Incorporating Attitudes in Airline Itinerary Choice: Modeling the Impact of Elapsed Time

by
Georg Wilhelm Theis

Diplom-Volkswirt
Universität Münster, Germany (1997)

Submitted to the Department of Civil and Environmental Engineering in Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the field of

Transportation Systems and Decision Sciences

at the
Massachusetts Institute of Technology

June 2011

Signature of Author ....
Department of Civil and Environmental Engineering
May 20, 2011

Certified by ....
Moshe Ben-Akiva
Edmund K. Turner Professor of Civil and Environmental Engineering
Thesis Co-Supervisor

Certified by ....
John-Paul Clarke
Associate Professor, School of Aerospace Engineering and
School of Industrial and Systems Engineering, at Georgia Institute of Technology
Thesis Co-Supervisor

Accepted by ....
Heidi M. Nepf
Chair, Departmental Committee for Graduate Students

THIS PAGE INTENTIONALLY LEFT BLANK

Incorporating Attitudes in Airline Itinerary Choice: Modeling the Impact of Elapsed Time

by
Georg Theis

Submitted to the Department of Civil and Environmental Engineering on May 20, 2011
in Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in

Transportation Systems and Decision Sciences

Abstract

Network airlines traditionally try to minimize the elapsed time of their passengers in connecting travel, based on the assumption that longer elapsed times would make their itinerary less competitive and thus reduce their revenue potential in a given origin destination market.

Contrary to the traditional assumption, we hypothesize that passengers assign a lower utility to connections close to the minimum connecting time published by airports. We hypothesize that there are three factors related to connection time that a customer might consider when choosing an itinerary: The risk of misconnection, the discomfort associated with rushing through an airport terminal, and the trust in an airline to provide reliable connections. Attitudes toward these constructs are latent and cannot be directly observed. In aggregate, we expect an n-shaped utility function for the time additional to the minimum connecting time, with increasing utility close to the minimum connecting time, followed by a time window of indifference, followed by decreasing utility due to value of time aspects.

We first present a case study using airline booking data that shows that up to 25% of passengers in a sample market voluntarily choose a longer connection, all else equal. In the subsequent chapters, a model to evaluate systematically the impact of length of connecting time is developed. We extend an existing airline survey to incorporate this question. A stated preference experiment is designed and conducted to collect choice data. Psychometric indicators are used to capture attitudes that are explained with socio-demographics and trip characteristics in a multiple indicator-multiple causes (MIMC) model. The MIMC model is then combined with the choice model to simultaneously estimate an integrated choice and latent variable model, quantifying the interactions of latent attitudes and connecting time.

The results demonstrate the non-monotonicity of connecting time utility and the disutility associated with short connecting times. The inclusion of the latent variables risk, rush and trust demonstrates the effect of these constructs on choice. Individuals who are risk-averse, rush averse or have a low level of trust into airlines' scheduling reliability have a higher utility for slightly longer connecting times. This is the first research in the airline choice literature to demonstrate the nonlinearity of connecting time utility. It is also the first research to include attitudes into itinerary choice models, thereby providing a richer explanation for passenger airline choice. Airlines can use the findings to better align their service offerings with their customers' preferences and at the same time reduce their operational costs.

Thesis Supervisor: Moshe E. Ben-Akiva
Title: Edmund K. Turner Professor of Civil and Environmental Engineering

Thesis Supervisor: John-Paul Clarke
Title: Associate Professor, School of Aerospace Engineering and School of Industrial Engineering and System Engineering at Georgia Institute of Technology

Acknowledgements

First, I would like to thank my advisors, Moshe Ben-Akiva and John-Paul Clarke, for their invaluable guidance and continuous support. My thanks go to Professor Ben-Akiva for encouraging me to pursue this research and for the thorough methodological and practical advice in all areas of life, including the fellow Dad conversations. I am indebted to Professor Clarke for giving me the freedom to define my area of research while offering guidance throughout the process.

I would also like to thank my other committee members. Cynthia Barnhart deserves my thanks for the warm welcome at the outset of my studies and for being available whenever I needed her support. I thank Peter Belobaba for excellent classes on airline management and for providing the insight into the airline distribution world, and Joan Walker for her methodological insights, her guidance during various modeling stages and for her encouragement during my final year at MIT.

Tom Adler and Rhona Dallison of Resource Systems Group were instrumental in setting up the survey and getting the data. Thank you also to Lutz Jungnickel of Deutsche Lufthansa AG for his help and advice setting up the case study. Many thanks to Denis Bolduc for supplying the estimation software and for his support.

I would also like to thank other faculty that made my stay at MIT an enjoyable one, among them Joe Coughlin, Charles Fine, Amadeo Odoni, Karen Polenske, Fred Salvucci, Joseph Sussman, Ian Waitz and Nigel Wilson.

I feel lucky to have had the chance to have engaging intellectual discussions with many fellow students at the Center for Transportation and Logistics and the International Center for Air Transportation. These friends and colleagues include Alexandra Mozdanowska, Alf Kohler, Chris Tsonis, Diana Dorinson, Emmanuel Carrier, Flora Garcia, Francis Carr, Hayley D. Reynolds, Jonathan Histon, Laurence Vigeant-Langlois, Maital Dar, Miguel Molina, Robin Riedel, Ryan Tam, Shiro Yamanaka, Tom Gorin, Tom Reynolds and Yasmine El-Alj.

I would like to thank the sponsors of my research, Nextor (FAA), the Cambridge-MIT Institute and the Sloan Foundation, which awarded me a fellowship for the most part of my time at MIT.

Finally, I would like to thank my family. My parents, for laying the foundations of the opportunities I am enjoying. My wife Jessica, for her willingness to accept tremendous opportunity costs during our time in Cambridge and for supporting my hedonistic endeavors in every possible way. My children, Nia and Simon, for every smile and the grounding they give me. All of you are the joy in my life.

THIS PAGE INTENTIONALLY LEFT BLANK

Table of Contents

THIS PAGE INTENTIONALLY LEFT BLANK

List of Figures

List of Tables

THIS PAGE INTENTIONALLY LEFT BLANK

Glossary

The following airline travel and operations related terms are used in the text according to the definitions outlined below.

Bank A bank is a time window in an airline's hub timetable with clustered arrivals and departures, aiming to insure short connections for passengers.

Block time The block time of a flight leg is the scheduled time from gate to gate, including time spent rolling from the origin gate to the origin airport's runway, time airborne, and time rolling from the destination airport's runway to the respective gate.

Buffer time Buffer time of a connection is the difference between minimum connecting time and the connecting time of that itinerary.

Connecting time The connecting time of an itinerary is the scheduled time between arrival of the first flight and scheduled departure of the subsequent flight.

Connection A passenger has a connection if she changes airplanes within her itinerary.

Elapsed time The elapsed time of an itinerary is the time between the scheduled departure of the first flight and the scheduled arrival of the last flight.

Flight A flight in passenger service is a travel from A to B without having to leave the plane. In the majority of cases, a flight is nonstop. However, flights can also have stops in between, with some passengers deboarding and others boarding the plane.

Flight leg A flight leg is a flight consisting of a single departure and a single arrival.

Hub A major airport of an airline that is used for passenger connections.

Itinerary An itinerary is a series of flights from a passenger's origin to her destination. An itinerary may contain only one flight or more than one flight.

Minimum connecting time The minimum connecting time (MCT) is the minimum feasible connecting time at an airport as published by the corresponding airport authority.

Path quality The path quality of an itinerary considers the number of stops or connections in that itinerary. A nonstop itinerary is considered to have a higher path quality than an itinerary with stops or connections.

Stop A stop is a landing and departure within a flight. Passengers do not have to leave the plane.

1 Introduction

1.1 Motivation

The business environment for U.S. network airlines has changed considerably over the past fifteen years. Whereas network airlines were able to yield record profits in the late nineties, they incurred record losses after September 11, 2001.

Since 2001, four of the six largest U.S. network carriers (Delta Air Lines, Northwest Airlines, US Airways, United Airlines) entered bankruptcy protection, and the last two (Delta and Northwest) emerged from bankruptcy in early 2007. Only American Airlines and Continental Airlines were able to avoid bankruptcy. During the bankruptcy-restructuring phase, airlines try to reduce their costs radically in order to regain their competitiveness. At the same time, airlines that are not under Chapter 11 protection have to follow the cost cutting lead in order not to fall behind. Airlines use bankruptcy protection primarily to reduce their obligations toward employees and aircraft leasing companies. However, they also investigate other unit cost reducing measures, such as aircraft and other resource utilization and operational and commercial procedures.

As part of an effort to reduce overall costs, several network airlines reconsidered their hub timetable design. Timetable design, i.e. the setting of departure times for a given set of flights, is one of the critical airline management decisions. On the one hand, airlines are under pressure to increase the utilization of their resources. Any idle time of airplanes, gates, ground equipment or personnel is costly to airlines. On the other hand, airlines have to observe the timetable's attractiveness to their customers. In order to sell their tickets, airlines need to assess the needs of their customers and take them into consideration when designing their timetables.

The recent discussion about "depeaking" of network airlines' timetables is an example of such a tradeoff (Flint, 2002). Network airlines' hub timetables traditionally were planned with peaks of departures and arrivals at certain times of day ("banks") because the airlines tried to keep the connecting times short for their connecting passengers. This was based on the assumption that passengers have a strong preference to minimize the scheduled elapsed time of their trip. However, these peaks result in high operational costs for the airline, the airport and air traffic control. Resources, such as ground equipment and ground personnel, have to be supplied for the peak level and are underutilized in off peak times. In addition, aircraft utilization suffers from banks in schedules: In order for an airplane to arrive within a bank, aircraft often have to be on the ground longer at spoke stations closer to the hub than the required minimum ground time. Empirical evidence also suggests that block times, i.e. the scheduled time of a flight leg from gate to gate, are longer into and out of banks versus off-bank flight legs. This occurs mainly because of congestion effects (Mayer & Sinai 2003).

Consequently, some network airlines have started to "depeak" their hub schedules. American Airlines reported that this led to a savings of five aircraft (Airline Business, 2002). At the same time, airlines have reported that depeaking leads to an increase in average connecting times for their passengers and thus an increase in itinerary elapsed time. American Airlines reports an increase in average connecting times of 7 minutes (Flint, 2002). Consequently, some network airlines are reluctant to depeak further since they fear that an increase in scheduled connecting time would have a detrimental effect on their market shares. This is based on the assumption that passengers would avoid itineraries with longer connecting times.

While it is fairly simple to calculate the operational efficiency gains of depeaking in terms of reduced fleet size, the effect on the demand side is much harder to estimate. How does a change in elapsed time impact passenger itinerary choice?

The traditional assumption in the literature is that an itinerary in a specific origin-destination market is less attractive with increased connecting times (e.g. Holloway, 2003). Mayer & Sinai (2003) point out that hub airlines want to "minimize passenger travel time spent ... waiting for flight connections". Mayer & Sinai (2003) also postulate "since passengers prefer shorter connections, longer connection times reduce the fares an airline can collect". The traditional assumption that passengers discount longer connections is mirrored in many airlines' network planning tools, market share models and hub optimizers (Schröder, 2004). Traditional reasons for minimizing elapsed time are explained in more detail in section 2.3.

However, the growth in air traffic might have had an impact on this traditional assumption. While airport terminals were expanded, published minimum connecting times of airports, i.e. the minimum time that is feasible for a connection according to the airport authority, have often stayed the same, leading to potential rush when connecting. At the same time, growth in air traffic in combination with limited system capacity has lead to an increasing variability in arrival times of flights, i.e. diversions from schedule. Since delays of inbound flights reduce passengers' actual connecting times, the risk of misconnection has increased.

As a result of these observations, we hypothesize that some passengers might prefer scheduled connecting times *higher* than the published minimum connecting time in order to avoid rush and in order to reduce the risk of misconnection. In this context, we hypothesize that there are three not directly observable (latent) variables related to connection time that a customer might consider when choosing an itinerary: his attitude toward the risk of misconnection, his attitude toward the potential discomfort associated with rushing through an airport terminal, and his trust in airlines to provide reliable connections. In aggregate, we expect an n-shaped utility function for scheduled connecting time, having a lower utility close to the minimum connecting time as published by the connecting airport, followed by a time window of relative indifference, followed by decreasing utility due to value of time aspects.

Research in passenger choice models has until now not covered the connecting process in detail. Some choice models use "elapsed time" and "number of connections" as attributes (see Chapter 3). However, none of the choice models focus on the effect different connecting time lengths have on passenger airline choice.

In addition, traditional models do not allow for the inclusion of passengers' attitudes in the choice model. Attitudes are a reflection of a passenger's values and tastes that influence his choice behavior (Polydoropoulou, 1997). The effects of attitudes toward rush or the risk of misconnection may influence the choice of itinerary and should thus be integrated into the passenger itinerary choice model.

Models that incorporate these factors and focus on connecting time can support a broad range of managerial decisions. If an airline knew that there is a time window of indifference toward connecting times, it could use the upper bound in hub timetable design and thereby increase the efficiency of its operations and the profitability of the airline. Identifying which passengers are more likely to be indifferent toward the length of connecting times would allow airlines to differentiate their service or reduce the number of passengers with short connections since these connections are prone to irregularities.

1.2 Research Objective and Approach

The purpose of this dissertation is to study whether minimizing elapsed time and thus minimizing connecting times maximizes passengers' utilities. This is the traditional assumption found in the literature. Contrary to this assumption, we hypothesize that the risk of misconnection, rush aversion and a low level of trust into airlines' scheduling reliability may lead to a non-monotonic curve of connecting time utility. In order to study this question, three types of data are used: Airline booking data, answers to a rating exercise, and a stated preference choice experiment.

Based on airline booking data, we analyze what share of passengers avoids short connections. Comparing two itineraries in the same origin destination market that are available on the day of purchase and that only differ in connecting time yields the share of passengers that voluntarily book a longer connection. Passengers who choose the longer connecting time, else equal, attain a higher utility due to the longer connecting time.

As part of a survey on airline itinerary choice, passengers are asked to rate statements regarding their connecting preferences on a scale from strongly disagree to strongly agree. For instance, passengers are asked whether they would avoid short connections due to the risk of misconnection. The answers give an indication whether or not minimizing elapsed time is generally preferred.

Based on the stated preference experiment, we test in a choice model if connecting time utility is non-monotonic. Airline itinerary choice is modeled as a function of e.g. fare, frequent flyer status, preferred airline, number of connections, elapsed time, and the buffer time at a connecting airport. Including the rating exercise responses and using them as manifestations of the latent variables allows the estimation of an integrated choice and latent variable model, thereby using attitudes to enhance the explanatory power of the choice model. The purpose here is to explain the taste variation regarding length of connecting time with the latent constructs of risk, rush and trust.

1.3 Summary of Findings and Contributions

Based on airline booking data, it is demonstrated that up to 25% of passengers book longer connections voluntarily for a sample origin destination market. This demonstrates that the phenomenon of avoiding short connecting times exists in a real market situation.

Responses to the rating exercise demonstrate that risk of misconnection, rush and airlines' scheduling reliability are taken into account by passengers when choosing their airline itinerary. 56% of the survey population somewhat or strongly agree with the statement "I try to avoid short connections due to the risk of misconnections", 59% of passengers somewhat or strongly agree with the statement "I like to take my time when connecting between flights", and 67% of respondents somewhat or strongly agree with the statement "Airlines sometimes underestimate the time needed to connect between flights". The answers to the rating exercise show that for a considerable share of passengers, minimizing elapsed time does not automatically maximize their utility.

The disutility of short connecting times and the nonlinearity of preferences regarding connecting time is demonstrated based on stated preference data. Passengers in aggregate have increasing utility when adding minutes to the connecting time close to the minimum connecting time as published by the airport authority. Increasing the connecting time further beyond a threshold value reduces the utility, leading to a non-monotonic curve of connecting time utility.

The behavior towards connecting times shows a significant amount of taste heterogeneity. The answers to the rating exercise serve as indicators for the latent constructs "risk", "rush" and "trust". These latent variables are quantified and explained with socio-demographic characteristics of the individuals. By integrating the latent variables into the choice model, it is demonstrated that these attitudes impact airline itinerary choice.

The methodological contribution is the development and implementation of a comprehensive tool to evaluate the impact of length of connecting time. A stated preference experiment was designed and executed to incorporate these effects. In addition, a rating exercise was designed to capture latent attitudes toward rush, risk and trust into airlines' scheduling reliability. Attitudes were quantified in a latent variable model and integrated in a choice and latent variable model, explaining preference heterogeneity by interacting latent variables with attributes of the alternatives.

The choice model results show that the utility of an itinerary is increased with a lower fare, when the individually preferred airline is offered, when the passenger is an elite member of the airline's frequent flyer program, with low number of connections, and low elapsed time. The results also show, however, that the utility of an itinerary is increased by adding minutes to the minimum connecting time when connecting. When the latent constructs risk tolerance, rush aversion and trust into airlines' scheduling reliability are explained with socio-demographics, results indicate that e.g. women are less risk tolerant and more rush averse than men in airline itinerary choice, that students are more risk tolerant than other employee groups, and that customers checking bags are less risk tolerant and more rush averse than those not checking bags. When the latent variables are integrated into the choice model, it is shown that they can provide a richer explanation of choice. Passengers with lower risk tolerance, higher rush aversion or lower trust into airlines' scheduling reliability have a higher utility of added minutes to the minimum connecting time of airports.

The results of this research refute the traditional assumption found in the literature that minimizing elapsed time maximizes passengers' utilities. In addition to the traditional "value of time" argument, it is demonstrated that passengers' attitudes toward the risk of misconnection, attitudes toward rush and passengers' trust into airlines' scheduling reliability play an important part in evaluating elapsed time. Including these constructs provides for a richer and more realistic explanation of airline itinerary choice.

1.4 Thesis Organization

This thesis is organized into seven chapters, including this introduction. Chapter 2 presents the relevance of elapsed time in a historical perspective and presents a case study using airline booking data. Chapter 3 reviews the literature in airline itinerary choice models. Chapter 4 introduces the analytical framework that is used in the subsequent analysis. It presents the general random utility framework, the behavioral assumptions and hypotheses and it describes the joint estimation approach incorporating choice model and latent variables used in this study. Chapter 5 gives an overview of the empirical data used in this study. It describes the survey approach, design and execution and summarizes the data. Chapter 6 reports the specification and estimation results of the model based on the previously described analytical framework and discusses the implications of the results. Finally, Chapter 7 presents the research's contributions, summarizes the findings and presents suggestions for future research.

2 Passenger Demand for Air Travel and Elapsed Time

2.1 Introduction

This chapter consists of three sections. In the first section, a brief historical background of airline distribution relevant to the discussion of elapsed time is given. In the second section, the reasons for the attributed importance of elapsed time in airline itinerary choice are presented and discussed. In the third section, a case study is presented. In the case study, booking data from a sample market is used to analyze what share of passengers voluntarily chooses a longer elapsed time for their itinerary.

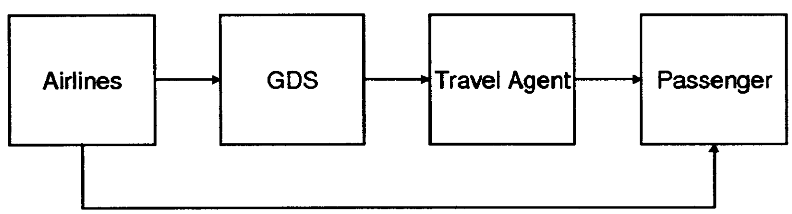
2.2 Historical Background

Since the 1950's, airline travel has seen tremendous growth rates. In the early days of air transportation, airlines operated under a strictly regulated environment. Routes and fares had to be applied for and approved by government authorities (see e.g. Goetz & Sutton, 1997). In this environment, airlines cooperated via interline agreements, honoring each other's tickets, since no one airline was able to cover all traveling needs for an individual. Because of this regulated environment with fixed fares and relatively low supply in comparison to today, the schedule of an airline was a key differentiating factor.

The deregulation of fares in the U.S. market brought about by the "United States Airline Deregulation Act of 1978" changed the relative importance of schedule by introducing fare differentiation into the market. With fares and restrictions potentially being different for two airlines, passengers could now tradeoff these attributes versus schedule quality. In the early 1980's, another key differentiator was introduced with the frequent flyer programs. Frequent flyer programs are a powerful customer retention program, giving incentives in form of elite status to perform most travel with one airline. In consequence, customers potentially moved from a "search by schedule" (with all else being equal), to a customer choice that involved two more powerful attributes, that is price and frequent flyer programs.

Distribution of airline tickets has changed considerably since the 1950's as well. At the beginning of mass air travel, airlines overwhelmingly sold tickets by phone or in their own sales offices. Passengers could call the airline or travel agents would call the airline to book tickets for their customers. With increasing automation, airlines offered travel agents booking engines, i.e. Computer Reservation Systems (CRS). With added functionality, these systems today are called "Global Distribution Systems" (GDS). In the traditional distribution channel setup (see Figure 2-1), Global Distribution Systems presented a bottleneck with the potential of abuse of power. Travel agents would almost exclusively rely on GDS to access an airline's inventory, giving the GDS in an unregulated environment the chance to abuse their power. Passengers could either call the airlines directly or, in order to reduce their transaction costs, rely on a travel agent.

Figure 2-1: Traditional Distribution Channels

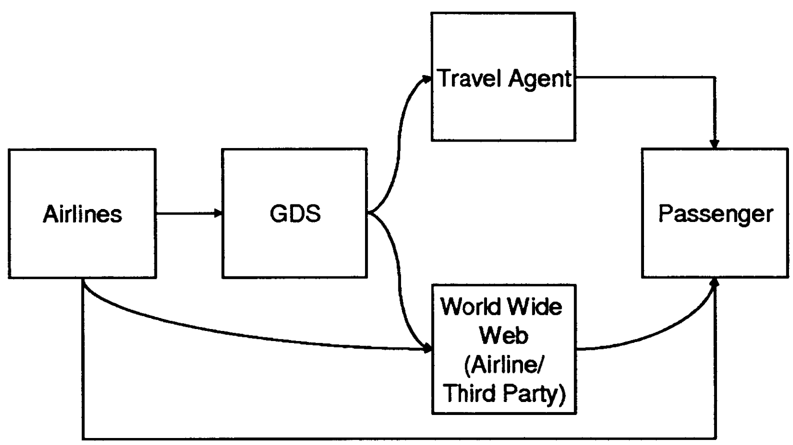


Source: adapted from Taubmann (2004)

In order to avoid that airlines owning GDS could exploit the system by giving their own itineraries preferred treatment, the U.S. government regulated the GDS. A "neutral" display was required, showing itineraries ordered by certain schedule criteria. The regulation also assured that all users had equal access to airline inventory. If an airline owning a GDS displayed itineraries in one GDS, it had to display it in all others as well. As a consequence of this regulation, GDS owner airlines couldn't force travel agents to subscribe to their own GDS if they wanted access to that airline's inventory. In addition, this led to transparency of market supply for the travel agent in one single transaction.

