications*: Cambridge University Press.

Luken, B. L., & Garrow, L. A. (2011). Multiairport Choice Models for the New York Metropolitan Area. *Transportation Research Record: Journal of the Transportation Research Board, 2206*(1), 24-31.

Marcucci, E., & Gatta, V. (2011). Regional airport choice: Consumer behaviour and policy implications. *Journal of Transport Geography, 19*(1), 70-84.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior *Frontiers in econometrics* (pp. 105-142). New York: Academic Press.

Pels, E., Nijkamp, P., & Rietveld, P. (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., Nijkamp, P., & Rietveld, P. (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.

Rhoades, D. L., Waguespack Jr, B., & Young, S. (2000). Developing a quality index for US airports. *Managing Service Quality, 10*(4), 257-262.

Rose, J. M., & Bliemer, M. C. (2009). Constructing efficient stated choice experimental designs. *Transport Reviews, 29*(5), 587-617.

Sawtooth. (2013). Sawtooth Software Inc: The survey software of choice: http://www.sawtoothsoftware.com/.

Scarpa, R., Thiene, M., & Train, K. (2008). Utility in willingness to pay space: a tool to address confounding random scale effects in destination choice to the Alps. *American Journal of Agricultural Economics, 90*(4), 994-1010.

Schiphol. (2013). The first phase of the Masterplan has gotten started. Retrieved 24 October, 2013, from http://www.schiphol.nl/Travellers/News/TheFirstPhaseOfTheMasterplanHasGottenStarted.htm

Seneviratne, P. N., & Martel, N. (1991). Variables influencing performance of air terminal buildings. *Transportation Planning and Technology, 16*(1), 3-28.

Skytrax. (2013). World Airport Survey - the Global Survey of Airport Excellence. Retrieved 7 november, 2013, from http://www.worldairportsurvey.com/

Statistics, U. S. B. o. L. (2013). CPI Inflation Calculator. Retrieved 8 january, 2014, from http://www.bls.gov/data/inflation_calculator.htm

Street, D. J., Burgess, L., & Louviere, J. J. (2005). Quick and easy choice sets: constructing optimal and nearly optimal stated choice experiments. *International Journal of Research in Marketing, 22*(4), 459-470.

Taussig, F. W. (1920). Principles of economics: Vol. II.

Train, K. (2009). *Discrete choice methods with simulation*: Cambridge university press.

Train, K., & Weeks, M. (2005). *Discrete choice models in preference space and willingness-to-pay space*: Springer.

Yeh, C. H., & Kuo, Y. L. (2003). Evaluating passenger services of Asia-Pacific international airports. *Transportation Research Part E: Logistics and Transportation Review, 39*(1), 35-48.

Zidarova, E. D., & Zografos, K. G. (2011). Measuring quality of service in airport passenger terminals. *Transportation Research Record: Journal of the Transportation Research Board, 2214*(1), 69-76.

Appendices

Appendix A Distinction of airport service quality dimensions 56
Appendix B Average airport access time and short-haul ticket prices for Dutch passengers 57
Appendix C Average short-haul air ticket prices of Dutch passengers 58
Appendix D Correlation matrices of separate blocks 59
Appendix E Full experimental design 60
Appendix F Variables in the dataset 61
Appendix G Socio-economic characteristics of the sample 62
Appendix H Benefits of effects coding opposed to dummy coding 63
Appendix I Rationale for including insignificant components of effects coded variables 64
Appendix J Complete estimation results for binary logit models 65