With the advent of the World Wide Web, the distribution of airline tickets has become much more diverse (see Figure 2-2). Airlines and passengers can now bypass the Global Distribution Systems by using airline or third party internet sites, making the GDS' position a much weaker one. At the same time, the market loses much of its former transparency, since not all airlines offer their itineraries in all systems. Passengers and travel agents incur much higher transaction costs if they want full information regarding supply today.

Figure 2-2: Modern Distribution Channels



Source: adapted from Taubmann (2004)

Both the changes following deregulation and the changes in distribution channels have an effect on the reasons for the attributed importance of elapsed time, which will be discussed in the following section.

2.3 Reasons for Attributed Importance to Elapsed Time

The importance of elapsed time is historically attributed to three factors: The screen position of an itinerary in global distribution systems (GDS), the decision window model (Soncrant & Hopperstad, 1998) and passenger preferences regarding lengths of trips and connections. Since the airline industry has changed dramatically in the last few years, some of the changes may lead to the conclusion that "elapsed time" has lost relative importance in passenger airline choice.

2.3.1 GDS Screen Position

The most prominent hypothesis about the importance of elapsed time is based on the GDS screen position. Global distribution systems have until recently been required by government regulation to supply a "neutral" display that shows itineraries in a certain order. The U.S. regulation called for elapsed time to be a "significant factor" (Department of Transportation, 2003):

> "(a2) Each integrated display offered by a system must either use elapsed time as a significant factor in selecting service options from the database or give single-plane flights a preference over connecting services in ranking services in displays.
>
> (b1) Systems may order the display of information on the basis of any service criteria that do not reflect carrier identity and that are consistently applied to all carriers, including each system owner, and to all markets."

Following this rule, several systems use path quality as their primary criterion of ranking, i.e. nonstop flights before direct flights (involving a stop without a plane change) and then connecting flights. Others use elapsed time directly or "displacement time", i.e. the deviation from the time requested by the system user (Department of Transportation, 2004).

The European Union regulation specifically requires connecting itineraries to be ranked by elapsed journey time (European Council, 1999):

> "1. Ranking of flight options in a principal display, for the day or days requested, must be in the following order unless requested in a different way by a consumer for an individual transaction:
>
> (i) all non-stop direct flights between the city-pairs concerned;
> (ii) all other direct flights, not involving a change of aircraft or train, between the city pairs concerned;
> (iii) connecting flights.

2. A consumer must at least be afforded the possibility of having, on request, a principal display ranked by departure or arrival time and/or elapsed journey time. Unless otherwise requested by a consumer, a principal display must be ranked by departure time for group (i) and elapsed journey time for groups (ii) and (iii)."

Figure 2-3 shows an example of a GDS screen for the itinerary Boston – Zurich from the Amadeus GDS.

Figure 2-3: Example of GDS Neutral Display

```
          AN1A06FEBBOSZRH

          ** AMADEUS AVAILABILITY - AN ** ZRH ZURICH.CH                 3 FR 06FEB 0000
                LX 053   F8 J9 C9 D9 Y9 S9 M9 /BOS E ZRH    1835     0755+1E0/332         7:20

                LH 423   F9 A5 O3 C9 D9 Z9 I9 /BOS E FRA 1  1645     0530+1E0/744
1st page        LH3720   C0 D0 Z0 I0 R0 Y0 B0 /FRA 1 ZRH    0645+1   0740+1E0/735         8:55

                AF 337   P7 F6 A2 J9 C9 D9 I9 /BOS E CDG2E  1755     0630+1E0/772
             WX*AF5100   C7 D7 Y9 S9 U9 K9 R9 /CDG2F ZRH M  0730+1   0850+1E0/142         8:55
-----------------------------------------------------------------------------------------------
                AF 321   P9 F8 A4 J9 C9 D9 I9 /BOS E CDG2E  2015     0850+1E0/772
2nd page     WX*AF5102   C9 D9 Y9 S9 U9 K9 R9 /CDG2F ZRH M  0955+1   1115+1E0/142         9:00

                NW 038   J4 C4 Z4 Y4 B4 M4 H4  BOS E AMS    1820     0705+1E0.D10
             KL:NW8411   J4 C4 Z4 Y4 B4 M4 H4  AMS   ZRH    0755+1   0930+1 0.737  TR     9:10
-----------------------------------------------------------------------------------------------
                AF 321   P9 F8 A4 J9 C9 D9 I9 /BOS E CDG2E  2015     0850+1E0/772
3rd page        LX 633   J9 C9 I9 Y9 S9 M9 L9 /CDG2B ZRH    1005+1   1125+1E0/320         9:10

                LH 423   F9 A9 O3 C9 D9 Z9 I7 /BOS E FRA 1  1645     0530+1E0/744
                LX1081   J9 C9 I9 Y9 S9 L9 O9 /FRA 2 ZRH    0715+1   0820+1E0/AR1         9:35
-----------------------------------------------------------------------------------------------
                CO 112   A0 D0 Z0 Y9 H9 K9 N9  BOS C EWR C  1600     1718  E0/738 6
4th page        CO 078   J9 D9 Z0 Y9 H9 K9 N9  EWR C ZRH    1800     0745+1E0/762         9:45

                BA 212   F8 A7 J8 C7 D6 I4 W6 /BOS 5 LHR 4  1805     0530+1E0/772
                BA 710   J9 C7 D7 I5 Y9 B0 L9 /LHR 1 ZRH M  0715+1   1000+1E0/320         9:55

              (without codeshare flights)
```

Source: Amadeus System, retrieval on 04 Feb 2004

The figure shows the first several lines per screen page. The first page displays the nonstop flight by Swiss Airlines (LX) on top. The subsequent connecting itineraries are sorted by elapsed time, which is displayed in the right hand column on the screen. The example shows that a difference of 15 minutes in elapsed journey time can decide whether an itinerary is displayed on the first page (LH423/LH3720) or on the third page (AF321/LX633).

Based on this ranking in GDS and the fact that around 80% of all bookings are made from the first screen, and more than 50% from the first line (House of Lords, 1998), many airlines feel that elapsed time has traditionally been important in passenger choice. However, this might be a confusion of correlation and causality: Bookings for itineraries on the first page might be made for other reasons than screen position. Since GDS' often display nonstop flights first, path quality might for example be more important than the mere fact that an itinerary is on the first screen.

The DOT ruled that display order regulations were to be terminated by July 31, 2004 (Department of Transportation, 2004). This may eventually lead to the disappearing of neutral displays and could reduce elapsed time's impact on screen position. Since the termination of regulation, the ranking is entirely up to the GDS supplier. GDS could try to sell the top spots on their screens or offer price reductions if airlines agreed to have their itineraries moved to a lower rank. In addition, they can exclude airlines from their display, making it necessary for consumers to search different sources if they want a full overview of supply.

Even in a still regulated environment like Europe, GDS screen position could only be relevant when a neutral display is used and no one involved in the decision process has any biases. However, airlines work hard to introduce incentives to all members of the decision chain in passenger bookings: Passengers are incentivized by frequent flyer programs, corporate travel offices by special discounts and travel agents by override commissions. Override commissions are special commissions given to a travel agency when it reaches a certain revenue level with one airline. Thus, a travel agent would try to maximize revenue with one specific airline if the other members of the decision chain had not voiced any preferences regarding their preferred itinerary choice.

The change in market shares of distribution channels is another factor that diminishes the importance of GDS and thus the potential of a "screen position hypothesis". Currently, approximately 50% of airline ticket sales in the U.S. are conducted online (Nielsen/NetRatings, 2005). If passengers book directly with airlines (either online or via telephone), GDS screen position is irrelevant. In addition, internet sites increasingly are used even by travel agents for airline bookings. These sites have never been regulated, i.e. may sort and bias by whatever criteria they choose.

An overview of the treatment of elapsed time on selected internet distribution channels is given in Appendix A, Table A-1. The table displays the treatment of elapsed time and connecting time on airline web sites (American Airlines, Continental Airlines, Delta Air Lines, Northwest Airlines, United Airlines, US Airways, Southwest Airlines), online travel agencies (Expedia, Orbitz, Travelocity), metacrawlers (Kayak, Mobissimo, Sidestep), and opaque search engines (Hotwire, Priceline). Metacrawlers do not sell tickets themselves but act as "meta-agents" who refer their site visitors to airline and online travel agency websites. Opaque search engines do not show airline information before payment. All of the non-airline sites have a fare driven approach on their default screen, sorting by best (lowest) fare. Only two of the airline websites, American Airlines' and US Airways' websites, provide a schedule driven search as their default screen. Results of a schedule driven are either sorted by number of connections, departure time or deviation from requested departure time. None of them use elapsed time as a default sorting criteria. In the initial presentation of search results, 5 of the 16 analyzed sites display separate flights and their departure and arrival times when offering a connecting itinerary. 11 of the 16 sites allow the user to actively sort by elapsed time. This overview shows that elapsed time has lost much of it display power when comparing the "bottleneck GDS era" to today's era of diverse distribution channels. Websites today sometimes even give warnings when connecting times are tight (see Appendix A, Figure A-1 (American Airlines website: "Connecting time either short or long") and Figure A-2 (ITA website: "The layover in Chicago has relatively little room for delays, and for this route a missed connection would likely be very inconvenient").

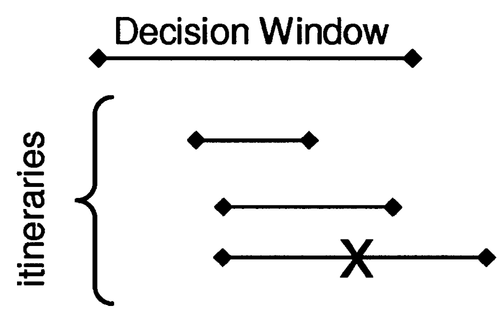
In conclusion, GDS and screen position are often cited to be the reason for a short "elapsed time" being important in passenger itinerary choice. However, deregulation of GDS, incentives across the decision chain and the advent of the internet have reduced the importance of "elapsed time" based on screen ranking in airline distribution.

2.3.2 Decision Window Model

The second important hypothesis about the relevance of elapsed time is based on the Decision Window Model (Soncrant & Hopperstad, 1998, and Boeing Commercial Airline Group, 1993). The Boeing Decision Window Model postulates that a passenger has a predetermined decision window, defined by an earliest departure time and the latest arrival time in which she wants to travel from A to B. The width of the window is a function of the "best", i.e. shortest travel time in the specific market, and the schedule tolerance around the ideal time of the individual. The model assumes that purpose of travel and length of trip influence the amount of schedule tolerance, with business travel and shorter trips reducing the schedule tolerance.

All itineraries within this window will be considered. Figure 2-4 displays an example for this selection process. The individual decision window is represented by the first horizontal line. Three itineraries are displayed below. While the first two itineraries fall within the decision window and thus remain in the choice set, the third itinerary is discarded before the next step. If no itinerary satisfies the decision window constraint, the model assumes that the passenger reconsiders and replans his decision window.

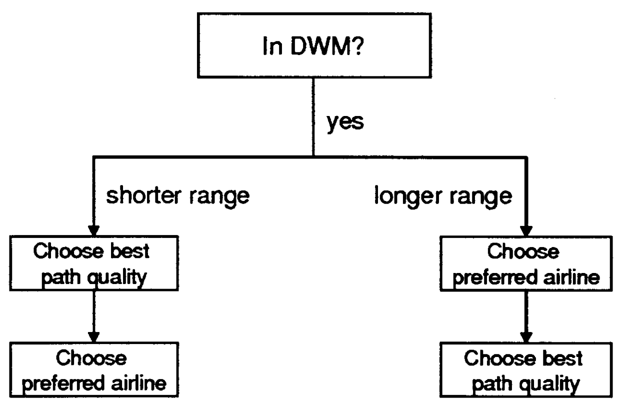
Figure 2-4: Example of Itinerary Selection in Boeing Decision Window Model



If an itinerary satisfies the decision window constraint, DWM assumes that passengers move to the next level of decisions (see Figure 2-5). On the second level, passengers choose either by path quality or by airline. Path quality is ranked from nonstop to direct to connecting flights. From all itineraries remaining in the choice set, a passenger choosing by best path quality would only allow best path quality itineraries to remain in the choice set. Another passenger, choosing by preferred airline, would eliminate all other airlines. The model postulates that a longer range increases the fraction of passengers choosing by airline preference.

The third level criterion in the Boeing Decision Window Model is defined by the second level criterion. If a passenger chooses by path quality on the second level, he would choose by preferred airline on the third level. If he chooses by preferred airline on the second level, he would choose by path quality on the third level.

Figure 2-5: Second and Third Level Decision Criteria of Decision Window Model



A longer elapsed time could cause an itinerary to go beyond a passenger's decision window and thus be excluded from the choice set.

The Decision Window Model can be critiqued in a number of ways. First, it does not include fare in its model, i.e. it assumes that choice is based solely on schedule quality and airline brand. Second, the DWM employs a lexicographic three level rule, not allowing for tradeoffs. For example, even if an itinerary by the preferred airline barely fell out of a postulated individual's decision window, it would be discarded from the choice set. However, this is highly unlikely, since consumer choice typically involves tradeoffs. Third, the foundation of the Decision Window Model is questionable. It assumes that a passenger defines an "ideal" itinerary time window isolated from supply considerations. Passengers could, on the contrary, first check possible itineraries and then finalize their plans at their destination.

In conclusion, while longer elapsed times increase the chance of not falling within an individual's decision window, it is questionable whether the model is an accurate representation of the passenger itinerary choice process.

2.3.3 Preference for Short Elapsed Times

The third rationale for the importance of elapsed time is the hypothesis that passengers *prefer* short elapsed times. It is a common statement in the literature that "passengers prefer shorter connections" and "discount longer connections" (Mayer & Sinai 2003), derived from the fact that time has a value in itself. However, this statement incorporates neither the risk of misconnection nor the effect of the traveler being rushed.

While passengers may well ex post prefer shorter actual connections, it is questionable whether they ex ante prefer shorter scheduled connections. With increasing congestion and delays in the 1990's, missed connections for passengers and luggage increased as well. Having a shorter scheduled connection increases the risk of misconnection and thus influences the expected elapsed time, accounting for a certain probability of misconnection. Passengers may discount the variance of elapsed time even more than a longer scheduled connection. For example, a passenger may be willing to accept a longer connection time in exchange for a higher probability of actually making that connection.

The second effect of short connections may be the feeling of being rushed. While airport terminals worldwide have been expanded, the minimum connecting times set by airports have typically remained constant as airports see small minimum connecting times as a strong competitive advantage. Even if misconnection was not an issue, passengers might prefer a relaxed transfer versus a rushed one at an airport. With the increase of the average population's age in the Western world, this factor might increase in importance in the future.

Another effect regards the value of time itself. Traditionally, connecting time is seen as idle time that is unproductive for the passenger. However, the information age reduces the opportunity costs of longer connecting times. Passengers today can use their cell phones or mobile devices in airports and be as productive as elsewhere. For other uses, airports today offer lounges, banks and shopping areas that allow to utilize the time in whatever way preferred. Consequently, time on ground may feel less onerous than without these options present.

2.3.4 Summary

All of these factors lead us to question whether minimizing elapsed time is as important for airlines as it may have been earlier and whether minimizing *scheduled* elapsed time is what passengers truly prefer. Table 2-1 gives a summary of the discussed items.

Table 2-1: Summary of Attributed Importance of Elapsed Time Discussion

	Primary Supporting Arguments	Contrary arguments
GDS Screen Position	80% of all bookings are from the first page, elapsed time relevant for screen position	- Prices are deregulated - GDS are deregulated - Incentives (FFP, etc.) - GDS' share decreases due to internet
DWM	Longer elapsed time may lead to "falling out of window"	- DWM doesn't include price - decision rule in reality may not be lexicographic
Passenger Preference	Value of time	- Risk of misconnection not taken into account (relevant is the discounting of variability in actual travel time); risk of misconnection has increased because of larger terminals, more air traffic congestion, more time consuming security measures, while MCT often stayed the same) - Rush aversion in combination with older average traveling age not taken into account - Information age reduces opportunity costs of longer transfer times (out of aircraft time does not have to be idle time, but can be used for cell phone conversations, wireless internet, etc.) - Terminals allow for better use of time (lounges, shopping areas, banks, bookshops, etc.)

2.4 Case Study: Passengers Voluntarily Avoiding Short Connecting Times

The goal of this case study is to get an indication whether or not passengers voluntarily avoid short connecting times, all other attributes of the itinerary being equal. In order to answer this question, we analyze what share of passengers voluntarily books a longer connection for a given itinerary.

We use data from a large network carrier. The analyzed booking period is August 11th 2005 until September 30, 2005 with flight departures from September 01, 2005 until September 30, 2005. Once per day between August 11th and September 30th, availabilities for all September departures on a selected origin destination market were logged. In addition, all bookings made for this Origin-Destination market in the defined booking period where logged. As part of the analysis, bookings were merged with availabilities in order to check whether the alternate itinerary was available in the booked reservation class on the day of the booking. If it was available, the passenger was counted in the following result table.

For the selected Origin-Destination Market, we analyze which inbound flight was chosen for a fixed outbound flight. Booked itineraries can either have a short connection time (the shortest connection available for sale by the airline) or a long connection (the next longer connection, leaving earlier from the spoke airport and transferring onto the same outbound flight as above). We define an itinerary as being booked voluntarily if the booked reservation class is available on both itineraries on the day of reservation. This ensures that there are no differences in airline, fare or earned mileage in frequent flyer programs between the two itineraries. Figure 2-6 exemplifies the approach. The first line shows itinerary A with a connecting time of 1 hour at the hub. The second line shows itinerary B. Itinerary B has the same outbound flight as itinerary A. However, the inbound flight into the hub is an earlier one than in itinerary A, with the connecting time at the hub being 2 hours. When choosing between the two itineraries, passengers can trade-off the smaller risk of misconnection and less rush of itinerary B versus the later departure time of itinerary A. We assume that both itineraries are known to the passenger. If only one itinerary was know to the passenger, it would in all likelihood be the shorter connection, since GDS as well as the carrier's internet distribution site rank by elapsed time.

The specific origin-destination market was chosen because it has a high frequency of in-bound flights into the hub, allowing for an incremental change in connecting time. The first flight is a short distance feeder flight (approximate flight time 1 hour), whereas the second flight is a long-distance flight. The minimum connecting time at this airport for this itinerary is 60 minutes. The long-distance destination is served up to five times a day from this specific hub. The airline providing the data has a market share of approximately 30% in this market, i.e. passengers would have other choices if they were willing to switch to a different airline. Due to the proprietary nature of this data, the exact origin, hub and destination have been disguised.

Table 2-2 displays the analyzed itineraries. The second and third column respectively display the departure and arrival time of the origin to hub flight. The fourth column displays the scheduled connecting time onto the outbound flight. The first row of each of the three itinerary pairs shows the long connection, whereas the second row of each pair shows the short connection.

Figure 2-6: Short and Long Connections in Alternative Itineraries

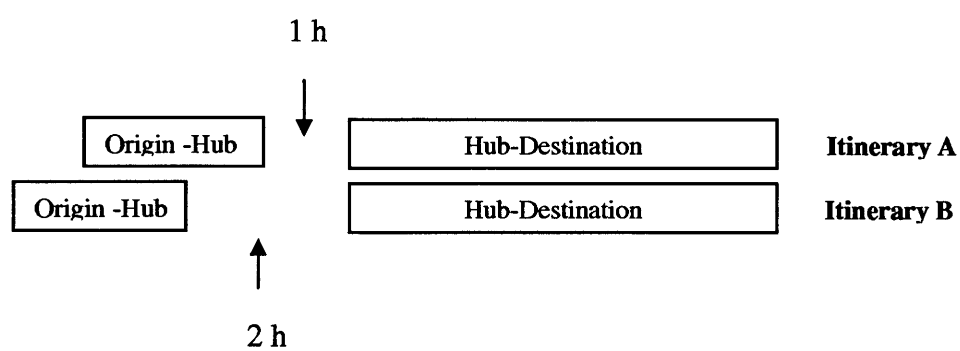


Table 2-2: Itineraries Analyzed in Case Study

	Departure Time	Arrival Time	Scheduled Connecting Time in Minutes		Departure Time
Origin-Hub	06:50	07:55	125	Hub-Destination	10:00
Origin-Hub	07:40	08:45	75	Hub-Destination	10:00
Origin-Hub	09:55	11:00	125	Hub-Destination	13:05
Origin-Hub	10:55	12:00	65	Hub-Destination	13:05
Origin-Hub	09:55	11:00	150	Hub-Destination	13:30
Origin-Hub	10:55	12:00	90	Hub-Destination	13:30

For example, a passenger could have booked the 0740 departure in reservation class "W". If "W" was also available at the time of booking for the itinerary departing at 0650, we assume that the passenger booked the shorter itinerary voluntarily. Another example would be a passenger booking the 0650 flight in "K" class. As long as the "K" class was also available for the itinerary leaving at 0740 at the time of booking, we assume that the longer connection was booked voluntarily. On the contrary, if the alternate itinerary was not available in the booked reservation class at the time of the booking, we disregard this passenger in subsequent analysis.

Results are segmented by cabin class and itinerary group, based on the three outbound flights. In aggregate, 20% of all passengers choose the long connection, and 80% the short connection. A differentiation by cabin class (see Table 2-4) does not yield large differences between the behavior of passengers who booked First and Business class versus those who booked Economy class.

Table 2-3: Aggregate Results of Case Study

Long	Short
20%	80%

Table 2-4: Results by Cabin Class

	Long	Short
First/Business	18%	82%
Economy	23%	77%

Table 2-5 shows the results differentiated by the individual itinerary pairs. One can see that the highest share of of passengers choosing the longer connection (25%) occurs when the short connection is 65 minutes and the flight is during midday. For the morning itinerary, only 9% choose the longer connection with 125 minutes versus the short connection with 75 minutes. For the third itinerary pair, even the short connecting time is 90 minutes, leading to a share of 90% for the short itinerary.

Table 2-5: Results by Connecting Times and Time of Day

Departure time feeder flight	Arrival time feeder flight	Connecting time	Departure time of trunk flight	First (abs.)	First (%)	Business (abs.)	Business (%)	Economy (abs.)	Economy (%)	Total (abs.)	Total (%)
06:50	07:55	**125**	10:00	0	0%	1	6%	2	14%	3	**9%**
07:40	08:45	**75**	10:00	2	100%	17	94%	12	86%	31	**91%**
09:55	11:00	**125**	13:05	1	20%	17	26%	18	25%	36	**25%**
10:55	12:00	**65**	13:05	5	80%	49	74%	54	75%	108	**75%**
09:55	11:00	**150**	13:30	0	0%	1	6%	3	15%	4	**10%**
10:55	12:00	**90**	13:30	3	100%	15	94%	17	85%	35	**90%**

The results indicate that of all passengers who have the choice to book a short connection versus a long connection, a share of up to 25% chooses the longer connection. This is an indication that passengers may take into account factors like the risk of misconnection or the avoidance of rush. The purpose of the subsequent chapters is to research these effects in more detail.