Appendix A Distinction of airport service quality dimensions

Service Quality Dimensions	Seneviratne and Martel (1991)	Lemer (1992)	Rhoades et al. (2000)	Yeh and Kuo (2003)	Fodness and Murray (2007)	Kuo and Liang (2011)	ACI (2013)	Skytrax (2013)
Concessions and amenities	Restaurants and shops	Restaurants and bars, choice of things to do	Restaurants and bars, Rest-room facilities, Retail- and duty free shops, Special services	Wash rooms, shops, restaurants, money exchange, cash machines	Shopping eating, productivity facilities	Wash rooms, shops, restaurants, money exchange, cash machines	Restaurants and shopping, washrooms	Airport shopping, food and beverages, rest areas
Circulation	Walking distance, Availability of space, Number of level changes	Walking distance/-time, level changes, difficulty of orientation, signposting, spatial logic, crowding	Inter-terminal transportation	Congestion level	Circulation effectiveness, movement efficiency	Congestion level	Signposting, walking distance	Wayfinding, walking distance
Aesthetics	Internal environment	Sound-, temperature and humidity levels, visual character		Cleanliness, lighting, ambience	Décor	Cleanliness, lighting, ambience	Cleanliness, Ambiance	Cleanliness
Courtesy of staff		Service justice, alternatives in case of delay, service reliability		Courtesy of staff	Interaction	Courtesy of staff, Competence of staff, response to complaints	Helpfulness of staff	Friendliness of staff
Information	Information		Information display	Information visibility			Flight information screens	Flight information screens, flight boarding announcements
Comfort	Availability of seats							Seats available
Waiting- and processing time	Waiting time	Service- and waiting times, reliability of time		Processing time		Processing time	Waiting times, baggage delivery times	Waiting times, baggage delivery times
Convenience	Facilities or devices available to facilitate processing activities			Availability of luggage carts		Availability of luggage carts		Availability of luggage carts
Security				Security		Security		
WIFI / Internet								WIFI and internet facilities
Access / Egress			Parking, ground transportation, Car rental services					
Other		Cost of concessions, recognition of social dimensions, Required time before departure.	Baggage claim, Boarding areas					Availability of smoking areas

Appendix B Average airport access time and short-haul ticket prices for Dutch passengers

Table B.1 Car access times of Dutch airports (in minutes)

Provincial capitals	Province Populations (P_i) *	Amsterdam Airport Schiphol **	Rotterdam The Hague Airport **	Eindhoven Airport **	Maastricht Aachen Airport **	Groningen Airport Eelde **	Average Access Times (D_i)
Haarlem	2728512	21	49	85	129	122	81,2
Rotterdam	3566427	44	12	66	111	138	74,2
Middelburg	381203	112	81	88	110	192	116,6
Utrecht	1247157	39	45	57	101	106	69,6
Den Bosch	2473255	61	58	28	72	128	69,4
Maastricht	1120726	126	119	61	15	180	100,2
Arnhem	2015876	72	75	53	92	98	78
Zwolle	1139284	70	83	82	121	55	82,2
Assen	489529	113	126	125	164	19	109,4
Groningen	580577	110	138	142	181	19	118
Leeuwarden	646262	89	119	135	174	52	113,8
Lelystad	399311	47	75	86	130	81	83,8

* Province population obtained from CBS StatLine (information retrieved on 28-08-2013)

** The car access times from the provincial capitals to the five major domestic airports are obtained from Google Maps (information retrieved on 28-08-2013)

If province population is denoted P_i; and average access times per provincial capital is denoted D_i, then the weighted average access times $\bar{x}$ (weighting factor is the population of the province), is calculated as follows:

$$\bar{x} = \frac{\sum_{i=1}^{n} D_i P_i}{\sum_{i=1}^{n} P_i} \tag{B.1}$$

Using this formula, the weighted average access time of the Dutch population to the domestic airports is 82,28 (≈ 80).

Appendix C Average short-haul air ticket prices of Dutch passengers

For three different departure and return dates, available flights and prices from the five main Dutch domestic airports to five major European capitals are retrieved (table B.1).

Table B.2 Air ticket prices from Dutch airports to five major European capitals

Nr. (n)	City pair	Departure date	Return date	Ticket fare (F_i)
1	Amsterdam - London	7-10-2013	14-10-2013	93,08
2	Rotterdam - London	7-10-2013	14-10-2013	139,10
3	Eindhoven - London	7-10-2013	14-10-2013	40,64
4	Amsterdam - Madrid	7-10-2013	14-10-2013	158,58
5	Rotterdam - Madrid	7-10-2013	14-10-2013	113,00
6	Amsterdam - Milan	7-10-2013	14-10-2013	86,90
7	Groningen - Milan	7-10-2013	14-10-2013	33,28
8	Amsterdam - Berlin	7-10-2013	14-10-2013	107,41
9	Eindhoven - Berlin	7-10-2013	14-10-2013	93,00
10	Amsterdam - Paris	7-10-2013	14-10-2013	98,22
11	Amsterdam - London	20-11-2013	27-11-2013	88,98
12	Amsterdam - Madrid	20-11-2013	27-11-2013	118,58
13	Amsterdam - Milan	20-11-2013	27-11-2013	70,50
14	Amsterdam - Berlin	20-11-2013	27-11-2013	73,58
15	Amsterdam - Paris	20-11-2013	27-11-2013	98,22
16	Rotterdam - London	20-11-2013	27-11-2013	147,79
17	Rotterdam - Madrid	20-11-2013	27-11-2013	83,00
18	Eindhoven - London	20-11-2013	27-11-2013	40,46
19	Eindhoven - Milan	20-11-2013	27-11-2013	92,40
20	Amsterdam - London	6-12-2013	13-12-2013	88,98
21	Amsterdam - Madrid	6-12-2013	13-12-2013	123,58
22	Amsterdam - Milan	6-12-2013	13-12-2013	82,80
23	Amsterdam - Berlin	6-12-2013	13-12-2013	75,63
24	Amsterdam - Paris	6-12-2013	13-12-2013	98,22
25	Rotterdam - London	6-12-2013	13-12-2013	139,10
26	Eindhoven - Madrid	6-12-2013	13-12-2013	83,00