3 Literature Review of Airline Passenger Choice Models

3.1 Introduction

This chapter presents a review of previous research in the area of airline itinerary choice. First, studies using revealed preference data will be discussed, followed by those using stated preference data. Some of these works include elapsed time in their models, and connections are primarily modeled as number of connections. To our knowledge, no research exists that specifically studies the non-monotonicity of connecting time or includes attitudes in airline itinerary choice modeling.

3.2 Airline Passenger Choice Models

Airline Passenger Choice models may be divided into two broad categories: Those using revealed preference data (booking data) and those using stated preference data (experimental data). Booking data seldom contains socio-demographic information and does not contain information on the initial choice set of the individual. It does, however, represent choices in a real world context, which may be more reliable than hypothetical scenarios. Stated preference data, on the other hand, has the advantage that the range of attribute levels can be extended, that multicollinearity among attributes can be avoided by survey design, and that the choice set is pre-specified. Major disadvantages of stated preference models are the fact that the respondent may be influenced by justification bias for his current choice ("inertia effect") or that the respondent diverts from his actual choice behavior to make an opinion statement, e.g. always choose the lowest price. (see Morikawa, 2002, for a more detailed discussion of advantages and disadvantages of revealed versus stated preference data). First, revealed preference studies will be discussed, followed by stated preference studies.

Kanafani & Sadoulet (1975) is one of the earliest multinomial choice studies in an airline choice setting. It uses traffic data from the North Atlantic to estimate a model that has relative fare difference, seasonal shifts, fare type and season dummies as variables. The study focuses primarily on the differences between (regulated) fare types and shows significant impacts of fare types and season.

Alamdari and Black (1992) give an overview of the use of the logit model in passenger choice until 1992. Some of these works, e.g. Kanafani & Ghobrial (1985) include the number of connections in their models, which has a negative coefficient value. However, the model does not include any more detail on the connection issue.

Coldren et al. (2003) and Coldren & Koppelman (2005) estimate aggregate air – travel itinerary share models at the city pair level. The model makes use of path quality, connections, carrier, carrier market presence, fares, aircraft size and type and time of day service characteristics. The model allocates total city-pair air travel volume to itineraries available in that market. The drawback of this approach is the use of average fare data in the model and the sole reliance on data from computer reservation systems that does not include bookings directly via airline systems. Consequently, there is a bias toward schedule importance. Itineraries were built using the United Airlines itinerary building engine, using distance based circuitry logic to eliminate unreasonable itineraries and minimum and maximum connection times to ensure "that unrealistic connections are not allowed." The dependent variable in the model is the number of passengers who booked each itinerary. Results indicate that path quality (nonstop, direct, single-connect, double-connect) is highly significant. However, since fares are averages, this result might be misleading because the tradeoff between fare and level of service was not adequately modeled. Coldren & Koppelman assess that "the results indicate that travelers making a connection strongly prefer the best connection (shortest ground time) among itineraries sharing a common leg at a transfer station". We do not know, however, if passengers chose the shortest connection because this was the only one offered to them. In addition, results show a "second-best connection time difference", representing a time difference penalty. However, it is not known whether the passengers would pay more for having a shorter connection.

Schröder (2004) describes the market share model employed by Lufthansa German Airlines. Data is extracted from the central reservation systems and ticket databases, and Lufthansa's market share on an origin-destination market is compared to the total market size. A connection builder is used to establish all itineraries from origin to destination. In a second step, those itineraries deemed infeasible are filtered from the set. Attributes relevant in this filtering process are path quality, total elapsed time, connecting time, number of connections, airport transfers across a city, market size, distance, and distance relative to shortest distance. Market shares on given itineraries are then modeled relative to all remaining connections as set up by the connection builder. As attributes of the choice model, "departure and arrival preferences", a competition variable, a code-share variable, elapsed time, path quality, aircraft type, an alliance variable and airline preference are used. Market models are estimated for 17 regions worldwide.

Nason (1980) is one of the first *Stated Preference applications* in the airline choice field. The choice is expressed in terms of socio-demographics, trip characteristics and the service level associated with each alternative. Since the model was developed in the regulated fare environment, it focuses on "ticket type choice" rather than on fares itself. Explanatory variables are trip purpose, number of cars in household, expected delay if flying standby, profession and a dummy variable for vacation destinations. Departing passengers at Boston – Logan airport are presented with a binary choice between a full fare ticket and a standby ticket.

Ndoh (1990) develops a nested multinomial logit model employing survey data from the 1986 Civil Aviation Authority survey conducted at four central England airports. Even though Ndoh focuses on airport choice, he incorporates travel related attributes like frequency, elapsed time and time of first leg when connecting into his model. The model does not, however, contain any socio-demographic or fare data.

Proussaloglou & Koppelman (1995) and Proussaloglou & Koppelman (1999) develop an extensive survey for nonstop markets Chicago – Denver and Dallas – Denver. Respondents make tradeoffs between carrier service attributes (frequency, quality of service, reputation), flight schedule attributes (schedule delay relative to the respondent's preferred flight times) and fare class attributes (fare level, advance purchase requirements, travel restrictions, cabin class related terminal and on-board service amenities, and penalties for canceling one's flight). The focus of Proussaloglou & Koppelman is the tradeoff between different attributes in a nonstop market. The quality of service was reflected by traveler's ratings that were used to extract factors that were then used directly as attributes in the choice model.

Bradley (1998) describes a study undertaken by Hague Consulting Group to study air traveler route choice in the Netherlands. While the focus was on competing airports, the study also included attributes of the itineraries themselves. Attributes of the choice model were departure airport with its associated access time, flight frequency, path quality and fare. The path quality was varied from nonstop flight, to connecting flight with short and connecting flight with long connecting time. Connecting times used were in the range of 45 min to 2 hr for European destinations and 1-3 hr for intercontinental flights. Trip characteristics of a recent trip and socio-demographics were collected but not used in the choice model, stating that the application of this model in the future would not be possible because that data is not readily available in a forecasting context. Models were estimated in segments by trip purpose and trip distance.

Table 3-1: Valuations Inferred from the Bradley (1998) Route Choice Model

	Business Europe	Business Intercontinental	Leisure Europe	Leisure Intercontinental
Value of Connecting Time (NLG/hr)	151	144	38	36
Penalty for connection (NLG)	198	225	62	54

Penalty values for each extra hour of connecting time were four to five times higher for business travelers than for leisure travelers. Connecting time was modeled linear-in-parameters, not allowing any conclusion on whether or not there is a penalty for very short connections. Similarly, penalties for connections as such were three to five times higher for business travelers than for leisure travelers (see Table 3-1).

Adler et al. (2005) and Hess et al. (2005) use a 2001 survey by Resource System Group to estimate models of airline itinerary choice. The models include the effects of airline, airport, aircraft type, fare, access time, flight time, number of connections, scheduled arrival time and on-time performance. Adler et al. develop logit as well as mixed logit models for the full survey population and segment the models by trip purpose into business and non-business travelers. Their results show the relative importance of number of connections in itinerary choice, however they do not specifically model the length of connecting time. Results indicate that air fare is the variable with most explanatory power, regardless of whose choice is modeled or which type of model is used. Connections are modeled as dummy variables for either one or two connections. Their results also reveal the substantial effect of frequent flyer status on itinerary choice, showing a much stronger airline preference for elite status program members.

3.3 Implications for this Research

Previous studies have found, as expected, that elapsed time and number of connections have a negative effect on passenger utility. However, none of these studies have focused specifically on connecting time by treating it in a way to study if there is a disutility of very short connecting times. Early studies have also omitted fare and frequent flyer status, factors that have a large impact on choice in the current airline choice context. What is missing so far is a comprehensive study focusing on elapsed time and especially connecting time, challenging the old assumptions about these two factors. In addition, passengers' attitudes, such as attitudes toward risk or rush, have not been included in passenger itinerary choice models so far. However, they can have a large impact on how consumers behave (Churchill 2004) and can give a behaviorally more realistic explanation of airline itinerary choice. In the following chapter, a methodology is presented that focuses on elapsed time and connecting time and includes attitudes in the choice model.

4 Analytical Approach

4.1 Introduction

In this chapter, the analytical approach will be presented. In section 4-2, potential decision rules and their applications to airline passenger choice are discussed. The behavioral assumptions and hypotheses are presented in section 4-3, and in the following section (4-4), possible data sources for modeling are explored. In the subsequent section (4-5), we present the basic choice model and its extensions. Section 4-6 introduces the concept of latent variables and details the modeling steps for the "Multiple Causes/Multiple Indicators Model" (MIMC) and the integration in an "Integrated Choice and Latent Variable Model" (ICLV). This chapter introduces the concept, while specification, estimation and results will be presented in chapter 6.

4.2 Decision Rule

Various decision rules could be applied to the choice problem, among them dominance, satisfaction, lexicographic rules and utility maximization (Ben-Akiva & Lerman, 1985). Dominance seldom leads to a unique solution and is thus disregarded. Both satisfaction and lexicographic rules do not allow for tradeoffs: Satisfaction rules eliminate alternatives if at least one of its attributes is below a predefined satisfaction level. Lexicographic rules (e.g. the decision window model) would rank first by only one attribute. In the case of the decision window model, all itineraries outside of the decision window would be removed from the choice set, even if other attributes, e.g. fare, were much more favorable in the removed alternative. Passenger airline choice, however, shows that tradeoffs are made: Passengers for instance choose a different itinerary if the price is reduced or other attributes are changed. Thus, the decision rule should allow for tradeoffs. In the utility decision rule, the decision maker maximizes the utility, whereby attributes can be traded off against each other. Since not all attributes of the alternatives and all characteristics of the decision maker(s) can be observed, the random utility approach is widely used. Random utility theory assumes that observed choice behavior is probabilistic and it is assumed that the decision maker maximizes his utility based on the attributes of the alternative. However, not all of these attributes are observable, making utility a random variable for the analyst, consisting of an observable component and a random component (Ben-Akiva, Lerman, 1985).

4.3 Definitions, Behavioral Assumptions and Hypotheses

The following definitions are used in the subsequent sections:

An *itinerary* is a series of flights from a passenger's origin to her destination. An itinerary may contain only one flight (i.e. a non-connecting itinerary) or more than one flight (i.e. a connecting itinerary)

Elapsed time of an itinerary is the time between the scheduled departure of the first flight and the scheduled arrival of the last flight.

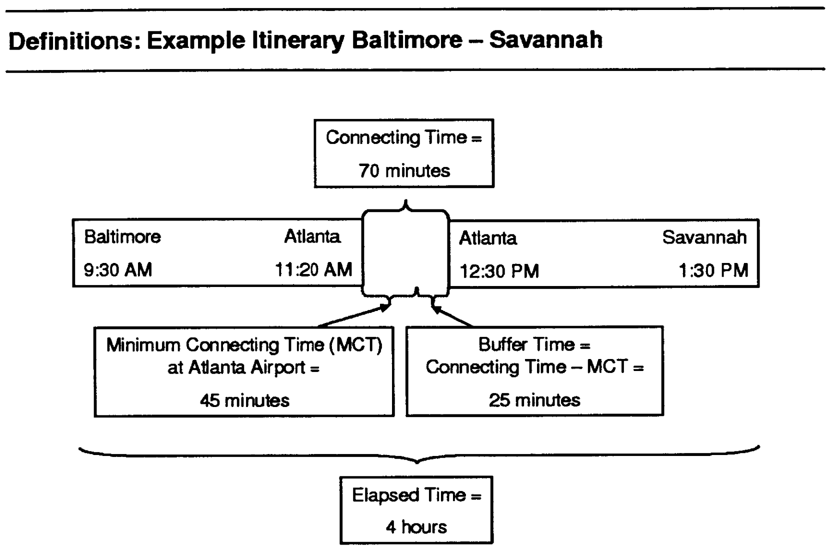
Connecting time is defined as the time between the scheduled arrival of a flight and the scheduled departure of a subsequent flight within an itinerary.

Minimum connecting time is the minimum feasible connecting time at an airport as published by the corresponding airport authority.

Buffer time is the difference between connecting time and minimum connecting time.

Figure 4-1 exemplifies these definitions. The graph shows an itinerary from Baltimore to Savannah. The flight at Baltimore leaves at 09:30 am, arriving at Atlanta airport at 11:20 am. The second flight leaves at 12:30 pm, and is scheduled to arrive at Savannah at 1:30pm. The elapsed time of this itinerary, i.e. the time from the first departure to the last arrival, is 4 hours. The connecting time at Atlanta airport, i.e. the time from arrival of the first flight to departure of the second flight, is 70 minutes. The minimum connecting time for Atlanta as published by the airport authority is 45 minutes, resulting in a buffer time of 25 minutes.

Figure 4-1: Definitions for Time Segments of an Itinerary



While many attributes are important in itinerary choice (e.g. fare, frequent flyer program membership, etc.), our hypothesis focuses on the effect of connecting time on itinerary choice. We hypothesize that there are three components associated with the utility of connecting times. First, passengers' value of time lets us assume that a longer scheduled connecting time leads to a lower utility level. Second, we assume that the risk of misconnection is a factor in assessing the length of connection time. Third, we assume that short connecting times (close to the minimum feasible as published by the airport) may create a feeling of rush for the passenger and thus would have a lower utility than longer connecting times.

The sum of all three effects, value of time, risk of misconnection and rush describes the utility of buffer time to an individual. We could assume, for instance, that the effect of rush can be mitigated by allowing 15 minutes of buffer time in addition to the minimum connecting time. Regarding the risk of misconnection, we could hypothesize that at a buffer time of 45 minutes, virtually all connections are successful. For value of time, we could assume a linear trend, reducing the utility with increasing buffer time. For illustrative purposes, the impact of the three components on buffer time utility is displayed in Figures 4-2 to 4-4.

Figure 4-2: Buffer Time Utility with Relatively Weak Risk of Misconnection Utility

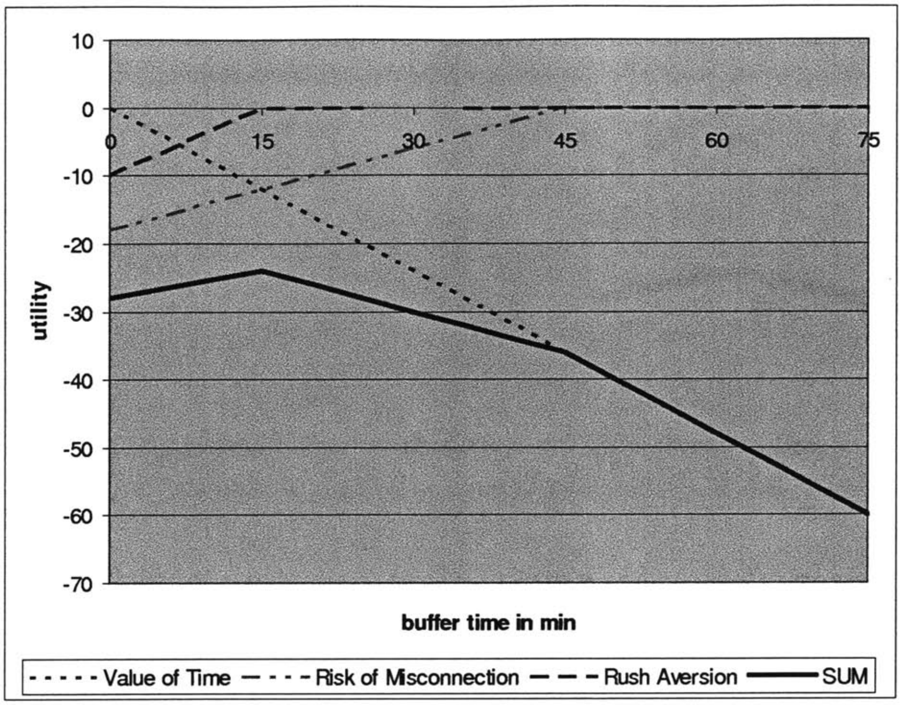


Figure 4-3: Buffer Time Utility with Relatively Strong Risk of Misconnection Utility

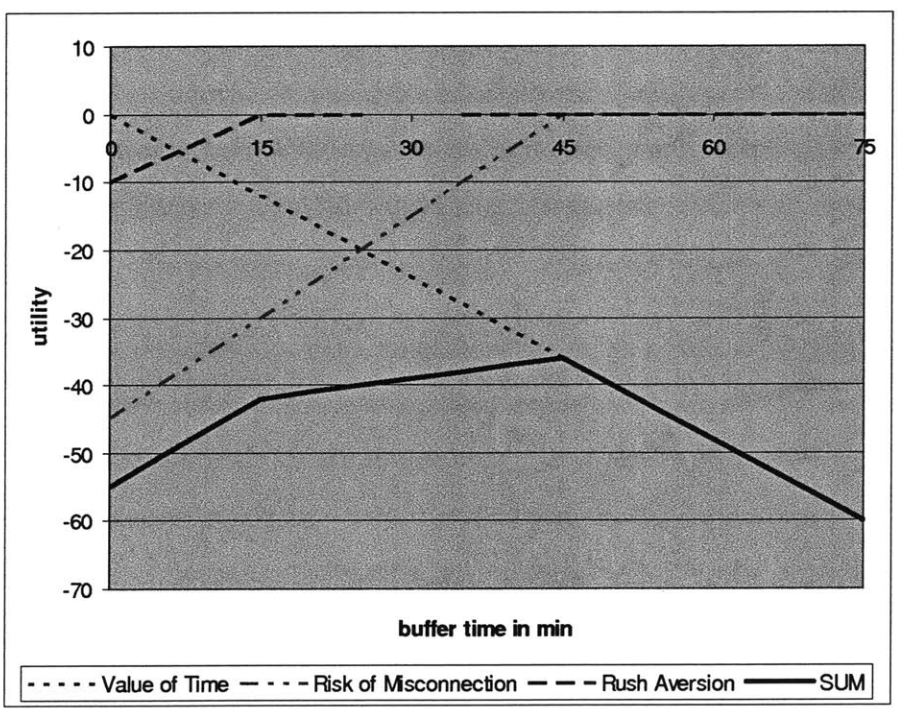
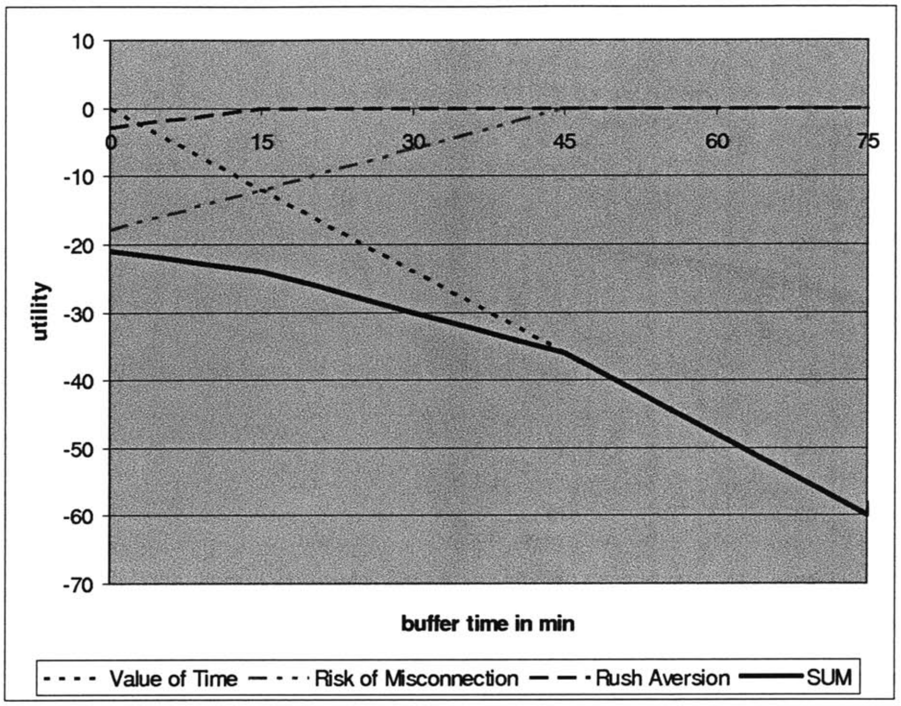


Figure 4-4: Buffer Time Utility with Relatively Weak Risk of Misconnection and Weak Rush Aversion Utility

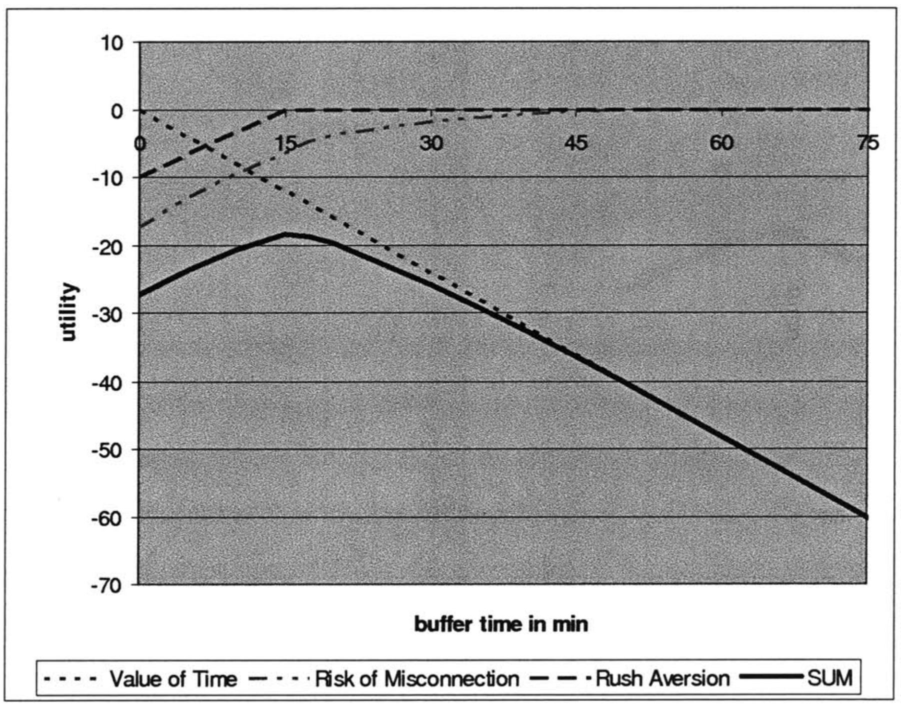


Depending on the relative impacts of risk of misconnection utility and rush aversion utility, the aggregate buffer time utility has different shapes, as seen in Figures 4-2 to 4-4. Figure 4-2 describes a situation where the effect of risk and rush are larger than that of value of time for the first 15 minutes of buffer time, leading to increasing buffer time utility. In the second segment, the effect of value of time is larger than that of risk of misconnection, reducing utility. In the final segment, only the effect of value of time is relevant, reducing the slope of the curve even more. Figure 4-3 differs from figure 4-2 in the middle segment. In this scenario, the effect of risk of misconnection is larger than that of value of time, leading to an increase in utility between 15 and 45 minutes. Figure 4-4 shows an example of a monotonic buffer time curve. In this case, already in the first segment from 0 to 15 minutes the effect of value of time is larger than that of the sum of risk and rush, leading to the monotonic curve. In this case, increasing buffer time decreases utility at any buffer time value.