* Ticket fare obtained from cheaptickets.nl (information retrieved on 31-08-2013)

The average air ticket price $\bar{x}$ is obtained by dividing the sum of the fares F_i by the number of n available flights, that is:

$$\bar{x} = \frac{\sum_{i=1}^{n} F_i}{n} \tag{C.1}$$

Using this formula, the average air ticket prices from the Dutch domestic airports to five major European capitals is 94,92 (≈ 95).

Appendix D Correlation matrices of separate blocks

Table D.1 Correlation matrix of version A

	FAR (A_1)	ACC (A_2)	CON (A_3)	CIR (A_4)	AES (A_5)	COU (A_6)	FAR (A_8)	ACC (A_9)	CON (A_{10})	CIR (A_{11})	AES (A_{12})	COU (A_{13})
FAR (A_1)	1											
ACC (A_2)	0	1										
CON (A_3)	0,45	0	1									
CIR (A_4)	0	0	0	1								
AES (A_5)	0	0	0	0	1							
COU (A_6)	0	0,89	0	0	0	1						
FAR (A_8)	-0,4	-0,1	0	-0,67	0,22	-0,22	1					
ACC (A_9)	0,4	0	0,45	-0,22	0	0,22	-0,4	1				
CON (A_{10})	-0,22	0,22	-0,5	0,5	0	0	0	-0,89	1			
CIR (A_{11})	0,22	0	-0,5	-0,5	-0,5	0	0	0	0	1		
AES (A_{12})	0,67	-0,22	0,5	0	0	-0,5	0	0	0	0	1	
COU (A_{13})	-0,22	-0,45	-0,5	0	0,5	-0,5	0	0	0	0	0	1

Table D.2 Correlation matrix of version B

	FAR (A_1)	ACC (A_2)	CON (A_3)	CIR (A_4)	AES (A_5)	COU (A_6)	FAR (A_8)	ACC (A_9)	CON (A_{10})	CIR (A_{11})	AES (A_{12})	COU (A_{13})
FAR (A_1)	1											
ACC (A_2)	0	1										
CON (A_3)	-0,45	0	1									
CIR (A_4)	0	0	0	1								
AES (A_5)	0	0	0	0	1							
COU (A_6)	0	-0,89	0	0	0	1						
FAR (A_8)	-0,7	-0,2	0,22	-0,22	0,22	0	1					
ACC (A_9)	0,2	-0,5	0	-0,45	0,22	0,45	0,4	1				
CON (A_{10})	0,45	-0,45	0	-0,5	0	0,5	0	0,89	1			
CIR (A_{11})	-0,22	-0,22	-0,5	0	0	0,5	0	0	0	1		
AES (A_{12})	0	-0,67	0	0,5	-0,5	0,5	0	0	0	0	1	
COU (A_{13})	0,22	0,45	-0,5	0	-0,5	-0,5	0	0	0	0	0	1

Appendix E Full experimental design

Table D.1 Final experimental design

Choice set #	Alternative 1						Alternative 2					
	FAR (a_1)	ACC (a_2)	CON (a_3)	CIR (a_4)	AES (a_5)	COU (a_6)	FAR (a_7)	ACC (a_8)	CON (a_9)	CIR (a_{10})	AES (a_{11})	COU (a_{12})
1	110 EUR	60 min	Poor	Crowded	Not clean	Friendly	80 EUR	105 min	Great	Not crowded	Not clean	Unfriendly
2	90 EUR	75 min	Poor	Crowded	Clean	Friendly	100 EUR	60 min	Poor	Not crowded	Not clean	Friendly
3	110 EUR	105 min	Poor	Not crowded	Clean	Unfriendly	90 EUR	105 min	Great	Crowded	Not clean	Friendly
4	100 EUR	90 min	Great	Crowded	Clean	Unfriendly	80 EUR	75 min	Poor	Crowded	Clean	Friendly
5	90 EUR	90 min	Poor	Not crowded	Not clean	Unfriendly	110 EUR	90 min	Great	Not crowded	Clean	Friendly
6	80 EUR	60 min	Great	Not crowded	Clean	Friendly	100 EUR	90 min	Great	Crowded	Clean	Unfriendly
7	100 EUR	75 min	Great	Not crowded	Not clean	Friendly	110 EUR	60 min	Poor	Crowded	Not clean	Unfriendly
8	80 EUR	105 min	Great	Crowded	Not clean	Unfriendly	90 EUR	75 min	Poor	Not crowded	Clean	Unfriendly
9	80 EUR	75 min	Poor	Not crowded	Not clean	Unfriendly	110 EUR	105 min	Poor	Crowded	Clean	Friendly
10	100 EUR	60 min	Poor	Not crowded	Clean	Unfriendly	90 EUR	90 min	Poor	Not crowded	Not clean	Friendly
11	110 EUR	90 min	Great	Not crowded	Not clean	Friendly	100 EUR	105 min	Poor	Not crowded	Clean	Unfriendly
12	110 EUR	75 min	Great	Crowded	Clean	Unfriendly	80 EUR	90 min	Poor	Crowded	Not clean	Unfriendly
13	100 EUR	105 min	Poor	Crowded	Not clean	Friendly	80 EUR	60 min	Great	Not crowded	Clean	Friendly
14	80 EUR	90 min	Poor	Crowded	Clean	Friendly	110 EUR	75 min	Great	Not crowded	Not clean	Unfriendly
15	90 EUR	60 min	Great	Crowded	Not clean	Unfriendly	100 EUR	75 min	Great	Crowded	Not clean	Friendly
16	90 EUR	105 min	Great	Not crowded	Clean	Friendly	90 EUR	60 min	Great	Crowded	Clean	Unfriendly

Appendix F Variables in the dataset

Table F.1 Description of the variables in the dataset

Variable	Description
Id	Unique number for each respondent
Age	Respondent's age
Fre	Number of times the respondent used air travel in the past 12 months
Gen	Respondent's gender (1: male, 2:female)
Edu	Respondent's education level (1: Primary, 2: LBO, 3: MAVO, 4: VMBO, 5: MBO, 6: HAVO, 7: VWO, 8:HBO, 9: WO) – Dutch educational system classification
Ver	Experiment version (1, 2)
Set	Choice set (1, 2, … 8)
Av1	Fixed to 1 (means that alternative 1 is available in this choice set)
Av2	Fixed to 1 (means that alternative 2 is available in this choice set)
Far1	Fare level of alternative 1
Acc1	Access time level of alternative 1
Con1	Concessions level of alternative 1
Cir1	Circulation level of alternative 1
Aes1	Aesthetics level of alternative 1
Cou1	Courtesy level of alternative 1
Far2	Fare level of alternative 2
Acc2	Access time level of alternative 2
Con2	Concessions level of alternative 2
Cir2	Circulation level of alternative 2
Aes2	Aesthetics level of alternative 2
Cou2	Courtesy level of alternative 2
Choice	Respondent's choice (1: first alternative, 2: second alternative)

Appendix G Socio-economic characteristics of the sample

Table G.1 Descriptive statistics of the socio-economic characteristics of the sample

Variable	Median	Mean	St. Dev.	Min	Max
Age	43	41,639	14,361	18	74
Gender	-	1,438	0,497	1	2
Air travel frequency	1	1,761	1,873	0	40
Education level	8	6,971	1,978	1	9