Since flight arrivals usually have a distribution similar to a normal distribution, a more realistic representation of risk of misconnection utility is an asymptotic one as displayed in Figure 4-5. By adding just a few minutes of buffer time to the minimum connecting time, a large portion of misconnections can be avoided. In this example, the aggregate buffer time utility increases up to a level of 15 minutes of buffer time and decreases subsequently.

The purpose of our research is to test whether this hypothesis of non-monotonic buffer times can be confirmed and whether there exists any heterogeneity regarding buffer time utility in the population.

Figure 4-5: Buffer Time Utility Hypothesis



4.4 Data Sources: Revealed versus Stated Preference

Data for a model incorporating connecting times can be retrieved from observed choices (revealed preference) or from choice experiments (stated preference). Our choice of data is largely based on feasibility. Historic booking data is available from GDS

(Global Distribution Systems) in form of MIDT (Market Information Data Transfer) files. These are data derived from the major GDS's and have been used previously in passenger choice models (see Chapter 3, e.g. Coldren & Koppelman, 2005). However, the data has several drawbacks. First, it only reveals the final choice of a passenger. The original choice set of the passenger remains unknown. Even though a universal schedule would show all scheduled flights, it would neither show whether these itineraries were actually available to the passenger at time of booking nor at what fare. Second, we have no socio-demographic information (e.g. frequent flyer membership) on the passenger across airlines. Third, MIDT data only gives us bookings via GDS, and only some of the bookings made directly through airline websites pass through GDS. Thus, MIDT data would potentially overestimate the importance of schedule since it does not capture all passengers who choose by airline (and book directly through the airline). Fourth, MIDT data does not give us any information on attitudes or perceptions of travelers. From the previous discussion we conclude that a more advantageous option to study the field of interest is to perform a survey. In a survey, the choice set of the respondent is known. Socio-demographic information can be collected so it can be incorporated into the modeling process. In addition, attitudes and perceptions can be captured that may provide a richer explanation of choice. The survey will be presented in chapter 5.

4.5 Choice Model

In this section and the next, the model framework is described. Equations, specification, estimation and results will be described in Chapter 6. In this first section (4.5), the basic choice model is described. In the next section (4.6), the latent variables risk, rush and trust that are hypothesized as being important to have a more realistic representation of itinerary choice will be introduced into the model.

Discrete choice models are widely used in applications in transportation, marketing, and other areas. A discrete choice is a choice between mutually exclusive and collectively exhaustive alternatives, for instance a choice of a certain itinerary in a given Origin-Destination market. The most prominent discrete choice model is the Multinomial

Logit Model (MNL) (Ben-Akiva, Lerman, 1985). MNL is a random utility maximization model, where we assume that the decision maker maximizes his utility while the analyst does not have complete information, making utility a random variable to the analyst. In the following section, the basic choice model is reviewed.

The utility an individual allocates to a certain alternative is modeled with a deterministic and a random part. The utility of an alternative i for individual n depends on the attributes of the alternative and the socio-demographic characteristics of the individual and an error term, where the error term is assumed to be i.i.d. extreme value distributed. In itinerary choice, the utility of an alternative could depend on fare, membership in frequent flyer program, number of connections, and other attributes and characteristics.

Figure 4-6: Framework for the Choice Model

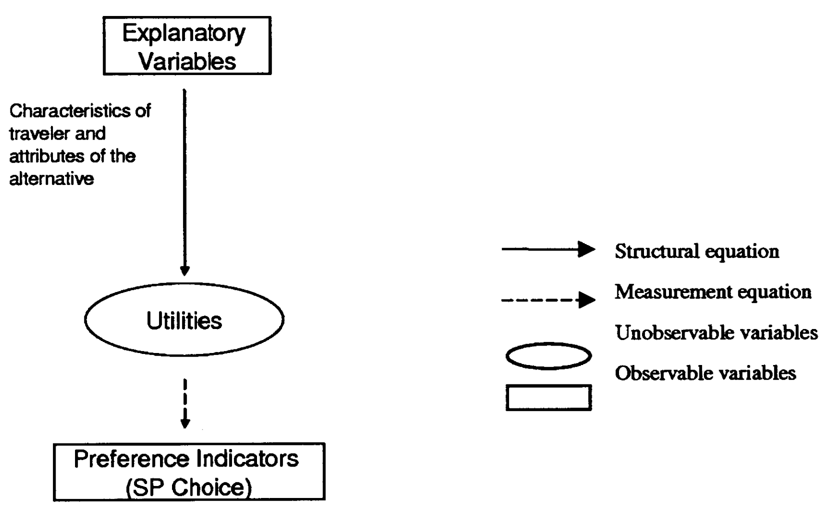


Figure 4-6 is a graphical representation of the framework. Observable variables are displayed in rectangles, whereas unobservable variables are shown as ovals. Solid arrows depict causal relationships (structural equations), while dotted arrows depict the relationship between a latent variable and its indicator (measurement equation). In the choice model, explanatory variables are either socio-demographics of the decision makers (e.g. gender, trip purpose) or attributes of the alternatives (e.g. fare, elapsed time, number of connections).

The choice model relies on important causal variables being directly observable. However, decisions are not only determined by attributes of the alternatives and observable characteristics of the decision maker, but also by underlying, non-observable attitudes of the decision maker. In the next section, the framework to specify the systematic part of the utility function in a more realistic way by including latent constructs is described.

4.6 Integration of Latent Variable Models

4.6.1 Introduction

Latent variables are unobservable to the analyst. Unlike attributes of the alternatives or socio-demographic characteristics, they are not directly measurable. Our hypothesis regarding buffer time is that the attitudes toward risk of misconnection, rush, and trust into airline's scheduling reliability are important causal factors in the decision making process when choosing an airline itinerary. Since these attitudes are latent, traditional discrete choice analysis methods cannot be used. In the following section, the framework of integrating latent variables into choice models is described.

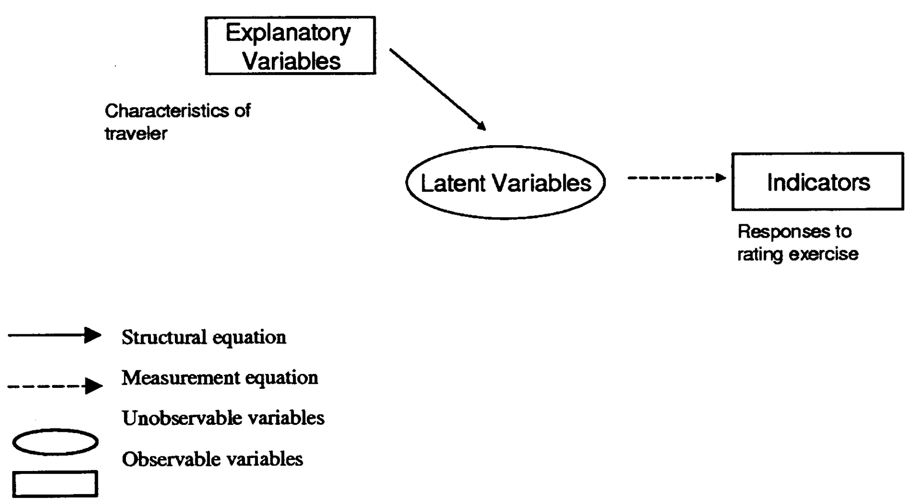
We hypothesize that there is heterogeneity in the assessment of connecting times, and latent variables (risk, rush, trust) are used to explain this heterogeneity. Latent variables are extracted from indicators, which are manifestations of the underlying latent variable. There are two steps involved in the integration of latent variables into the model. The first one is the MIMC model, consisting of measurement equations with the psychometric indicators represented as a function of the latent variables, and a structural equation, explaining the latent variables with observed variables. The second step is the integration of the choice and the latent variable model.

4.6.2 Multiple Indicator Multiple Causes Model (MIMC)

In a MIMC model, indicators are used as manifestations of latent variables. The latent variables are described by socio-demographic characteristics of the individual. MIMC or structural equation models have been pioneered by the work of Jöreskog (1973), Keesling (1972) and Wiley (1973). For a comprehensive review of structural equation modeling, see Bollen (1989), Kline (2005) or Schumacker & Lomax (2004).

The MIMC model consists of measurement equations and structural equations. In the measurement equations, indicators are explained as a function of latent variables, whereas the structural equation explains the latent variables with socio-demographics of the individual. Figure 4-7 shows the MIMC model. The next step will be to integrate the MIMC model with the choice model.

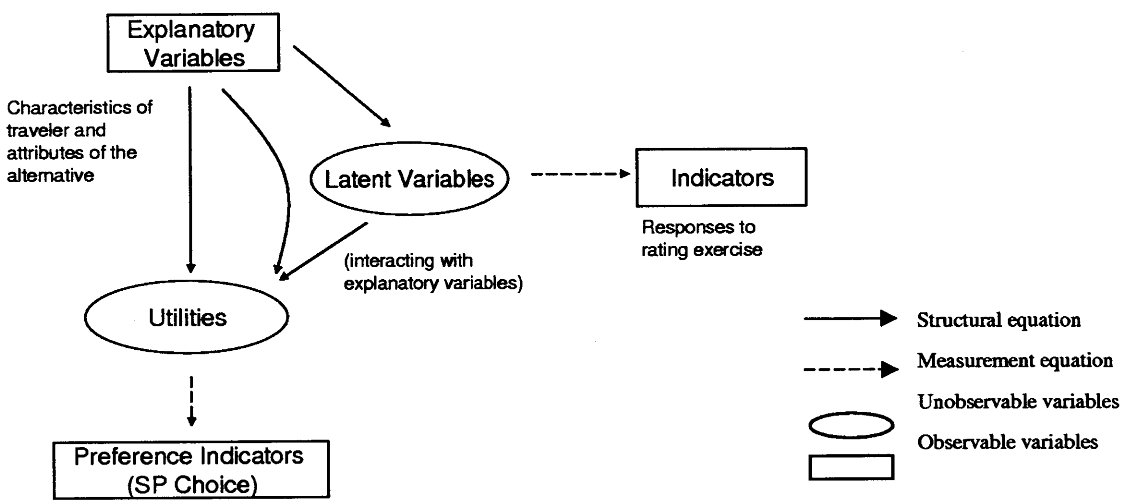
Figure 4-7: Multiple Indicator Multiple Causes Model



4.6.3 ICLV Model (Integrated Choice and Latent Variable Model)

The Integrated Choice and Latent Variable Model integrates the MIMC model and the choice model in one model. Figure 4-8 is a graphic representation of the model. The formulation is based on the work by McFadden (1986), Cambridge Systematics (1986), Ben-Akiva and Boccara (1987), Ben-Akiva et al. (1999), Walker (2001), Walker and Ben-Akiva (2002), Morikawa, Ben-Akiva and McFadden (2002), Ben-Akiva et al. (2002a), Ben-Akiva et al. (2002b), Bolduc and Giroux (2005) and Bolduc, Boucher and Alvarez-Daziano (2008).

Figure 4-8: Integrated Choice and Latent Variable Model



In the integrated choice and latent variable model, all previously described components are estimated simultaneously. There are two structural equations, one explaining the latent variables with socio-demographic characteristics, and the other one explaining the utility with attributes of the alternatives and latent variables. Latent variables are used like any observable variable in the choice model. If they represent perceptions of an alternative, they can be used as a stand alone variable explaining utility. In our case, latent variables represent an individual's attitudes towards certain constructs (e.g risk), i.e. provide socio-demographic information that either can be included alone in the utility function or interacted with attributes of the alternatives. In this dissertation, both alternatives are air itineraries and the attitudes are characteristics of individuals. A non-interacted inclusion in the utility function would not make sense, since risk or rush do not favor one of the alternatives per se. (This is contrary to a mode choice model, where attitude toward rush could make an individual favor e.g. car vs. public transit). Consequently, the latent variables are interacted with attributes of the alternative in order to analyze the impact of the latent variable on the valuation of the attribute.

The integrated choice and latent variable model also has two measurement equations. The first measurement equation is the one introduced as part of the MIMC model. It explains the indicators as a function of the latent variables. The second measurement equation explains the choice in the stated preference experiment.

The parameters are estimated by maximum likelihood estimation, leading to consistent and efficient estimators. Simulation is required in estimation as the likelihood function for the integrated model consists of multi-dimensional integrals. The specific specification and likelihood function will be described in Chapter 6.

4.7 Conclusion

In this chapter, the behavioral assumptions and hypotheses and the general framework and approach to test the hypothesis of non-monotonic buffertime utility was presented. It is assumed that individuals have a lower utility for connecting times close to the minimum connecting time versus slightly longer connecting times. Factors for this behavior are hypothesized to be the value of time, attitudes toward risk of misconnection and rush aversion and trust into airlines' scheduling reliability. We assume that utility for buffer time increases with a certain increase in time and then decreases based on value of time aspects. We present the general framework to test this hypothesis and to analyze any heterogeneity regarding buffer time. The first step is a choice model incorporating buffer time. In the second step, indicators are used to extract the latent variables (risk, rush, trust) and these latent variables are explained with socio-demographic variables of the individuals. In the third step, the choice model and the latent variable are jointly estimated in order to produce a more behaviorally realistic model.

5 Survey Design and Summary Statistics

5.1 Introduction

Based on the framework discussed in the previous chapter, a survey instrument was designed to collect the data used in the estimation of a choice and latent variable model. This chapter describes the survey. The questionnaire design is described in the following section. Section 5.3 describes the execution of the survey. Finally, section 5.4 presents descriptive statistics of the survey data.

5.2 Questionnaire Design

The survey described here is part of a larger study that is performed annually by Resource System Group (see e.g. Adler, 2005). The survey was adapted and expanded to serve the purpose of this dissertation. In this chapter, the sections of the survey are described in more detail. Respondents were first asked to provide information on a recent trip (5.2.1). They were then asked to describe their preferences regarding airports and airlines (5.2.2). The following stated preference experiment is the central element of the survey, giving respondents the choice between different itineraries (5.2.3). The next step of the survey is a rating exercise, collecting information on attitudes toward risk, rush and trust in airlines' scheduling reliability from the respondent (5.2.4). Finally, respondents are asked to provide socio-demographic information (5.2.5). Each part of the survey will be described in detail in the following sections.

5.2.1 Information on Previous Trip

At the beginning of the survey, the respondent was asked to describe her most recent U.S. domestic air trip. Trips made using frequent flyer miles or using any other kind of coupon or voucher did not qualify. The survey covered the outbound portion of the respondent's air trip.

The items describing the recent trip and selection choices are summarized in Table 5-1. These trip characteristics were selected and collected because it was hypothesized that they could explain the choice of itinerary and give an explanation for the heterogeneity in attitudes. For the pre-purchase phase, the collected characteristics were trip purpose, ticket sponsor, party size, the preferred arrival time, and the used distribution channel. Choice behavior and attitudes might vary systematically based on these criteria.

The itinerary, pre-flight and flight characteristics describe the attributes of the trip itself. Itinerary characteristics are those attributes of the trip that are inherent to the itinerary, e.g. date and time of the itinerary, airline, connecting airport, etc. Pre-flight characteristics are access mode and time, the time the passenger arrives at the airport, and whether or not the passenger has checked baggage. Flight characteristics are the expected on-time performance of the flight and the aircraft type flying the route. All of these are expected to influence the passengers' choice behavior.

Table 5-1: Relevant Surveyed Trip Characteristics

Category	Item	Alternatives/Entry
Pre-purchase	Trip Purpose	Business, Vacation, Visit friends/relatives, attend school/college, other
	Distribution Channel	Travel Agent, Airline, Third Party Online
	Ticket sponsor	Personally, company, family/friends
	Party Size	1,2,3,4,5, >5
	Preferred local arrival time	[hours, minutes, am/pm]
Itinerary Characteristics		
	Month	Jan-Dec
	Day of week	Su-Sa
	Nights away	1,2,3,4,5,6,7,8-14,15-21,>21
	Class of service	Economy, Business, First Class
	Airline	[Select from list]
	Departure Airport	[city, airport name, or airport code]
	Arrival Airport	[city, airport name, or airport code]
	Fare	[positive integer]
	Scheduled local departure time	[hours, minutes, am/pm]
	Scheduled local arrival time	[hours, minutes, am/pm]
	Itinerary type	Nonstop, connections [number]
	Connecting airport (if applicable)	[city, airport name, or airport code]
	Scheduled transfer time (if applicable)	[hours, minutes]
Pre-Flight	Access Mode	Auto, taxi, shuttle, bus, train
	Access time	[hours, minutes]
	Time at airport before departure	[hours, minutes]
	Checked baggage	Yes, no
	Food	Brought from home, bought on the way to airport, bought at airport, bought on the plane, complimentary meal on board
Flight	On-time Performance	Yes, no/hours, minutes
	Aircraft type	Wide body, standard jet, regional jet, propeller

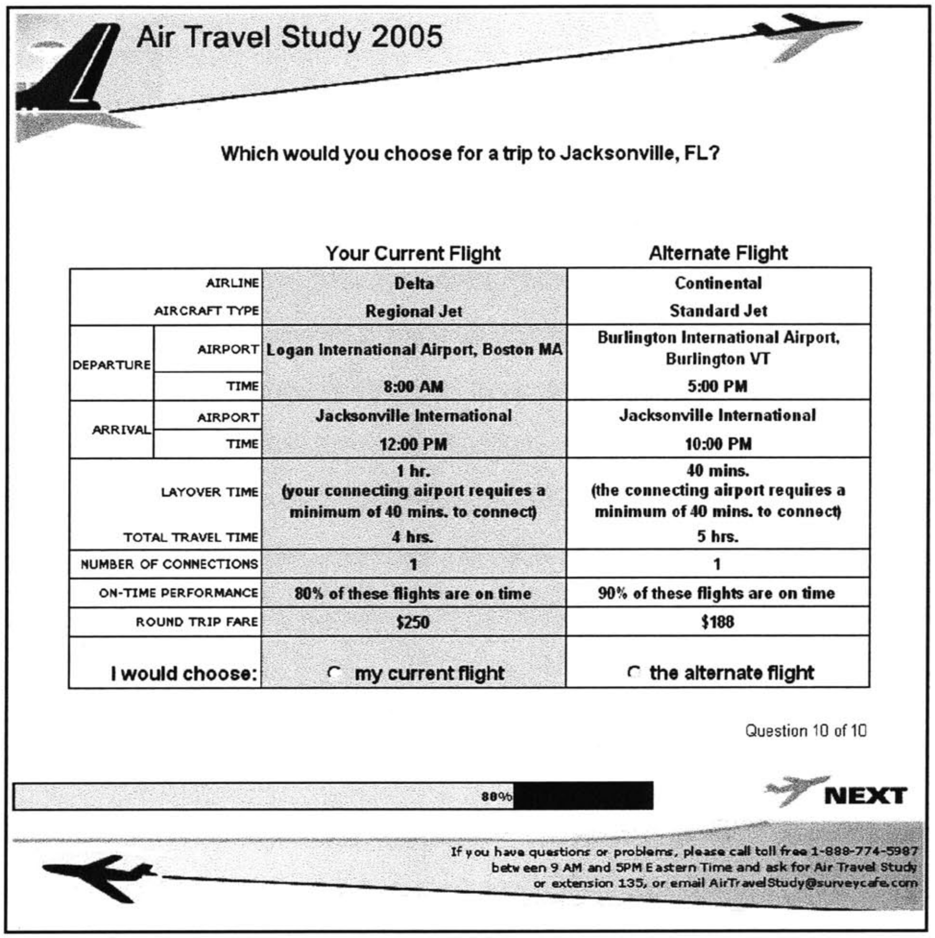
5.2.2 Preferences Regarding Specific Airports and Airlines

In the next section of the questionnaire, respondents were asked to reveal their preferences regarding specific airports and airlines. These preferences are used in the stated choice experiment to design the alternative itinerary. Respondents were first asked to rank the three airlines that they most prefer and to identify one that they least prefer based on their perception of quality, ignoring price. Secondly, respondents were asked to identify the departure airport closest to their home. The survey results are also used by other authors for airport choice models, hence the detailed data collection on airports. Up to five airports within a 150-mile radius of the identified airport were extracted from an airport database and displayed to the respondent. Respondents were asked to identify three airports they would consider using as alternatives to the recent trip departure airport, assuming that suitable flights were available from these airports. Respondents were allowed to choose from the displayed airports or identify other alternative airports. Once they had chosen three airports, they were asked to rank these and the departure airport of the recent trip from most preferred to least preferred. In addition, they were asked to provide their access times to the three alternate airports.

5.2.3 Stated Preference Experiment

The stated preference experiment is the core of the survey. Respondents were given eight pairs of flight itineraries from which to choose one each (see Figure 5-1). The first itinerary is the one representing the recent trip of the respondent, based on the information provided in the first section of the survey. The information provided to the respondent for the first itinerary is the name of the airline, aircraft type, departure and arrival airports, departure and arrival times, layover time, total travel time, number of connections, and fare as reported. In addition to the reported connecting time, the respondent was informed about the minimum connecting time as published by his connecting airport. The minimum connecting time is the time an airport publishes as the minimum feasible time needed for a connection. An itinerary has to have at least this minimum connecting time, but often has longer connecting times. This data is based on the listings available in the OAG database (OAG, 2004). On-time performance is added as an attribute with levels of 50, 60, 70, 80, or 90% on-time performance for the recent trip. This was done to avoid a bias toward the recent trip, where the on-time performance outcome is actually known to the respondent.

Figure 5-1: Screen Shot of Stated Preference Survey



The second alternative was designed using the attributes and levels that are summarized in Table 5-2. The design of the second alternative was random, with one level randomly picked for each attribute, unless otherwise described below. Airline and departure airport were drawn from the ranking list provided by the respondent.

Table 5-2: Stated Preference Experiment Attributes and Levels

Attributes	Levels
Airline	1st, 2nd, 3rd, last ranked
Departure airport	1st, 2nd, 3rd, 4th ranked
Aircraft type	wide body, standard jet, regional jet, propeller
Fare	-50, -25, +-0, +25, +50 % of fare of recent trip
On-time performance	50, 60, 70, 80, 90 %
Number of connections	0, 1, 2
Minimum connecting time	30, 40, 50, 60 min
Buffer Time	0. 15, 30, 45, 60, 90, 150, 210 min
Arrival Time Deviation	-2, -1, 0, 1, 2 hrs vs. preferred arrival time as provided

Aircraft type levels were customized based on the travel time of the recent trip. For trips shorter than 2.5 hours, aircraft type was randomly drawn from the set {standard jet, regional jet, propeller}. For trip lengths of at least 2.5 hours but shorter than four hours the aircraft type was randomly drawn from the set {regional jet, standard jet}, and for itineraries of at least 4 hours the aircraft type was drawn from the set {standard jet, wide body}. The fare was varied relative to the recent trip's fare of the respondent, and on-time performance was again varied from 50 to 90%.

If the recent trip was shorter than 4 hours, the number of connections was drawn from {0,1}. For trips that were at least 4 hours long, the number of connections was randomly drawn from the set {0,0,1,1,2}, i.e. a nonstop itinerary or a one stop connection had a 40% probability of being chosen, whereas a two stop connection had a 20% chance of being chosen. This was done since two stop connections occur less frequently in U.S. domestic trips.

The minimum connecting time was randomly varied between 30 and 60 minutes. The attribute "buffer time" resembles the additional layover time. Adding the minimum connecting time and the buffer time results in the connecting time or layover time of the itinerary. Finally, the "arrival time deviation" attribute reflects the deviation from the preferred arrival time as reported by the respondent at the outset of the survey.

The attributes described above where used as input factors to the displayed itinerary. For airline and departure airport, the ranks were replaced by the names of the airline and airport. The fare level was used as a percentage deviation from the fare of the previous trip in order to calculate the fare of the alternative itinerary. On-time performance, number of connections and minimum connecting time levels were presented without any further calculations. The buffer time was added to the minimum connecting time level and was displayed as "layover time" of the alternative. Finally, the arrival time deviation was used to calculate the arrival time of the alternative. Given that arrival time, the departure time of the itinerary was calculated. Arrival and departure times are shown in local time, i.e. in case there were time zone differences between departure and arrival airports, these were accounted for.

5.2.4 Rating Exercise

The rating exercise is designed as part of this dissertation and is set up to collect the following attitudinal characteristics:

- The willingness to rush vs. a shorter scheduled elapsed time
- The willingness to risk misconnection vs. shorter scheduled elapsed time
- The perception of an airline's trustworthiness in providing reliable connections

Questions were set up with a 5 point likert scale (see Figure 5-1), with ratings from "strongly disagree", "somewhat disagree", "neither agree or disagree", "somewhat agree" to "strongly agree". In the questionnaire, the statements were shuffled to disguise the underlying expected attitudinal characteristics.

Figure 5-2: Rating Exercise (for overview purposes only, actual survey shuffled)

	strongly disagree				strongly agree
[1. Willingness to rush through airport vs. shorter scheduled elapsed time]					
I like to take my time when connecting between flights	o	o	o	o	o
It's hard for me to find my way through airports	o	o	o	o	o
I don't think time at airports is wasted because I can shop, eat, or work at airports	o	o	o	o	o
I don't mind being rushed at a connecting airport if this means I'll arrive at my final destination earlier	o	o	o	o	o
I enjoy time having extra time at airports	o	o	o	o	o
[2. Willingness to take risk of missing flight vs. shorter scheduled elapsed time]					
I usually arrive at the check-in counter just before the check-in deadline	o	o	o	o	o
Catching my scheduled connecting flight is of great importance to me	o	o	o	o	o
I try to avoid short connections because of the risk of either me or my luggage missing the connecting flight	o	o	o	o	o
Given two itineraries that only differ in connecting time, I always choose the one with shorter connecting time	o	o	o	o	o
I'm willing to accept the risk of a missed connection if this gets me earlier to my destination most of the time	o	o	o	o	o
[3. Perception of airlines' trustworthiness in providing reliable connections]					
Airlines only sell connections that they expect passengers could make	o	o	o	o	o
Airlines sometimes underestimate the time needed to connect between flights	o	o	o	o	o
It is the passenger's responsibility to plan for a sufficient transfer time when booking a connecting itinerary	o	o	o	o	o
I make sure that the planned connecting time is adequate for me when booking a connecting itinerary	o	o	o	o	o

5.2.5 Socio-Demographics

The collected socio-demographics are listed in Table 5-3. In addition to collecting information on gender, age, employment status and type, and household income, several items specific to the survey topic were collected. First, respondents were asked about the number of U.S. domestic air trips they made in the previous year. Secondly, they were asked to name their membership level in the frequent flyer program for all airlines that they had ranked earlier in the questionnaire. Finally, respondents were asked whether they had missed a connecting flight in the previous two years.

Table 5-3: Collected Socio-Demographics

Gender	Female, Male
Age	18-21, 22-29, 30-39, 40-49, 50-59, 60-69, 70+
Employment status	Full-time, part-time, self-employed, student, retired, homemaker, unemployed
Employment type	Clerical/secretarial, executive/managerial, professional/technical, sales/buyer, teacher/professor, retail/service, other non-office worker
Industry	List of industry types
Household's annual income I before taxes	$I<10$, $10 \leq I<20$, $20 \leq I<30$, $30 \leq I<40$, $40 \leq I<50$, $50 \leq I<75$, $75 \leq I<99$, $100 \leq I$
How many air trips have you made in the last year that were in the U.S.?	[Number of U.S. domestic trips]
Membership in Frequent Flyer Program	For all ranked airlines: ∞ Not a member ∞ Standard level ∞ Elite 1st level ∞ Elite Highest Level
Missed connecting flight in the past two years	Yes, no

5.3 Sampling and Survey Execution

The survey was administered in May 2005. Subscribers of the survey tool "surveycafe.com" were invited to take part in the Air Travel Study if they had made a paid U.S. domestic air trip within the past 12 months. If they did not qualify but someone else in their household did, this person was allowed to answer the survey. Based on this sampling procedure, the survey is not a probability sample, meaning that statements cannot be considered representative for the whole population. However, the survey can be used for modeling and for conditional statements for specific socio-demographic groups or respondents with certain trip attributes.

All information was collected via the internet tool "surveycafe.com". In total, 5331 registered users of surveycafe.com were e-mailed invitations to participate in the survey, and the entries were accepted on a first come first serve basis until the target size was reached. As an incentive, all respondents received a $7 voucher for a dessert.

5.4 Data Cleaning

Of the 621 respondents, two had input insufficient data and were eliminated from further analysis. The median survey duration was 28 minutes. Survey durations under 10 minutes were considered unreliable and were excluded from further analysis. The answers to the rating exercise were used for an additional check whether respondents gave sincere answers. Since some of the statements contradict each other, it would have been expected to have different rankings for different questions. Data of respondents who answered with the same level of agreement for all 14 rating questions (e.g. strongly agree) was eliminated from further analysis. After data cleaning, 583 cases with 4664 stated preference responses remained in the data set.

5.5 Summary Statistics

In the following section, descriptive statistics of the surveyed data are presented. In the next sub-section, the trip characteristics of the respondents' recent trip are summarized. Following, socio-demographics and answers to the rating exercise are presented.

5.5.1 Trip Characteristics

The trip characteristics of the respondents are outlined in Table 5-4. Of the reported trips, 18% were for business purposes, and 75% for either vacation or visiting friends or relatives.

73% of respondents booked their ticket online, on either the airline's website (48%) or on an independent online travel portal (25%). This figure is higher than that of the total U.S. population. Nielsen/NetRatings reports that "nearly 50%" of airline ticket sales in the U.S. have been conducted online in the first half of 2005. (Nielsen/NetRatings 2005). Nielsen also reports that of these 50%, half are conducted via online travel agencies and the other half via airline websites. For domestic itineraries, Air Canada reports that in Q2 2006, 50% of tickets are sold via the airline's website and another 12% via call centers, numbers that are very similar to those seen in the survey (Air Canada, 2007).

Of all respondents, 44% traveled alone, and another 32% as a party of two. 48% of respondents spent four nights or a fewer number of nights away from home for this trip. 77% of respondents checked baggage for their flight. This compares to a value of 70% of customers checking bags of a major North American carrier on its domestic itineraries (due to competitive nature of the business, airline name cannot be mentioned).

Of special interest in this study is the number of connections of the recent trip (see Table 5-5). Relative to all respondents, 68.5% had no connections, 28.6% had one connection, 2.7% had two connections, and 0.2% had three connections. In comparison, of all U.S. domestic travelers, 65.7% had no connections, 31.6% one connection, 2.4% two connections, and 0.2% three connections, based on 2004 data from the U.S. DOT (DB1B Market coupon data).

Table 5-4: Selected Surveyed Trip Characteristics

What was the primary reason you went from x to y?	Frequency	Percent
Business	107	18.4
Vacation	178	30.5
Visit friends or relatives	261	44.8
Attend school/college	1	.2
Other	36	6.2
How did you acquire your ticket?		
Travel agent	72	12.3
Directly from the airline (call center or ticket counter)	73	12.5
Online - using an airline's web site	278	47.7
Online - from a travel site not hosted by an airline	146	25.0
Other	14	2.4
Number in travel party		
1 (I traveled alone)	255	43.7
2 (I traveled with one other person)	189	32.4
3 (I traveled with two other people)	54	9.3
4 (I traveled with three other people)	55	9.4
5 (I traveled with four other people)	6	1.0
6 or more (I traveled with 5 or more other people)	24	4.1
What class of service did you use?		
Economy or coach	526	90.2
Business	39	6.7
First class	18	3.1
How many nights were you away on your trip		
0 (left and returned the same day)	6	1.0
1 night	13	2.2
2 nights	71	12.2
3 nights	107	18.4
4 nights	81	13.9
5 nights	67	11.5
6 nights	52	8.9
7 nights	74	12.7
8 to 14 nights	72	12.3
15 to 21 nights	12	2.1
22 nights or more	28	4.8
Did you check any baggage for your flight		
Yes	446	76.5
No	137	23.5

Table 5-5: Level of Service Characteristics of Respondents' Reported Recent Trip

Did you have a nonstop flight?	Frequency	Percent
Yes, nonstop flight	363	62.3
No, made one or more stops or connections	220	37.7
Total	583	100.0
How many connections did you have?		
0	36	6.2
1	167	28.6
2	16	2.7
3	1	.2
Total	220	37.7

5.5.2 Socio-Demographics

Of all respondents, 61% were female. Most of respondents' households had an annual income of $50,000 to $74,999, with 22.3% having an income of $100,000 or more. 89% of all respondents are between the ages of 22 and 59, and 8% are 60 years or older. Finally, 13% of respondents confirmed having missed a connecting flight in the previous two years.

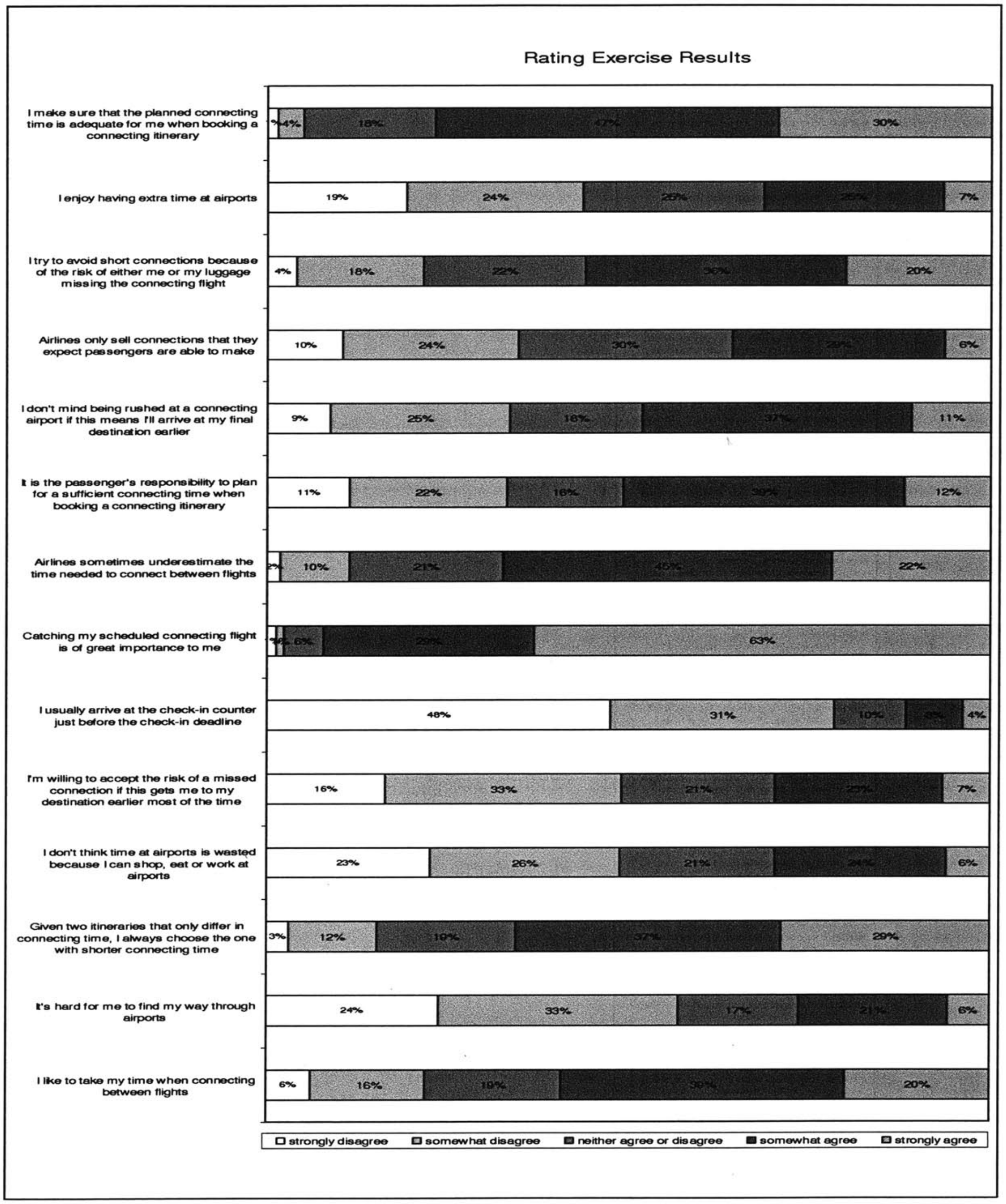
Table 5-6: Selected Socio-Demographics Summary Statistics

	Frequency	Percent
Gender		
Female	355	60.9
Male	228	39.1
Annual household income		
Under $10,000	7	1.2
$10,000 to $19,999	18	3.1
$20,000 to $29,999	38	6.5
$30,000 to $39,999	56	9.6
$40,000 to $49,999	69	11.8
$50,000 to $74,999	157	26.9
$75,000 to $99,999	108	18.5
$100,000 or more	130	22.3
Age		
18 to 21	17	2.9
22 to 29	99	17.0
30 to 39	164	28.1
40 to 49	161	27.6
50 to 59	94	16.1
60 to 69	38	6.5
70+	10	1.7
Have you missed a flight at a connecting airport in the last two years?		
Yes	76	13.0
No	507	87.0

5.5.3 Rating Exercise

The results of the rating exercise are summarized in Figure 5-2. 77% of all respondents somewhat or strongly agree that they make sure that the connecting time is adequate when booking their itinerary, and 56% of respondents somewhat or strongly agree that they avoid short connections because of the misconnection risk. On the other hand, 48% somewhat or strongly agree that they do not mind being rushed at a connecting airport. 92% of respondents somewhat or strongly agree that catching their connecting flight is of great importance to them. All results can be viewed in Figure 5-3 (see also histograms in Appendix C).

Figure 5-3: Results of Rating Exercise



5.6 Conclusion

This is the first survey of airline passenger choice that explicitly focuses on connecting flights. By designing the survey to include connecting time and minimum connecting time, the respondent is given an anchor to compare the connecting time to. The rating exercise allows the inclusion of attitudes toward different connecting times in choice and latent variable models.

The descriptive statistics of the rating exercise show that more than three quarter of respondents take into account the connecting time when making their itinerary choice. The respondents are approximately split evenly between those trying to avoid short connecting times and those that do not mind being rushed when connecting. The modeling results described in the next chapter will extract the latent variables from the indicators in the rating exercise and connect socio-demographics with those latent variables, thereby explaining how these socio-demographics impact attitudes and airline itinerary choice.

6 Results

In this chapter, the specification search and modeling results are described. The chapter starts with the choice model specification and results. The choice model is the first building block of the Integrated Choice and Latent Variable Model. In the second section (6-2), the latent variable (MIMC) model is introduced. In the third section (6-3), the results of the integration of the choice and the latent variable model (ICLV) are detailed.

6.1 Choice Model

6.1.1 Introduction

In random utility maximization, the value of utility is a random variable. The utility an individual allocates to a certain alternative is modeled with a deterministic and a random part. The utility U_{in} of an alternative i for individual n depends on the attributes of the alternative and the socio-demographic characteristics of the individual (X_{in}) and an error term (ε_{in}). The error term is assumed to be i.i.d. extreme value distributed. The vector of unknown parameters β is estimated. The structural equation [6-1] is comprised of the described deterministic and random part. The measurement equation [6-2] reveals the choice. It is 1 if the respective alternative was chosen and 0 otherwise, assuming utility maximization.

Structural Equation:

$$U_{in} = X_{in}\beta + \varepsilon_{in} \qquad [6\text{-}1]$$

Measurement Equation:

$$y_{in} = \begin{cases} 1, \text{ if } U_{in} = \max\limits_{j}\{U_{jn}\} \\ 0, \text{ otherwise} \end{cases} \qquad [6\text{-}2]$$

where:

n denotes an individual, $n = 1, \ldots, N$.

i,j denote alternatives, $i, j = 1, \ldots, J$.

y_{in} is the choice indicator (equal to 1 if alternative i is chosen, and 0 otherwise).

U_{in} is the utility of alternative i as perceived by individual n.

X_{in} is a $(1 \times K)$ vector of explanatory variables describing n and i.

β is a $(K \times 1)$ vector of unknown parameters.

ε_{in} are i.i.d. extreme value random variables.

The individual choice probability is then

$$P(y_{in} = 1 \mid X_n; \beta) = \frac{e^{(X_{in}\beta)}}{\sum_{j \in C_n} e^{(X_{jn}\beta)}} \qquad [6\text{-}3]$$

where :

C_n is the choice set faced by individual n, comprised of J_n alternatives

Numerous models were estimated in the specification search (described in the next section), using the attributes available in the survey results as outlined in Chapter 5. Figure 6-1 shows the final model specification and Table 6-1 displays the estimation results.

6.1.2 Choice Model Specification and Estimation Results

The alternative specific constant for the current trip can be explained by various factors. First, the current trip is a reported actual trip, i.e. the respondent does not associate any uncertainty with this trip. There may also be a justification of past behavior present in the SP exercise or an inertia effect, reflecting a bias toward the current trip. The alternative specific constant may also capture attributes that are not included in the variables.

In the specification search, fare was first treated linearly and then as a log transform. The latter led to a better model fit, representing decreasing marginal returns with higher fare and implying a lower sensitivity to absolute fare changes at higher fare levels.

Frequent flyer program membership was available at four levels in the survey data, top-elite, base-elite, standard and none. The frequent flyer program membership included here is in respect to the airlines displayed in the alternatives. Even if the characteristic of the traveler in both alternatives is "none", the respondent could be a frequent flyer program member of an

Figure 6-1: Choice Model Specification

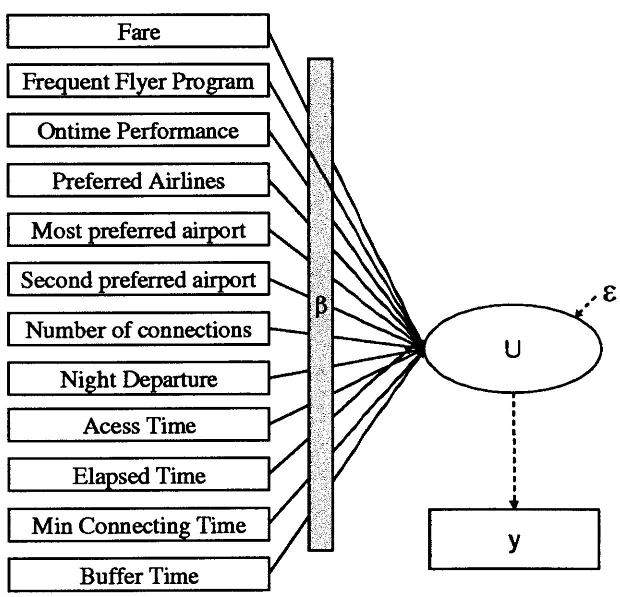


Table 6-1: Estimation Results for Choice Model

Parameter	Coefficient Value	Standard Error	T-test
Alternative specific constant current trip	1.0216	0.0608	16.8
ln (fare)	-3.978	0.133	-29.9
Elite frequent flyer program member	0.472	0.197	2.4
Standard frequent flyer program member	0.201	0.0973	2.1
On-time performance in percentage points	0.0117	0.00253	4.6
Preferred airlines	0.148	0.0668	2.2
First airport choice	0.800	0.0933	8.6
Second airport choice	0.305	0.0972	3.1
Number of connections	-0.488	0.133	-3.7
Night departure	-0.359	0.177	-2.0
Access time in min	-0.00432	0.000775	-5.6
Elapsed time in min	-0.00319	0.000468	-6.8
Minimum Connecting Time in min	-0.00596	0.00300	-2.0
Buffer time < 15 min in min	0.0151	0.00626	2.4
Buffer time 15-59 min in min	-0.00754	0.00333	-2.3
Buffer time > 60 min in min	-0.00186	0.00141	-1.3

Number of estimated parameters:	16
Number of observations:	4664
Final log-likelihood:	-1656.37
Adjusted rho-square:	0.4827

airline that is not part of the choice set. As part of the specification search, constants were estimated for all four levels, with no membership being the base level. The results indicate no significant difference between a model with all four levels versus one in which the first two levels are combined to a joint level called "elite frequent flyer program member." Therefore, only three levels were used for subsequent modeling and estimation.

On-time performance is described by a percentage value depicting on-time departures within 15 minutes of schedule. A preliminary test with dummy variables for the different attribute levels (50%, 60%, 70%, 80%, 90% on-time performance) showed that on-time performance can be modeled in a linear fashion.

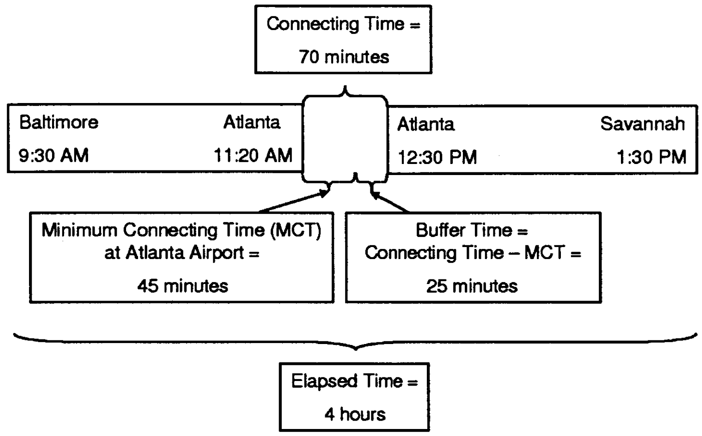
Of the four airports ranked by the respondent, only the first two were significantly different, with the third and fourth airport being the base. The estimation values show that airport is far more important than airline, which intuitively makes sense. Even though access time is captured separately, other attributes of airports such as ground transport, parking facilities or lounge access could be relevant in this regard.