Figure G.1 Histogram for age of the respondents

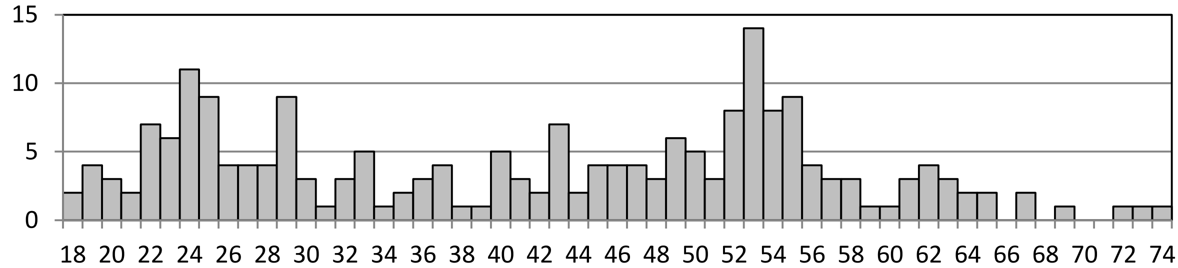


Figure G.2 Histogram for gender of the respondents

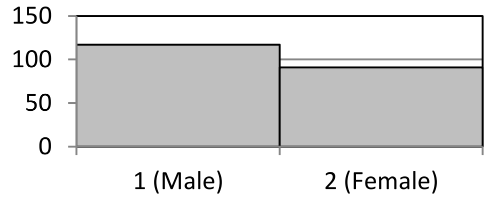


Figure G.2 Histogram for air travel frequency of the respondents

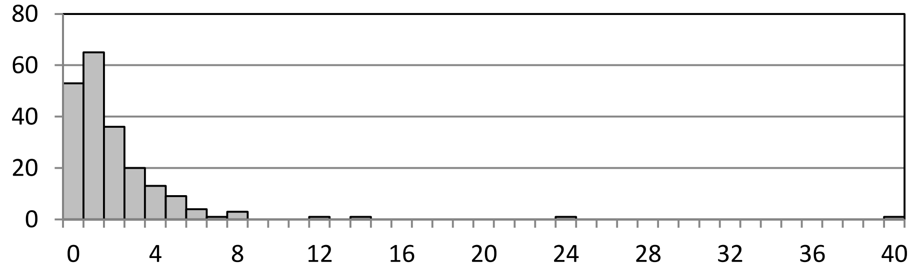
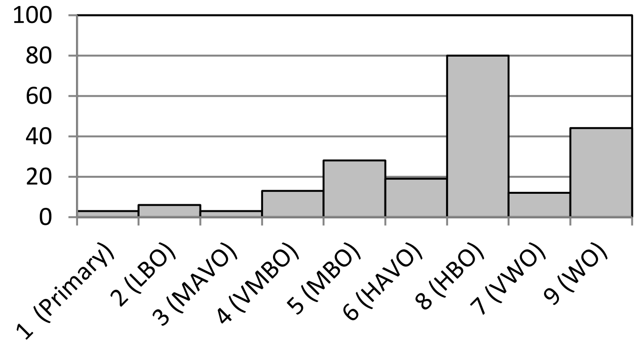


Figure G.4 Histogram for education level of the respondents



Appendix H Benefits of effects coding opposed to dummy coding

Here we explain the rationale for using effects coding as opposed to dummy coding. The explanation is based on Hensher et al. (2005, pp. 119-120).

For simplicity's sake, assume that the choice for an airport depends only on the fare level of the airports FAR_i. Furthermore, the researcher wants to control for the interaction effect of the gender of the passenger with the fare level $FAR_i \cdot GEN_q$. Then the systematic utility Vpassenger q derives from an airport i, is specified as follows:

$$V_{iq} = \beta FAR_i + \gamma FAR_i \cdot GEN_q \tag{H.1}$$

First suppose the researcher *dummy codes* the gender variable such that GEN_q takes the value of 1 if the passenger is a male and the value of 0 if the passenger is a female. Subsequently, the systematic utility function of a male passenger is:

$$V_{iq} = \beta FAR_i + \gamma FAR_i \cdot 1 = \beta FAR_i + \gamma FAR_i \tag{H.2}$$

and the systematic utility function of a female passenger is:

$$V_{iq} = \beta FAR_i + \gamma FAR_i \cdot 0 \; = \; \beta FAR_i \tag{H.3}$$

Examination of Eq. (H.3) shows that the fare coefficient β is not measuring a mean fare coefficient (irrespective of which gender) but actually measures the fare coefficient for females. Hence, by using dummy coding structures the researcher perfectly confounds the main effect with one of the gender categories (*i.e.* females).

Now suppose the researcher *effects codes* the gender variable such that GEN_q takes the value of 1 if the passenger is a male and the value of -1 if the passenger is a female. Hence, the systematic utility function of a male passenger is:

$$V_{iq} = \beta FAR_i + \gamma FAR_i \cdot 1 \tag{H.4}$$

and the systematic utility function of a female passenger is:

$$V_{iq} = \beta FAR_i + \gamma FAR_i \cdot -1 \tag{H.5}$$

Note that the fare coefficient β is no longer perfectly confounded with a particular category. Instead β reflects an *overall mean effect*, with $\beta \pm \gamma$ reflecting the effects of specific categories.

Appendix I Rationale for including insignificant components of effects coded variables

Here we explain the rationale for including insignificant components of effects coded variables. The explanation is based on Hensher et al. (2005, p. 351).

For simplicity's sake, assume that the choice for an airport depends only on the fare level of the airports FAR_i. Moreover, the researcher controls for the interaction effect of the education level of the passenger with fare level. The researcher specifies three categories of education level (low, middle, high) and uses the effects coding structure we used in this thesis (see table 5.2). Then the systematic utility Vpassenger q derives from an airport i, is specified as follows:

$$V_{iq} = \beta FAR_i + \gamma_1 FAR_i \cdot EDUh_q + \gamma_2 FAR_i \cdot EDUm_q \tag{I.1}$$

Subsequently, the systematic utility function of a passenger of the higher education category is:

$$V_{iq} = \beta FAR_i + \gamma_1 FAR_i \cdot 1 + \gamma_2 FAR_i \cdot 0 = \beta FAR_i + \gamma_1 FAR_i \tag{I.2}$$

the systematic utility function of a passenger belonging to the middle education category is:

$$V_{iq} = \beta FAR_i + \gamma_1 FAR_i \cdot 0 + \gamma_2 FAR_i \cdot 1 = \beta FAR_i + \gamma_2 FAR_i \tag{I.3}$$

and the systematic utility function of a passenger belonging to the middle education category is:

$$V_{iq} = \beta FAR_i + \gamma_1 FAR_i \cdot -1 + \gamma_2 FAR_i \cdot -1 = \beta FAR_i - \gamma_1 FAR_i - \gamma_2 FAR_i \tag{I.4}$$

Now suppose that after initial estimation the interaction effect between fare- and the higher education category $FAR_i \cdot EDUh_q$ is significant, while the interaction effect between fare- and the middle education category $FAR_i \cdot EDUm_q$ is not significant. Hence, the researcher decides to exclude the latter interaction effect and re-estimates the model. Note that this does not change the systematic utility function of a passenger belonging to the higher education category. However, the systematic utility function of a passenger belonging to the middle education level changes into:

$$V_{iq} = \beta FAR_i + \gamma_1 FAR_i \cdot 0 = \beta FAR_i \tag{I.5}$$

Examination of Eq. (I.5) shows that – by excluding the interaction between fare- and middle education category $FAR_i \cdot EDUm_q$ – the fare coefficient β is perfectly confounded with the middle education category. If the researcher is interested in estimating *overall mean effects* (such as we are), then this is undesirable and hence the researcher should still include the insignificant interaction effect between fare- and middle education category.

Appendix J Complete estimation results for binary logit models

Table J.1 Complete estimation results for base model

	BL_{base}			
	Parameter	Robust s.e.	Robust T-test	p-value
Marg. utility of income (δ)	0.0816	0.0049	16.59	0
Access time (min.)	-0.5600	0.0350	-15.81	0
Great availability of restaurants and shops (1,0)	2.39[a]	1.0600	2.27	0.02
Uncrowded (1,0)	2.17[b]	1.1400	1.9	0.06
Clean (1,0)	13.70	1.0400	13.14	0
Friendly staff (1,0)	11.40	0.9710	11.78	0
Incl. socio-economics	NO			
Observations	1664			
Parameters	6			
Final LL	-766.524			
Rho-square	0.335			
Adj. rho-square	0.330			