The number of connections of an itinerary was initially modeled as three constants (zero connections, one connection, two connections). Another estimation run showed that a model with number of connections modeled linearly was not significantly different from the first model. The disutility associated with the length of the connecting time is modeled separately and is described below.

Night departures are defined as departures between midnight and 5am. This is the time window in which currently in the industry almost no flights depart, as it is perceived as being highly disadvantageous to the customer, both in terms of time of day and access to the airport. As expected, the coefficient has a negative sign, revealing the disutility of night departures. Access time is the time needed from home to the specific airport of the alternative.

Time from departure airport to arrival airport is captured by various coefficients. Figure 6-2 shows a detailed description of the time windows an itinerary is comprised of.

Figure 6-2: Time Windows of an Example Itinerary



A connecting itinerary consists of various elements. The total time from departure of the first flight to arrival of the last flight is defined as "elapsed time". In the estimation results, elapsed time has a value closer to zero than access time, which can possibly be explained by the higher discomfort associated with the trip to the airport vs. the trip from origin airport to destination airport.

The time from arrival of the first flight to departure of the second flight is defined as connecting time. In the model, connecting time is divided into two subelements, the minimum connecting time specific to the connecting airport and the additional time, defined as "buffer time". Minimum connecting time has a stronger effect than elapsed and access time. This could be due to the fact that the minimum connecting time is a proxy for the size of the connecting airport, with higher minimum connecting times equivalent to larger airports. The sum of minimum connecting time and buffer time produces the connecting time. In the survey, respondents were given the elapsed time, connecting time and minimum connecting time and had to infer the available buffer time. The details of the buffer time estimation will be described in the next section.

In addition to the time-oriented aspects discussed above, estimation runs were performed with "schedule delay" and "time of day departure time" components. "Schedule delay" is defined as the difference between the preferred arrival time and the actual arrival time. "Time of Day" was modeled by using three coefficients, one of them being the departure time, the second one the squared departure time, and the third one the cubed departure time. Both approaches, schedule delay and time of day departure time, didn't improve the model fit. This might be caused by the limited variability of departure and arrival times in the data set. Interactions were tried between trip purpose and buffer time, however, these produced insignificant coefficient estimates for the interactions and did not significantly improve the model fit.

One may hypothesize that frequency of flights for the flight leg from hub to destination could have an influence on choice. In the stated preference experiment, neither the hub location nor the frequency of outbound flights was given, so the hub location could only be inferred from the airline choice and the frequency from the hub-destination pair by well versed travelers. Regardless of the second leg's frequency however, a misconnection creates a burden of uncertainty and transaction costs. Customers often need to see an overcrowded service counter, later flights may be fully booked and checked baggage may not be rerouted properly on subsequent flights. Based on these points, it is believed that the frequency of the second leg has limited influence on choice.

Aircraft type (wide body jet, narrow body jet, regional jet, prop aircraft) was included in preliminary model specifications and estimations. However, no aircraft type related coefficient could statistically significantly improve the model fit and thus aircraft type is disregarded in the final specification. Compared to previous choice models in the literature (e.g. Adler, 2005), this shows that passengers might have become accustomed to flying in regional jets and prop airplanes, which in previous models were perceived negatively versus larger aircraft types.

6.1.3 Alternative Specifications for Buffer Time

In order to test the hypothesis regarding buffer time described in chapter 4, buffer time was modeled in various ways. Buffer time was modeled linearly and several different piece-wise linear approximations were tested. In addition, two power transformations were tested.

In order to test our n-shaped disutility curve hypothesis for buffer times, a linear parameter for buffer time would be insufficient. In the linear specification [6-4], only one parameter is estimated. We split the range of buffer times into separate parts, leading to the estimation of 3 separate coefficients. Below is an example for cutoffs at 30 minutes and 75 minutes [6-5 to 6-8].

Linear specification: $\beta_{buffer} \times buffer\ time$ [6-4]

Piecewise linear specification:

$$\beta_{buffer_{0\_30}} \times B_1 + \beta_{buffer_{30\_75}} \times B_2 + \beta_{buffer_{gt75}} \times B_3 \qquad [6\text{-}5]$$

with

$$B_1 = \begin{cases} t & if\ t < 30 \\ 30 & otherwise \end{cases} \qquad [6\text{-}6]$$

$$B_2 = \begin{cases} 0 & if\ t < 30 \\ t-30 & if\ 30'' \ t < 75 \\ 45 & otherwise \end{cases} \qquad [6\text{-}7]$$

$$B_3 = \begin{cases} 0 & if\ t < 75 \\ t-75 & otherwise \end{cases} \qquad [6\text{-}8]$$

With a log likelihood test between the restricted model [6-4] and the unrestricted model [6-5], the hypothesis can be tested that all coefficient values are equal.

A second type of specification is a parabola based one. In this case, the maximum of a negative parabola is hypothesized to be either at 15 or at 30 minutes, leading to the following specification:

$$(\beta_{buffer} - 15)^2 \qquad [6\text{-}9]$$

$$(\beta_{buffer} - 30)^2 \qquad [6\text{-}10]$$

Table 6-2 shows the estimation results for buffer time only for alternative models. Except for the changes regarding buffer time, the same model as in Table 6-1 was estimated. The other parameters of the model remained stable. The linear specification leads to a negative coefficient value, as expected. This reveals that when specified as one linear parameter, increased buffer time yields lower overall utility values.

The next three models have piecewise linear specifications. The first one has cut off values at 30 and 75 minutes, the second one at 15 and 60 minutes, and the third one at 15, 30, and 45 minutes.

Table 6-2: Alternative Buffer Time Specifications and Estimation Results (Buffer Time Parameters shown only)

	linear specification			piecewise linear with breaks at 30,75			piecewise linear with breaks at 15,60		
	Buffertime	value	(t-test)	Buffertime	value	(t-test)	Buffertime	value	(t-test)
	linear parameter	-0.00194	-2.0653	<30	0.00626	1.52	<15	0.0151	2.41
				30-74	-0.00698	-1.85	15-60	-0.00754	-2.26
				>75	-0.00190	-1.16	>60	-0.00186	-1.32
No. of parameters		14			16			16	
Final LL		-1660.91			-1658.56			-1656.37	
adjusted ρ^2		0.4819			0.4820			0.4827	

	piecewise linear with breaks at 15,30,45			parabola with peak at 30 min			parabola with peak at 15 min		
	Buffertime	value	(t-test)	Buffertime	value	(t-test)	Buffertime	value	(t-test)
	<15	0.0163	2.42	(Bftime - 30)^2	-1.775E-05	3.24	(Bftime - 15)^2	-1.403E-05	-2.89
	15-30	-0.0188	-1.57						
	30-45	0.0013	0.12						
	>45	-0.0025	-1.92						
No. of parameters		17			14			13	
Final LL		-1656.27			-1661.25			-1658.80	
adjusted ρ^2		0.4824			0.4818			0.4826	

For the first two specifications, we assume two breaks, i.e. three pieces and want to test which values lead to a better adjusted rho square value. Assuming a minimum connecting time of 45 minutes, these cut off values would resemble overall connecting times of 75 min and 120 min for the first model and 60 minutes and 105 minutes for the second model. All have in common that the first piecewise linear coefficient is positive, supporting the hypothesis about lower utility close to the minimum connecting time. An increase in buffer time leads to a higher utility value of the alternative. The second coefficient values in all three piecewise linear models are negative, resembling the fact that a further extension of the buffer time and thus connecting time reduces utility values.

The subsequent two models are specified with a power transformation. The rationale for this is an upside down parabola shaped utility for buffer time. The first model has the maximum at 30 minutes, while the second one has the maximum at 15 minutes. Both models yield significant coefficient values, however, the models are not superior based on the adjusted ρ^2 value versus the piecewise linear specification with cut off values at 15 and 60 minutes.

Figure 6-3 to Figure 6-8 are graphical representations of the piecewise linear buffer time coefficients. Utility values are the sum of buffer time and elapsed time, since an increase in buffer time also increases the elapsed time of a trip. The hypothesis about increasing utility levels close to the minimum connecting time cannot be discarded. Contrary to our hypothesis, however, the middle section of the buffertime curves does not show a plateau, depicting a window of indifference. After the first cutoff point, utility decreases. The smaller negative slope of the last part of the curve might be explained with decreasing values of absolute minutes at higher buffer time values.

Figure 6-7 and Figure 6-8 show the graphs of the parabola specification. Since the cumulative effect of additional buffer time is the sum of the buffer time coefficient and the elapsed time coefficient, the curves do not show the typical parabola graphical representation.

For further analysis, the model with cut off values at 15 and 60 minutes will be used, since based on the adjusted ρ^2 value it is the statistically superior specification.

Figure 6-3: Linear Buffer Time Specification

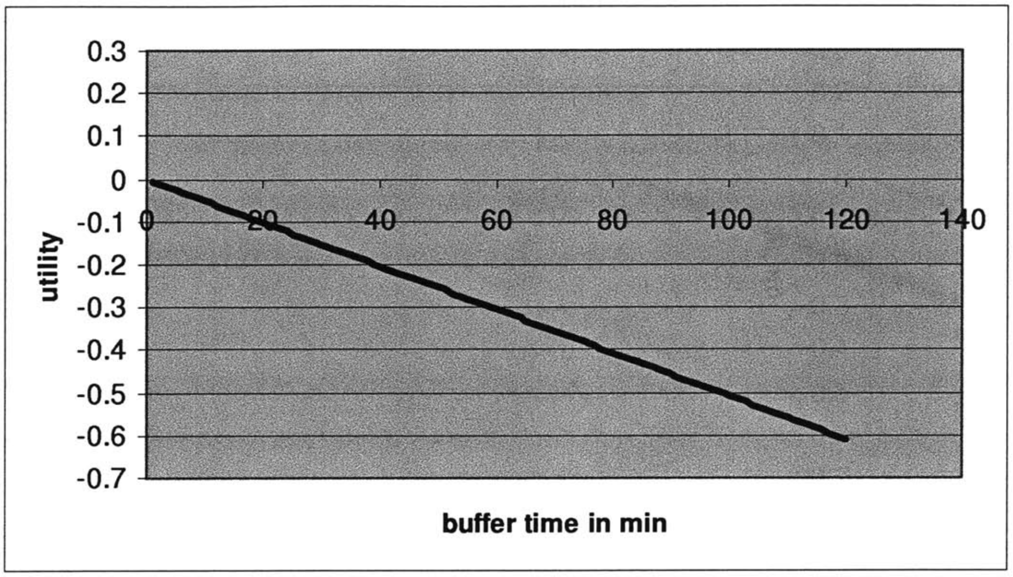


Figure 6-4: Piecewise Linear Specification at 30 and 75 minutes

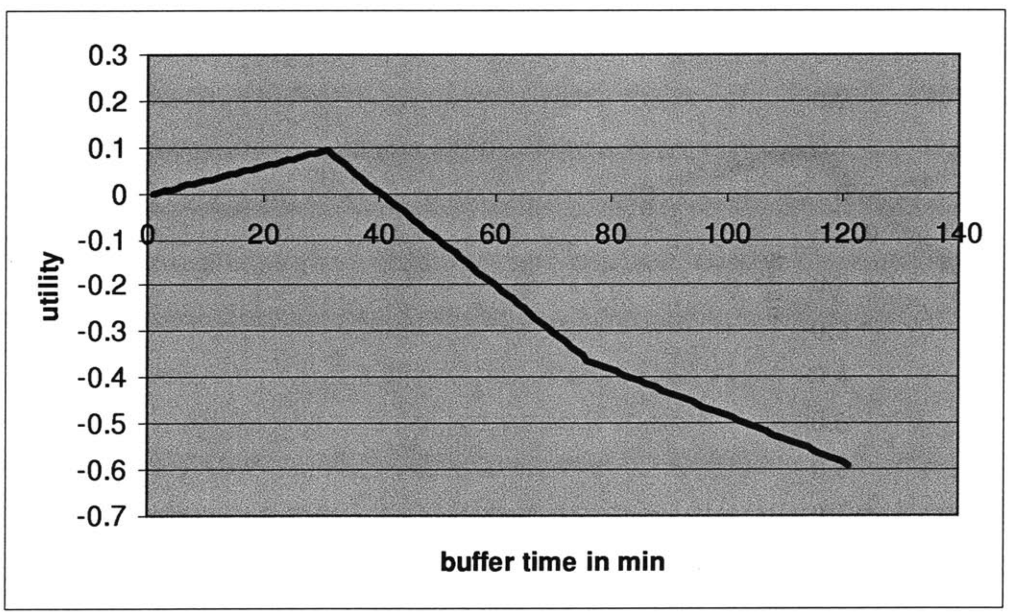


Figure 6-5: Piecewise Linear Specification at 15 and 60 minutes

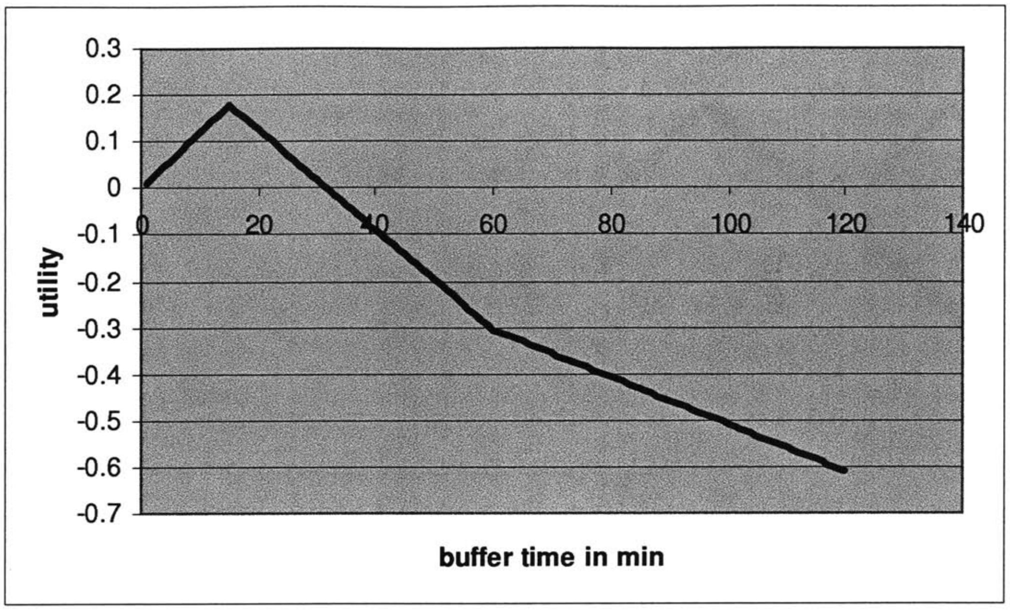


Figure 6-6: Piecewise Linear Specification at 15, 30 and 45 minutes

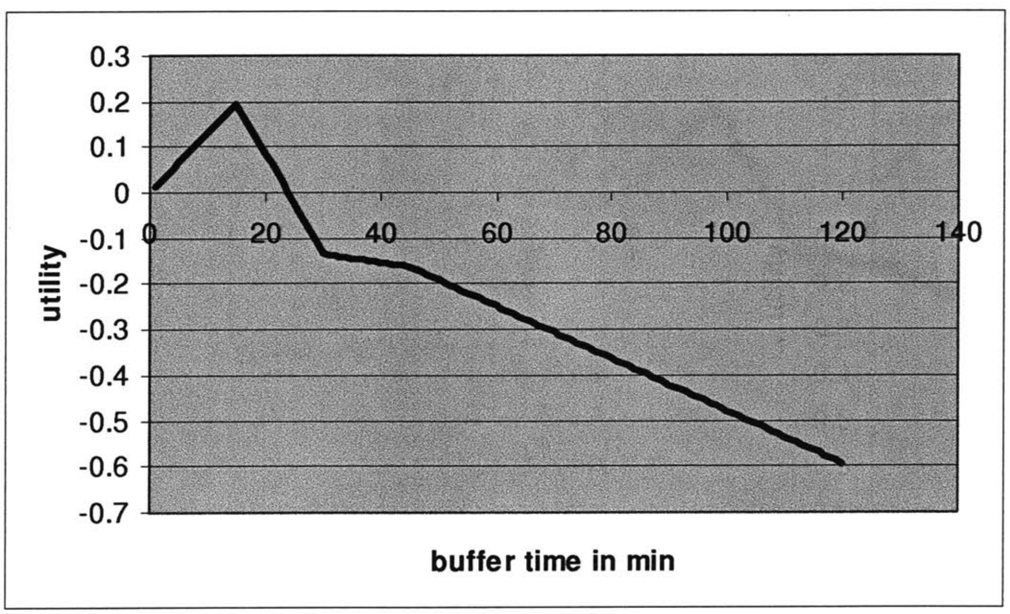


Figure 6-7: Parabola Specification with (Buffertime – 30 Minutes)^2

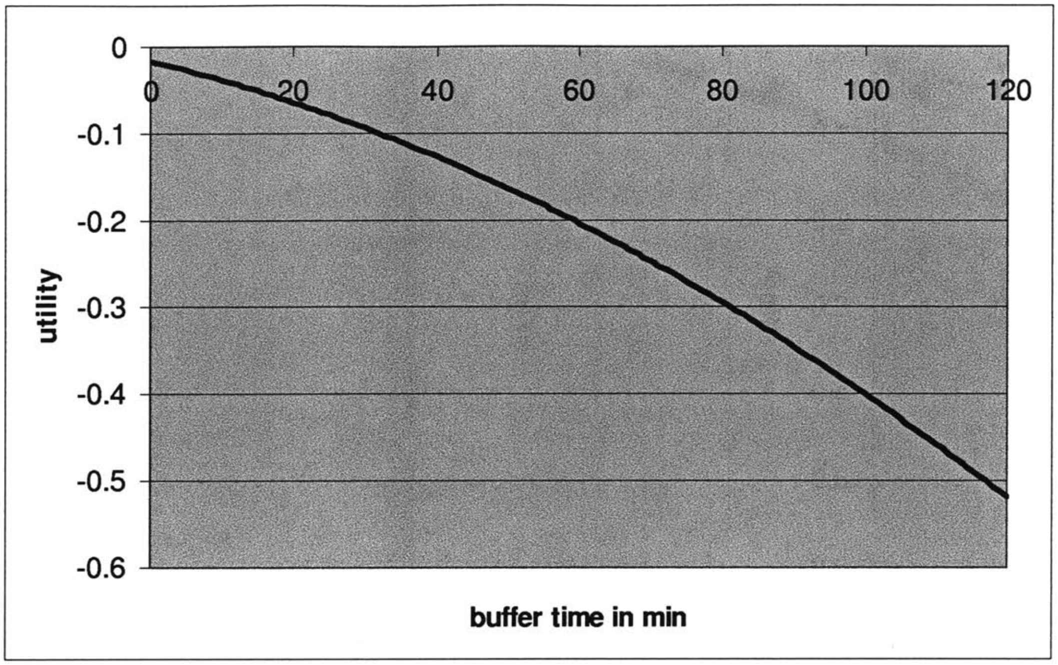
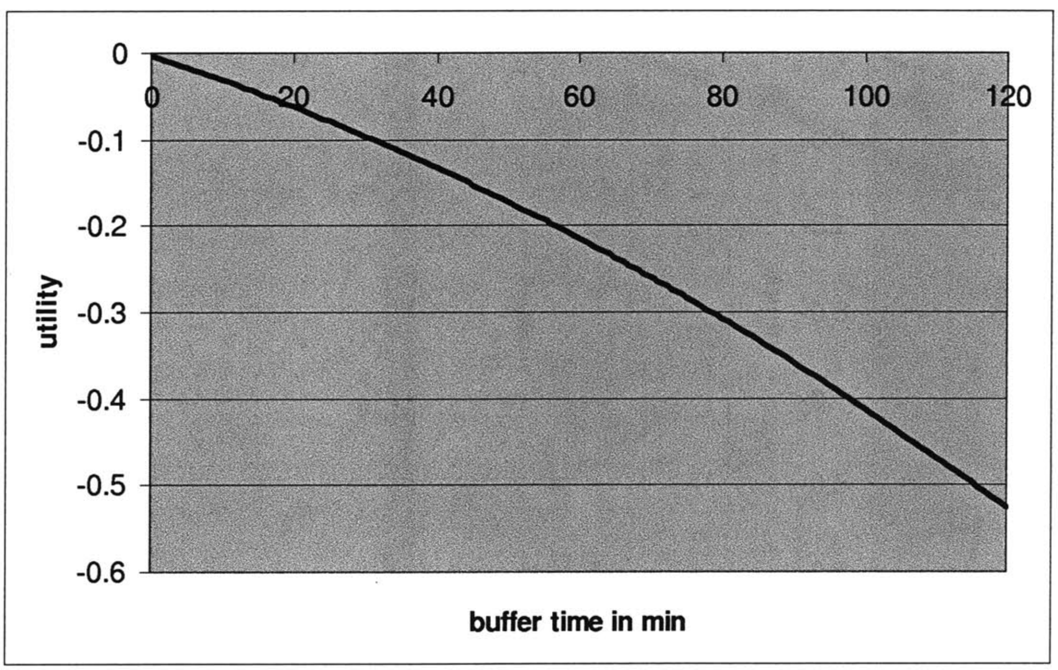


Figure 6-8: Parabola Specification with (Buffertime – 15 Minutes)^2



6.1.4 Willingness to Pay Values

Figure 6-9 shows the willingness to pay based on the final choice model (Table 6-1). The third column shows the willingness to pay in U.S. dollars, and the fourth column the willingness to pay in minutes.

Willingness to pay in $ is calculated by dividing the respective attribute coefficient by the fare coefficient and then multiplying this value by the average fare of $335 in order to account for the logarithmic fare coefficient. Willingness to pay in minutes is calculated by dividing the respective attribute coefficient by the coefficient value of elapsed time.

If the passenger is an elite member of the respective airline's frequent flyer program in the alternative, this is valued at $39.75 or 148 minutes of the passenger's time versus not being a member of the program. If the passenger is only a standard member, this value is reduced to $16.93 or 63 minutes. On time performance is valued at $9.85 or 37 minutes for every 10% increase in on time performance. Preferred airlines are the top two preferred airlines reported by the respondent. The value is relatively low at $12.46. However, since there likely is a correlation between elite membership in frequent flyer programs and preferred airlines, the cumulative value for such an airline would be $12.46 + $39.75 = $52.21. Values for first airport choice versus last, and second airport choice versus last, are in the expected magnitude and order with $67.37 for the first airport and $25.69 for the second airport. One connection has a negative value of $41.10. This value resembles the disutility of a connection as such, regardless of the time involved. Night departures are valued at -$30.23, revealing the inconvenience of nightly trips. Access time is valued higher than elapsed time ($21.83 vs. $16.12 per hour), representing possibly the discomfort associated with ground transport to the airport.

Fifteen minutes of additional buffer time, based on the minimum connecting time, are valued at $15.05. This means that on average, individuals would be willing to pay this amount in order to receive the additional buffer time. This is contrary to the traditional assumption that passengers discount longer connecting times.

Figure 6-9: Willingness to Pay Values in US$ for Average Fare of $335 and in Min of Elapsed Time

	in US$	in min
Airline where passenger is Elite frequent flyer program member	39.75	148
Airline where passenger is standard frequent flyer program member	16.93	63
On-time performance increase of 10%	9.85	4
Preferred airlines vs. last	12.46	46
First airport choice vs. last	67.37	251
Second airport choice vs. last	25.69	96
Number of connections	-41.10	-153
Night departure (midnight to 5am)	-30.23	-113
Access time in min	-0.36	-1.4
Elapsed time in min	-0.27	-1.0
Minimum Connecting Time in min	-0.50	-1.9
Buffertime < 15 min in min	1.27	4.7
Buffertime 15-59 min in min	-0.63	-2.4
Buffertime > 60 min in min	-0.16	-0.6

The results show that the hypothesis that additional buffer time from 0 to 15 minutes has a positive utility and thus a positive willingness to pay cannot be discarded. This is counterintuitive to classic microeconomics, where additional required time would lower utility based on value of time aspects. In the next section, latent constructs (attitudes toward risk, rush and trust in airlines' scheduling reliability) will be introduced to better explain the underlying mechanisms that lead to a higher utility with higher buffer time and to capture the heterogeneity of this effect amongst respondents.

6.2 Multiple Indicator Multiple Causes Model (MIMC)

The Multiple Indicator Multiple Causes Model (MIMC) is the second building block of the Integrated Choice and Latent Variable Model. With the Multiple Indicator Multiple Causes Model, latent constructs are introduced into the model to better understand and capture attitude based heterogeneity in the model.

6.2.1 Model Specification

In the Multiple Indicator Multiple Causes Model, latent variables are extracted from indicators and explained with socio-demographic characteristics of the respondents. The model consists of a measurement equation and a structural equation.

The purpose of the measurement equation is to extract latent variables from the indicators represented by the rating exercise. The indicators are displayed in Figure 6-10 and the framework of the final specification is displayed in Figure 6-11. The measurement model consists of a system of 14 equations (R=14), one for each indicator variable, extracting three latent variables (L=3). Based on preliminary analyses, one indicator (I 7) from the original rating exercise was constrained to 0. The specification of the measurement equation was based on a priori assumptions and the design of the rating exercise

Measurement Equation:

$$I_n = \kappa + AX^*{}_n + \nu_n, \; \nu_n \sim \mathrm{N}\,(0, \Sigma_\nu \text{ diagonal}) \qquad [6\text{-}11]$$

with

I_n = vector of indicators of latent variables of individual n (R x 1)

κ = vector of intercepts (Rx1)

A = matrix of unknown parameters (R x L)

X^*_n = vector of latent explanatory variables of individual n (L x 1)

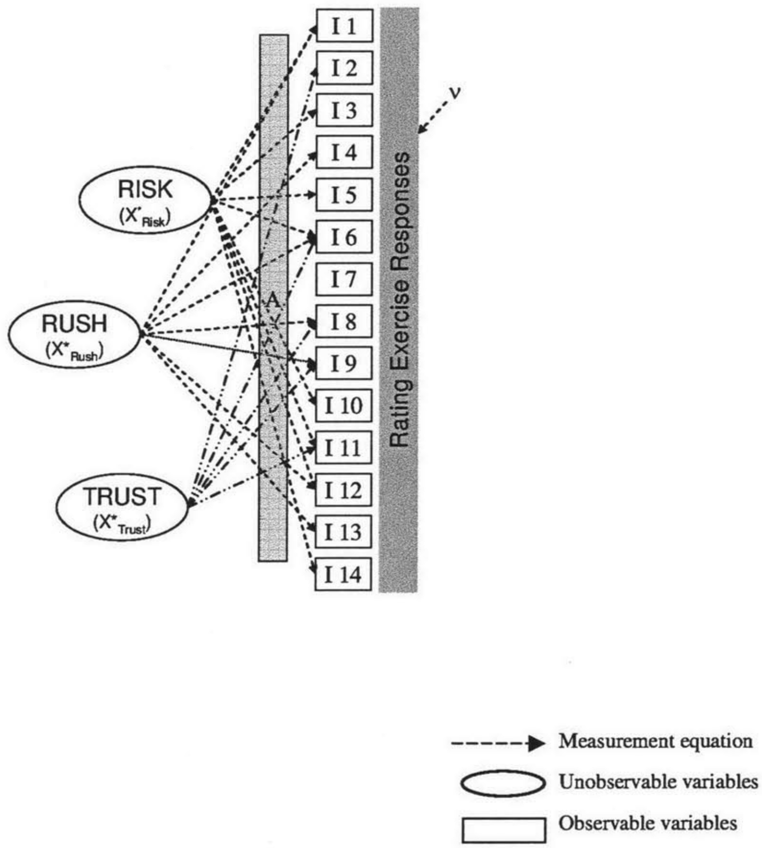
ν_n = vector of disturbances of individual n (R x 1)

The error vector is assumed to be normally distributed. Covariances in the error term are set to 0 based on an assumption of conditional independence, i.e. all correlation between indicators is assumed to be explained by the latent variable, resulting in a diagonal matrix with R variances. In addition, we assume indicators to be on a continuous scale. Discrete indicators (based on the 5 point likert scale used) would have increased the number of parameters which, with given observations, was judged not to be sufficient. A constant vector is included in the measurement equation since indicators are in their absolute form, not in deviation form.

Figure 6-10: Indicators for MIMC Measurement Equation

I 1	I like to take my time when connecting between flights
I 2	It's hard for me to find my way through airports
I 3	Given two itineraries that only differ in connecting time, I always choose the one with shorter connecting time
I 4	I don't think time at airports is wasted because I can shop, eat, or work at airports
I 5	I'm willing to accept the higher risk of a missed connection if this gets me to my destination earlier most of the time
I 6	I usually arrive at the check-in counter just before the check-in deadline
I 7	Catching my scheduled connecting flight is of great importance to me
I 8	Airlines sometimes underestimate the time needed to connect between flights
I 9	It is the passenger's responsibility to plan for a sufficient connecting time when booking a connecting itinerary
I 10	I don't mind being rushed at a connecting airport if this means I'll arrive at my final destination earlier
I 11	Airlines only sell connections that they expect passengers are able to make
I 12	I try to avoid short connections because of the risk of either me or my luggage missing the connecting flight
I 13	I enjoy having extra time at airports
I 14	I make sure that the planned connecting time is adequate for me when booking a connecting itinerary

Figure 6-11: Specification for Measurement Equation of MIMC Model



In figure 6-12, I_r are the R indicators from the rating exercise (see Figure 6-10) and X^* are the L=3 latent variables:

X^*_1 = Risk

X^*_2 = Rush

X^*_3 = Trust

Figure 6-12: Specification of Measurement Equation

$$\begin{pmatrix} I1 \\ I2 \\ I3 \\ I4 \\ I5 \\ I6 \\ I8 \\ I9 \\ I10 \\ I11 \\ I12 \\ I13 \\ I14 \end{pmatrix} = \begin{pmatrix} \kappa_1 \\ \kappa_2 \\ \kappa_3 \\ \kappa_4 \\ \kappa_5 \\ \kappa_6 \\ \kappa_8 \\ \kappa_9 \\ \kappa_{10} \\ \kappa_{11} \\ \kappa_{12} \\ \kappa_{13} \\ \kappa_{14} \end{pmatrix} + \begin{pmatrix} \alpha_1 & \alpha_2 & 0 \\ \alpha_3 & \alpha_4 & \alpha_5 \\ \alpha_6 & 0 & 0 \\ 0 & \alpha_7 & 0 \\ \alpha_8 & 0 & 0 \\ \alpha_9 & \alpha_{10} & \alpha_{11} \\ 0 & \alpha_{12} & \alpha_{13} \\ 0 & \alpha_{14} & \alpha_{15} \\ \alpha_{16} & \alpha_{17} & 0 \\ \alpha_{18} & 0 & \alpha_{19} \\ \alpha_{20} & \alpha_{21} & 0 \\ 0 & \alpha_{22} & 0 \\ \alpha_{23} & 0 & 0 \end{pmatrix} \begin{pmatrix} X^*_1 \\ X^*_2 \\ X^*_3 \end{pmatrix} + \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \\ \nu_4 \\ \nu_5 \\ \nu_6 \\ \nu_8 \\ \nu_9 \\ \nu_{10} \\ \nu_{11} \\ \nu_{12} \\ \nu_{13} \\ \nu_{14} \end{pmatrix}$$

Structural Equation:

The structural equation explains the latent constructs with socio-demographics and trip characteristics. The structural equation is defined as follows

$$X^*_n = \Lambda \tilde{X}_n + \omega_n \qquad \omega_n \sim N(0, I) \qquad [6\text{-}12]$$

where

Λ = matrix of unknown parameters (L x M)

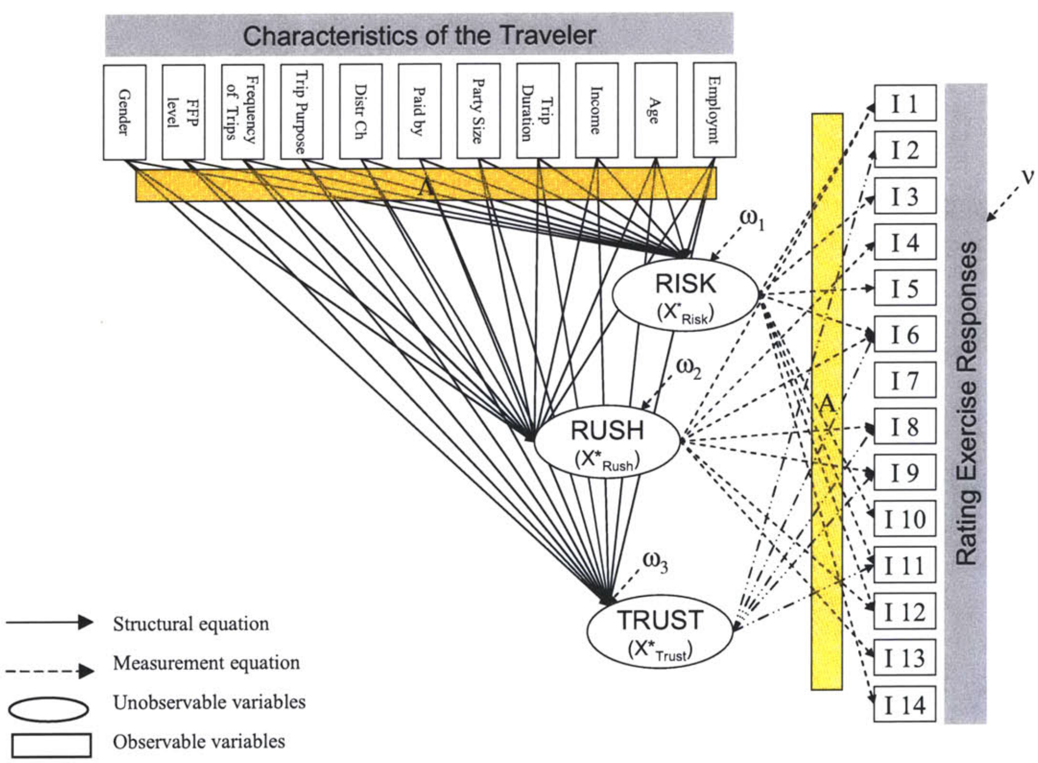
$\widetilde{X}_n$ = vector of explanatory variables causing the latent variables (M x 1)

ω_n = vector of disturbances of individual n (L x 1)

We assume normally distributed error terms and independence of the error terms, i.e. covariances are set to 0. For normalization purposes, the variance of each error term is set to one to set the scales of the latent variables. No constants are included in the structural equation since $\widetilde{X}_n$ are in deviation form.

The specification of the complete model with measurement and structural equation can be seen in Figure 6-13.

Figure 6-13: Specification of MIMC Model



Assuming that the error components are independent, the likelihood function is written as

$$f(I_n \mid X_n;\alpha,\lambda,\Sigma_\upsilon,\Sigma_\varepsilon) = \int_{X^*} f_M(I \mid X^*;\alpha,\Sigma_\upsilon)\, f_S(X^* \mid \tilde{X}\;;\lambda,\Sigma_\omega)\, dX^* \qquad [6\text{-}13]$$

Substituting in 6-13, we obtain

$$f(I_n \mid X_n;\alpha,\lambda,\Sigma_\upsilon,\Sigma_\varepsilon) = \int_{X^*} \prod_{r=1}^{R} \frac{1}{\sigma_{\upsilon_r}} \phi\left[\frac{I_r - \kappa_r - X^*\alpha_r}{\sigma_{\upsilon_r}}\right] * \prod_{l=1}^{L} \phi\left[X^*_l - \tilde{X}\,\lambda_l\right] dX^* \qquad [6\text{-}14]$$

where ϕ is the standard normal density function.

6.2.2 MIMC Model Results

The following figures show the results of the MIMC model estimation. Estimation was performed with a software program provided by Bolduc (see Bolduc and Giroux, 2005) with maximum simulated likelihood.

In Figure 6-14, the results of the measurement equation of the 'Multiple Indicator Multiple Causes Model' are displayed. The measurement equation consists of 13 equations, one for each of the rating statements used in estimation. The latent variable risk tolerance has high positive factor loadings for statements 3, 5 and 10, e.g. "I'm willing to accept the risk of a missed connection if this gets me to my destination earlier most of the time". Risk tolerance has negative factor loadings for statements 1 and 12, e.g. "I try to avoid short connections because of the risk of either me or my luggage missing the connecting flight".

The latent variable "rush aversion" has high factor loadings on statements 1, 4 and 11, e.g. "I like to take my time when connecting between flights" and "I enjoy having extra time at airports". The high factor loading on statement 4, "I don't think time at airports is waisted because I can shop, eat or work at airports" is an indication that the opportunity costs of a longer connection are low for individuals with rush aversion, making it easier to accept a longer connecting time. There is also a positive factor loading on the statement "It's hard for me to find my way through airports", which may add to the rush aversion of the individual when connecting.

The third latent variable, "trust in airlines' scheduling reliability", has high absolute factor loadings on statements 7 and 10. Statement 7, "Airlines sometimes underestimate the time needed to connect between flights", loads negatively on trust, as would be expected. Statement 10, "Airlines only sell connections that they expect passengers could make", has a positive factor loading on trust and shows the trust into airlines' scheduling reliability.

The statements of the rating exercise are used as indicators for the latent variables in the measurement equation. The purpose of the structural equation is to relate socio-demographics and trip characteristics to these latent variables. The goal is to describe the latent variables with these individuals' characteristics, thereby being able to assign a certain level of risk tolerance, rush aversion or trust to an individual based on his or her characteristics. In the future, this would allow to assess someone's latent variable levels based on socio-demographics and trip characteristics without having to resort to having the individual perform a rating exercise.

The squared multiple correlation R^2 is calculated as

$$R^2 = 1 - \frac{\text{error variance}}{\text{variance of dependent variable of equation}} \qquad [6\text{-}15]$$

Figure 6-14: Measurement Equation of MIMC Model

$\hat{\kappa} + \hat{\mathbf{A}} =$ intercept + (Risk Tolerance, Rush Aversion, Trust) loadings

$\hat{\kappa}$	Risk Tolerance	Rush Aversion	Trust	R^2	Item	
3.52(163.6)	−0.328(−16.9)	0.467(21.5)	0	0.30	I like to take my time when connecting between flights	I 1
2.52(109.3)	−0.154(−6.6)	0.107(4.9)	−0.339(−13.1)	0.12	It's hard for me to find my way through airports	I 2
3.76(172.9)	0.547(26.6)	0	0	0.28	Given two itineraries that only differ in connecting time, I always choose the one with shorter connecting time	I 3
2.65(126.9)	0	0.967(38.5)	0	0.66	I don't think time at airports is wasted because I can shop, eat, or work at airports	I 4
2.71(127.0)	0.672(28.1)	0	0	0.36	I'm willing to accept the risk of a missed connection if this gets me to my destination earlier most of the time	I 5
1.90(64.5)	0.357(15.2)	0.120(6.2)	−0.0655(−2.7)	0.12	I usually arrive at the check-in counter just before the check-in deadline	I 6
3.77(194.6)	0	0.214(12.7)	−0.539(−21.4)	0.37	Airlines sometimes underestimate the time needed to connect between flights	I 8
3.19(140.1)	0	0.192(8.6)	0.136(5.6)	0.04	It is the passenger's responsibility to plan for a sufficient transfer time when booking a connecting itinerary	I 9
3.17(153.8)	0.763(31.8)	−0.115(−6.2)	0	0.50	I don't mind being rushed at a connecting airport if this means I'll arrive at my final destination earlier	I 10
2.98(163.5)	0.242(12.5)	0	0.477(19.2)	0.29	Airlines only sell connections that they expect passengers could make	I 11
3.51(172.0)	−0.425(−20.3)	0.282(14.6)	0	0.26	I try to avoid short connections because of the risk of either me or my luggage missing the connecting flight	I 12
2.76(132.0)	0	0.924(37.2)	0	0.63	I enjoy having extra time at airports	I 13
4.01(213.3)	−0.211(−13.0)	0	0	0.07	I make sure that the planned connecting time is adequate for me when booking a connecting itinerary	I 14

(t-stats in parentheses)

Figure 6-15: Structural Equation of MIMC Model

$\hat{\Lambda}' =$

	Risk Tolerance	Rush Aversion	Trust
Female	−0.245(−5.0)	0.0967(2.3)	−0.187(−3.2)
Missed	0.194(2.9)	−0.256(−4.2)	−0.859(−10.5)
Elite	0.231(3.3)	0.507(8.4)	0.229(2.9)
Two to five trips	−0.098(−2.1)	0.0116(0.25)	0.168(2.6)
More than five trips	0.271(4.0)	0.121(2.0)	0.0353(0.4)
Business	−0.0201(−0.2)	0.438(5.3)	0.460(4.9)
Vacation	−0.0154(−0.3)	0.0389(0.9)	0.0430(0.7)
Online	0.0150(0.3)	0.0845(1.9)	0.333(5.2)
Company	−0.260(−2.4)	−0.193(−2.2)	−0.450(−4.1)
Alone	−0.00950(−0.2)	0.0274(0.5)	0.106(1.5)
Couple	−0.063(−1.1)	−0.163(−3.0)	−0.0666(−1.0)
Max three nights	0.293(4.6)	0.0280(0.5)	−0.0845(−1.1)
Four to seven nights	0.208(3.7)	−0.145(−2.8)	−0.0182(−0.3)
Employed	−0.056(−1.0)	0.114(2.1)	−0.0294(−0.4)
Student	0.339(2.7)	0.0916(0.8)	0.243(1.6)
Age < 30	−0.0151(−1.1)	0.00650(0.5)	−0.0597(−3.8)
Age 30 *to* 50	−0.00497(−1.1)	0.00245(0.6)	0.00497(0.9)
Age 50+	−0.0361(−6.4)	−0.00831(−1.8)	−0.00540(−0.7)
Income $< 30K$	−0.00341(−0.5)	−0.00533(−0.9)	0.0450(4.5)
Income $30K$ *to* $75K$	−0.000760(−0.4)	0.0226(1.3)	0.000419(0.2)
Income $> 75K$	−0.000290(−0.2)	−0.012(−9.4)	−0.00123(−0.8)
Bags	−0.177(−3.2)	0.126(2.6)	−0.000550(0.0)
Sq. Mult. Corr.	0.12	0.08	0.16

(t-stats in parentheses)

In Figure 6-15, the results of the structural equation of the MIMC model are displayed. The structural part of the model consists of three equations, one for each latent variable. All variables are dummy variables, except for age and income, which are specified piecewise linear. The variable "missed" stands for having missed a connection in the past 12 months. "Elite" are passengers that have Elite status on any airline. Number of trips ("two to five", "more than five") is based on air travel in the past twelve months. Trip purpose "business" and "vacation" are based on survey participants' responses about their last trip. "Online" stands for customers having booked their ticket via the internet. "Company" is a trip paid for by the respondent's company. "Alone" and "Couple" relate to the party size, and "max three nights" and "four to seven nights" to the length of stay. Income is defined in USD per annum. "Bags" stands for customers having checked bags for the flight. In the next paragraph, socio-demographics that are significant at the 95% level will be discussed.

Being female, having the ticket paid for by one's company, an age higher than 50, and having checked luggage reduces the risk tolerance value. Having the ticket paid for by one's company may be an indication for the importance of arriving on time at one's destination, thereby reducing the tolerance to misconnection risk. Having checked one's luggage may decrease the tolerance for risk because in that case, not only oneself might misconnect, but also the baggage. Even when arriving on time, risking to not having one's luggage may reduce the risk tolerance.

On the other hand, the level of risk tolerance is increased when being an elite member of any frequent flyer program, when taking more than five trips per year, on trips with up to seven nights duration, and when being a student. The first two characteristics describe an individual who is traveling often, making buffers in aggregate more expensive to the traveler. If the traveler would add 30 to 60 minutes of buffer to a weekly trip, this would add up to 26 to 52 hours of buffer time per year. In this case, the individual may be willing to risk missing one or several connections per year. For someone only traveling once or twice per year, inexperience and the fact that the absolute cost of a buffer is much lower than in the case described above may lead to lower risk tolerance. Risk tolerance is also increased by having missed a connection in the past two years. An explanation could be that someone who is risk tolerant has experienced the consequences of being risk tolerant by missing a connection.

Rush aversion is increased by being female, by being an elite member of any frequent flyer program, by going on a business trip, and by having checked luggage. It may seem counter-intuitive that being an elite member increases risk tolerance and also increases rush aversion. This shows that these are two distinct concepts. Elite members may be willing to risk a misconnection, but dislike the rush associated with a short connection. When going on a business trip, the rush aversion may be associated with the fact that rush could inhibit the concentration for the following meeting or simply lead to unwanted physical stress. Rush aversion is decreased by a party size of two versus a larger party, by trip length of four to seven nights and by a household income larger than $75,000. Parties of three or more may have greater rush aversion because of coordination needs and because the slowest in the group defines the overall speed going through the terminal. Having missed a connection in the past two years decreases rush aversion as well. As in the case of risk and having missed a connection, there may be a cause and effect relationship, i.e. someone who has a low value of rush aversion would book short connections and thus would be prone to missing a connection.