Table J.2 Complete estimation results for full model

	BL_{full}			
	Parameter	Robust s.e.	Robust T-test	p-value
Marg. utility of income (δ)	0.0961	0.0065	14.81	0
Access time (min.)	-0.5440	0.0355	-15.32	0
Great availability of restaurants and shops (1,0)	3.57	1.0600	3.37	0
Uncrowded (1,0)	1.49[c]	1.0800	1.38	0.17
Clean (1,0)	12.50	1.0200	12.29	0
Friendly staff (1,0)	9.44	0.9680	9.75	0
Fare * age	0.0093	0.0065	2.3	0.02
Fare * air travel frequency	0.0413	0.0355	3.66	0
Fare * gender	-0.0971	1.0600	-1.84	0.07
Fare *middle education level	0.1240	1.0800	1.75	0.08
Fare * higher education level	-0.2360	1.0200	-2.13	0.03
Access time * age	0.0108	0.9680	4.23	0
Access time * air travel frequency	-0.0200	0.0041	-2.07	0.04
Access time * gender	-0.0883	0.0113	-2.63	0.01
Access time * middle education level	0.0413	0.0529	0.89	0.38
Access time * higher education level	-0.1080	0.0705	-1.61	0.11
Availability of restaurants and shops * age	-0.15	0.1110	-2.04	0.04
Availability of restaurants and shops * air travel frequency	-0.95	0.0026	-2.37	0.02
Availability of restaurants and shops * gender	0.46	0.0097	0.48	0.63

Availability of restaurants and shops * middle education level	-0.95	0.0336	-0.74	0.46
Availability of restaurants and shops * higher education level	1.89	0.0466	1.15	0.25
Uncrowded * age	0.04	0.0671	0.54	0.59
Uncrowded * air travel frequency	0.46	0.0725	1.32	0.19
Uncrowded * gender	-0.20	0.4020	-0.19	0.85
Uncrowded * middle education level	0.39	0.9580	0.29	0.77
Uncrowded * higher education level	-1.90	1.2800	-1.05	0.29
Clean * age	0.13	1.6400	1.65	0.10
Clean * air travel frequency	-0.23	0.0764	-0.94	0.35
Clean * gender	-0.16	0.3490	-0.16	0.88
Clean * middle education level	-1.37	1.0100	-1.01	0.31
Clean * higher education level	1.02	1.3400	0.54	0.59
Friendly staff * age	0.10	1.8100	1.41	0.16
Friendly staff * air travel frequency	0.60	0.0754	2.25	0.02
Friendly staff * gender	-1.03	0.2440	-1.09	0.27
Friendly staff * middle education level	-0.63	1.0200	-0.5	0.62
Friendly staff * higher education level	-1.99	1.3600	-1.09	0.28
Incl. socio-economics	ALL			
Observations	1664			
Parameters	36			
Final LL	-721.686			
Rho-square	0.374			
Adj. rho-square	0.343			

Table J.3 Complete estimation results for final model

	BL_{final}			
	Parameter	Robust s.e.	Robust T-test	p-value
Marg. utility of income (δ)	0.0950	0.0061	15.55	0
Access time (min.)	-0.5450	0.0344	-15.82	0
Great availability of restaurants and shops (1,0)	3.10	1.0000	3.1	0
Uncrowded (1,0)	1.83[d]	1.0500	1.74	0.08
Clean (1,0)	12.20	0.9660	12.67	0
Friendly staff (1,0)	9.67	0.9150	10.57	0
Fare * age	0.0086	0.0037	2.37	0.02
Fare * air travel frequency	0.0351	0.0128	2.73	0.01
Fare * gender	-0.1160	0.0430	-2.69	0.01
Fare *middle education level	0.0743	0.0578	1.29	0.2
Fare * higher education level	-0.2310	0.0813	-2.84	0
Access time * age	0.0103	0.0024	4.29	0
Access time * air travel frequency	-0.0223	0.0098	-2.27	0.02
Access time * gender	-0.0966	0.0310	-3.12	0
Access time * middle education level	0.0236	0.0422	0.56	0.58

Access time * higher education level	-0.1180	0.0567	-2.08	0.04
Availability of restaurants and shops * age	-0.16	0.0653	-2.46	0.01
Availability of restaurants and shops * air travel frequency	-0.86	0.3960	-2.16	0.03
Uncrowded * air travel frequency	0.46	0.3220	1.44	0.15
Clean * age	0.13	0.0692	1.84	0.07
Friendly staff * age	0.13	0.0625	2.02	0.04
Friendly staff * air travel frequency	0.66	0.2810	2.35	0.02
Incl. socio-economics	SELECTED			
Observations	1664			
Parameters	22			
Final LL	-725.868			
Rho-square	0.371			
Adj. rho-square	0.352			