Trust in airlines' scheduling reliability is increased by being an elite member of any frequent flyer program, by going on a business trip and by booking online. Elite members may have a better, fact based, perception of airlines scheduling reliability and thus may have higher trust in airlines than non-elite members. Even in the case of misconnections, these may not be as burdensome for elite members since they would get special treatment by the airline because of their status. Individuals who do not trust airlines or who are uncertain about airlines' scheduling reliability may be inclined to avoid booking online. The value for trust is reduced by being female, by having missed a connection in the past two years, by having the ticket paid for by one's company, and by low income. In the case of having missed a connection, it seems like a logical conclusion to have a reduced level of trust. Low income could be an indicator for a lower level of education, which may be related to the level of trust in airlines' scheduling reliability.

The structural equation of the MIMC model explains the latent variables risk, rush and trust with socio-demographics and trip characteristics. In the next step, the values for the latent variables are interacted with attributes of the alternative in order to include the latent constructs in the choice model and thereby to provide a more realistic representation of airline itinerary choice.

6.3 Integrated Choice and Latent Variable Model

6.3.1 Model Specification

In joining the MIMC model with the choice model, the integrated choice and latent variable model (ICLV) specification is created. As described in chapter 4, the ICLV model consists of two structural and two measurement equations. The first structural equation is the choice model utility function, including the latent variables.

Structural Equations:

$$U_n = X_n\beta_1 + X^*_n\beta_2 + \varepsilon_n \qquad \varepsilon \sim \text{i.i.d. extreme value} \qquad [6\text{-}16]$$

$$X^*_n = \Lambda \tilde{X}_n + \omega_n \qquad \omega_n \sim N(0, I) \qquad [6\text{-}17]$$

where:

U_n = vector of utilities of J alternatives i

X_n = vector of attributes of alternatives i for individual n

X^*_n = vector of latent variables (L x 1)

β_1 = vector of unknown parameters relating to X_n

β_2 = vector of unknown parameters relating to X^*_n

Λ = matrix of unknown parameters (L x M)

$\tilde{X}_n$ = vector of explanatory variables causing the latent variables (M x 1)

ω_n = vector of disturbances of individual n (L x 1)

Disturbances of the latent variable function are assumed to be normally distributed. The error terms ε and ω are assumed to be independent, and variances for ω are set to 1 for identification purposes.

The defined latent variables are characteristics of the individuals n and thus need to be interacted with attributes of the alternative. Specifically, three models are presented where buffer time is interacted with one of the latent variables. The purpose is to explain heterogeneity in sensitivity to buffer time, e.g. if risk tolerant individuals have lower utility values from buffer time than risk averse individuals.

Measurement Equations:

$$I_n = \kappa + AX^*_n + \nu_n,\ \nu_n \sim \mathrm{N}\,(0, \Sigma_\nu \text{ diagonal}) \qquad [6\text{-}18]$$

$$y_{in} = \begin{cases} 1, & \text{if } U_{in} = \max\limits_j \{ U_{jn} \} \\ 0, & \text{otherwise} \end{cases} \qquad [6\text{-}19]$$

with

I_n = vector of indicators of latent variables of individual n (R x 1)

κ = vector of intercepts (R x 1)

A = matrix of unknown parameters (R x L)

X^*_n = vector of latent explanatory variables of individual n (L x 1)

ν_n = vector of disturbances of individual n (R x 1) where $\nu \sim \mathrm{N}\,(0, \Sigma_\nu)$

Disturbances of the indicators are assumed to be independent, i.e. the covariances are set to 0. The distribution of the error terms is assumed to be normal.

Likelihood Function:

In order to build the likelihood function, we start with the choice model, following Ben-Akiva et al. (2002):

$$P(y_n \mid X_n;\beta) \qquad [6\text{-}20]$$

We now add the structural equation of the latent variable model:

$$P(y_n \mid X_n;\beta,\lambda,\Sigma_\omega,\Sigma_\varepsilon)=\int_{X^*} P(y_n \mid X_n,X^*_n;\beta)f_s(X^* \mid \tilde{X}_n;\lambda,\Sigma_\omega)dX^* \qquad [6\text{-}21]$$

In the next step, we add the measurement equation, introducing the indicators:

$$f(y_n,I_n \mid X_n;\alpha,\beta,\lambda,\Sigma_\varepsilon,\Sigma_\nu,\Sigma_\omega)=\int_{X^*} P(y_n \mid X_n,X^*;\beta)f_s(X^* \mid \tilde{X}\ ;\lambda,\Sigma_\omega)f_M(I_n \mid X^*;\alpha,\Sigma_\nu)dX^*$$

[6-22]

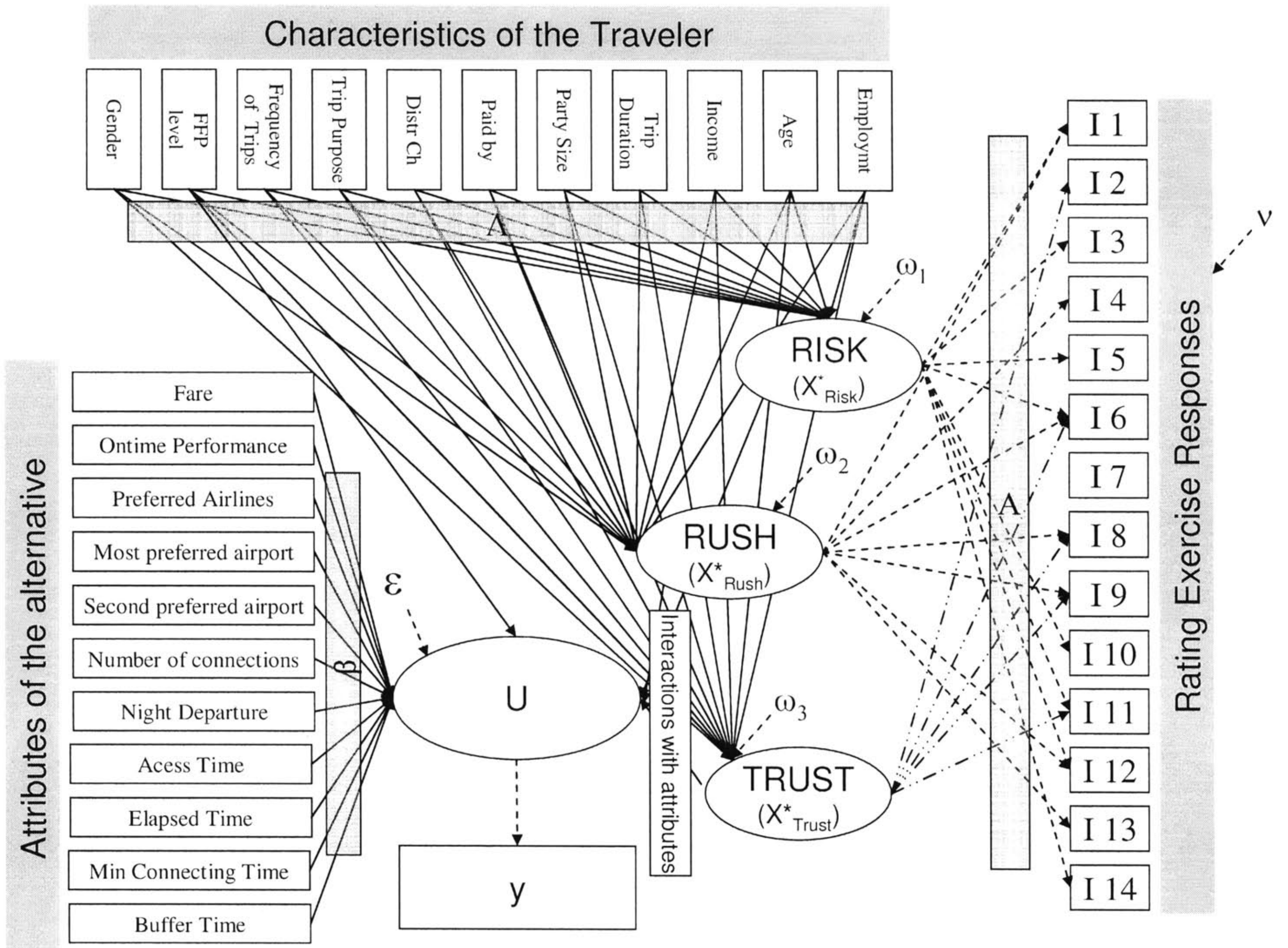
Finally, we insert the equations for the choice model including interactions with latent variables, the MIMC structural equation, and the MIMC measurement equation:

$$f(y_n,I_n \mid X_n;\alpha,\beta,\lambda,\Sigma_\varepsilon,\Sigma_\nu,\Sigma_\omega)=\int_{X^*}\left(\frac{e^{X_{in}\beta_1+X'_{in}X^*_n\beta_2}}{\sum_{j\in C_n} e^{X_{jn}\beta_1+\hat{X}_{jn}X^*_n\beta_2}}\right)*\prod_{l=1}^{L}\phi\left[X^*_l-\tilde{X}_n\lambda_l\right]*\prod_{r=1}^{R}\frac{1}{\sigma_{\upsilon_r}}\phi\left[\frac{I_m-X^*\alpha_r}{\sigma_{\upsilon_r}}\right]dX^*$$

[6-23]

The first term on the right hand side is the classical logit formulation. In addition to the attributes of the alternatives, the latent variables are interacted with specific attributes, in our case buffer time (X'X*). The second part of the right hand side corresponds to the structural equation of the latent variable model. It can be simplified to exclude the standard deviation term since the variance of the structural equation is set to 1 for identification purposes. The third term represents the measurement model of the latent variable model.

Figure 6-16: ICLV Model



6.3.2 ICLV Model Results

In this section, the results of the Integrated Choice and Latent Variable Model (ICLV) are described. Since the model interacting all latent variables with buffer time in the same model did not converge, separate models were run for each latent variable interacting with buffer time. Estimation was done by Maximum Simulated Likelihood with a software provided by Denis Bolduc (see Bolduc, D. and A. Giroux, 2005). In the following sections, the results of the three models are described. Section 6.3.2.1 displays results for the model interacting risk tolerance and buffer time. Section 6.3.2.2 describes results for the model interacting rush aversion and buffer time, and section 6.3.2.3 presents results for the model interacting the latent variable trust with buffer time. In section 6.3.2.4, a summary of the Integrated Choice and Latent Variable models is presented.

6.3.2.1 Risk Tolerance

In the first integrated model, the latent variable "risk tolerance" and buffer time are interacted and included in the choice model. In addition, number of connections is interacted with rush and trust. Table 6-3 and Figures 6-17 and 6-18 show the estimation results. As expected, a positive risk tolerance reduces the slope of the buffer time utility for the first 15 minutes, i.e. for a risk tolerant individual an increase of buffer time does not yield as much utility as for a risk averse individual.

Both interactions of number of connections with rush and trust produce insignificant results. The structural and measurement equation results of the latent variable model are very similar to those of the MIMC model, leading to the same interpretation as presented in the previous section.

Table 6-3: ICLV Model: Risk Tolerance and Buffer time

	Coefficient	Standard Error	T-test
Alternative specific constant current trip	1.103	0.078	14.1
ln (fare)	-4.008	0.148	-27.0
Elite frequent flyer program member	0.516	0.208	2.5
Standard frequent flyer program member	0.198	0.103	1.9
Ontime performance in percentage points	0.0115	0.003	4.5
Preferred airlines	0.147	0.07	2.1
First airport choice	0.791	0.10	7.9
Second airport choice	0.311	0.093	3.3
Number of connections	-0.367	0.132	-2.8
Night Departure	-0.424	0.157	-2.7
Access time in min	-0.00443	0.001	-6.7
Elapsed time in min	-0.00318	0.000	-11.7
Minimum connecting time in min	-0.00716	0.003	-2.3
Buffer time < 15 min in min	0.0126	0.005	2.4
Buffer time 15-59 min in min	-0.00399	0.002	-1.9
Buffer time > 60 min in min	-0.00145	0.002	-0.9
Interactions:			
Buffer time < 15 min in min and risk tolerance	-0.0134	0.005	-2.4
Buffer time 15-59 min in min and risk tolerance	0.00891	0.004	2.5
Buffer time > 60 min in min and risk tolerance	-0.00237	0.002	-1.5
Number of connections and rush aversion	0.0406	0.062	0.7
Number of connections and trust	0.0889	0.072	1.2

Number of Observations: 4664
Final log-likelihood: -1653.65

Figure 6-17: Measurement Equation of ICLV Model with Risk Tolerance

$\hat{\kappa} + \hat{\mathbf{A}} =$

$\hat{\kappa}$	Risk Tolerance	Rush Aversion	Trust	R^2		
3.52(162.8)	−0.327(−16.7)	0.466(21.4)	0	0.30	I like to take my time when connecting between flights	I1
2.52(108.9)	−0.153((−6.6)	0.106(4.8)	−0.340(−13.0)	0.12	It's hard for me to find my way through airports	I2
3.76(172.4)	0.548(26.6)	0	0	0.28	Given two itineraries that only differ in connecting time, I always choose the one with shorter connecting time	I3
2.65(126.9)	0	0.965(39.0)	0	0.66	I don't think time at airports is wasted because I can shop, eat, or work at airports	I4
2.71(126.4)	0.673(28.0)	0	0	0.36	I'm willing to accept the risk of a missed connection if this gets me to my destination earlier most of the time	I5
1.90(64.1)	0.357(15.0)	0.121(6.3)	−0.0665(−2.7)	0.12	I usually arrive at the check-in counter just before the check-in deadline	I6
3.77(194.1)	0	0.214(12.6)	−0.540(−21.2)	0.37	Airlines sometimes underestimate the time needed to connect between flights	I8
3.19(139.9)	0	0.192(8.6)	0.134(5.5)	0.04	It is the passenger's responsibility to plan for a sufficient transfer time when booking a connecting itinerary	I9
3.17(153.9)	0.764(32.0)	−0.112(−6.0)	0	0.50	I don't mind being rushed at a connecting airport if this means I'll arrive at my final destination earlier	I10
2.98(163.2)	0.241(12.3)	0	0.476(−19.0)	0.29	Airlines only sell connections that they expect passengers could make	I11
3.51(171.6)	−0.425(−20.2)	0.280(14.5)	0	0.26	I try to avoid short connections because of the risk of either me or my luggage missing the connecting flight	I12
2.76(131.8)	0	0.924(37.0)	0	0.63	I enjoy having extra time at airports	I13
4.01(212.6)	−0.212(−13.1)	0	0	0.07	I make sure that the planned connecting time is adequate for me when booking a connecting itinerary	I14

(t-stats in parentheses)

Figure 6-18: Structural Equation of ICLV Model with Risk Tolerance

$$
\hat{\Lambda}' = \begin{bmatrix}
-0.241(-4.9) & 0.0985(2.4) & -0.186(-3.2) & \textit{Female} \\
0.202(3.0) & -0.258(-4.2) & -0.855(-10.4) & \textit{Missed} \\
0.233(3.3) & 0.506(8.3) & 0.238(3.0) & \textit{Elite} \\
-0.0920(-1.9) & 0.0107(0.2) & 0.167(2.5) & \textit{Two to five trips} \\
0.273(4.1) & 0.129(2.1) & 0.0348(0.4) & \textit{More than five trips} \\
-0.0248(-0.2) & 0.442(5.3) & 0.458(4.8) & \textit{Business} \\
-0.0113(-0.2) & 0.0376(0.8) & 0.0374(0.6) & \textit{Vacation} \\
0.0145(0.3) & 0.0813(1.9) & 0.333(5.2) & \textit{Online} \\
-0.245(-2.2) & -0.194(-2.2) & -0.451(-4.0) & \textit{Company} \\
-0.0232(-0.4) & 0.0374(0.7) & 0.102(1.5) & \textit{Alone} \\
-0.0616(-1.0) & -0.165(-3.1) & -0.0691(-1.0) & \textit{Couple} \\
0.288(4.6) & 0.0273(0.5) & -0.0843(-1.1) & \textit{Max three nights} \quad K \\
0.200(3.6) & -0.142(-2.7) & -0.0164(-0.3) & \textit{Four to seven nights} \\
-0.056(-1.0) & 0.107(2.0) & -0.0260(-0.4) & \textit{Employed} \\
0.343(2.7) & 0.0851(0.7) & 0.255(1.7) & \textit{Student} \\
-0.0156(-1.2) & 0.00528(0.4) & -0.0590(-3.7) & \textit{Age} < 30 \\
-0.00481(-1.1) & 0.00287(0.7) & 0.00473(0.9) & \textit{Age}30 \textit{ to } 50 \\
-0.0366(-6.5) & -0.00801(-1.7) & -0.00506(-0.7) & \textit{Age } 50+ \\
-0.00357(-0.5) & -0.00577(-1.0) & 0.0453(4.5) & \textit{Income} < 30K \\
-0.000865(-0.5) & 0.0273(1.6) & 0.000240(0.1) & \textit{Income}30 \textit{ to } 75K \\
-0.000279(-0.2) & -0.0121(-9.5) & -0.00109(-0.7) & \textit{Income} > 75K \\
-0.184(-3.3) & 0.134(2.8) & -0.00585(0.1) & \textit{Bags} \\
\textit{Risk Tolerance} & \textit{Rush Aversion} & \textit{Trust} &
\end{bmatrix}
$$

Sq. Mult. Corr.	0.13	0.07	0.14

(t-stats in parentheses)

In order to understand how many respondents are displaying monotonic behavior based on this model, first the distribution of the risk tolerance variable is plotted in Figure 6-19, calculated by using the respondents' characteristics and the structural equation of the latent variable model. The median of this latent variable is -0.0158, with a standard deviation of 0.378. We then calculate the minimum value for the latent variable that results in a monotonic buffer time curve. To achieve a result of utility = 0 for the first fifteen minutes of buffer time, the sum of the coefficient values for "elapsed time", "buffer time < 15 min" and the interaction value require to be zero. The sum of the first two coefficients equals 0.00942. For the entire sum to be zero, the interaction value is required to be – 0.00942. With the interaction coefficient value being -0.0134, this leads to a corresponding risk tolerance value of -.00942/-.0134 = 0.703. Higher values of the latent variable risk represent monotonic behavior, while lower values represent non-monotonic behavior regarding buffer time. Based on the distribution of the risk tolerance variable, 95.9% of respondents show non-monotonic behavior regarding buffer time and 4.1% of respondents show monotonic behavior. Figure 6-20 shows the buffer time utility curve for the monotonic threshold value of 0.703 and for the median of the distribution.

Figure 6-19: Cumulative Distribution of Risk Tolerance Variable

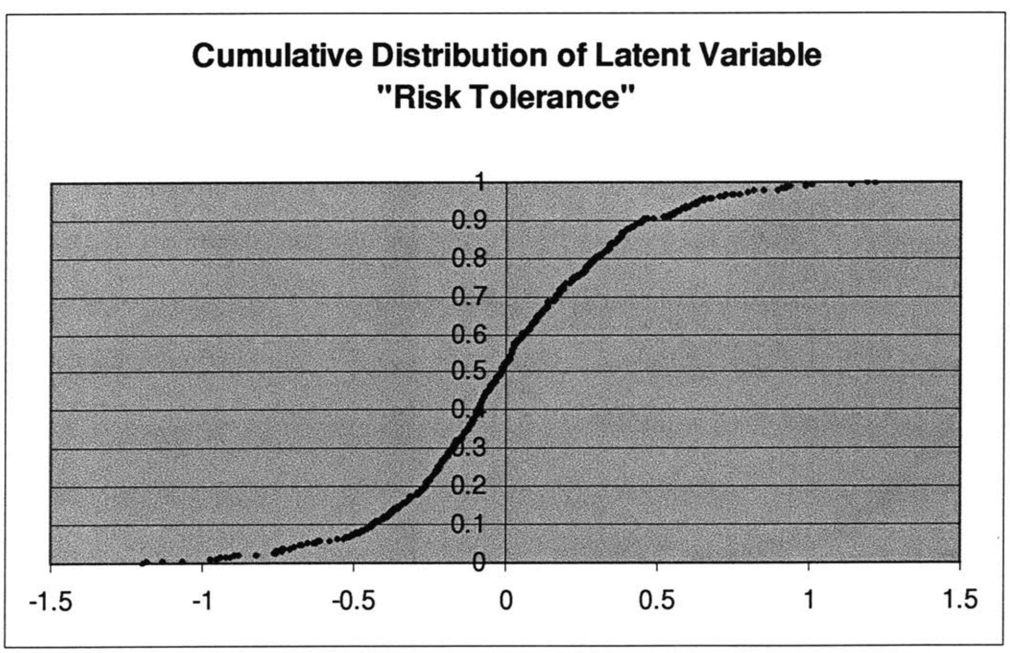
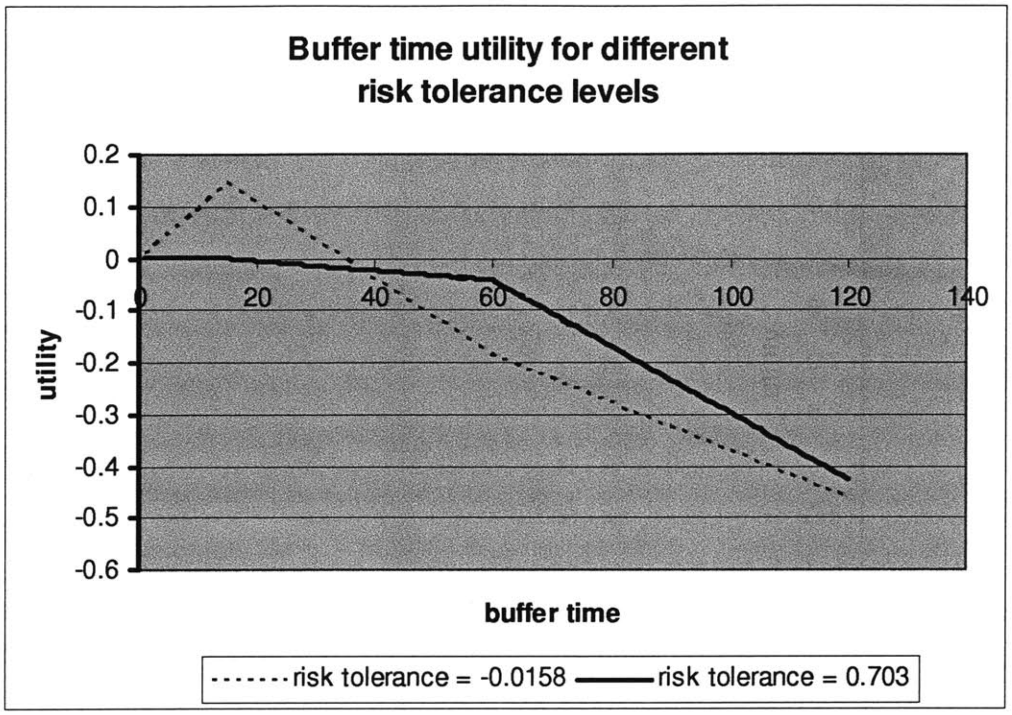


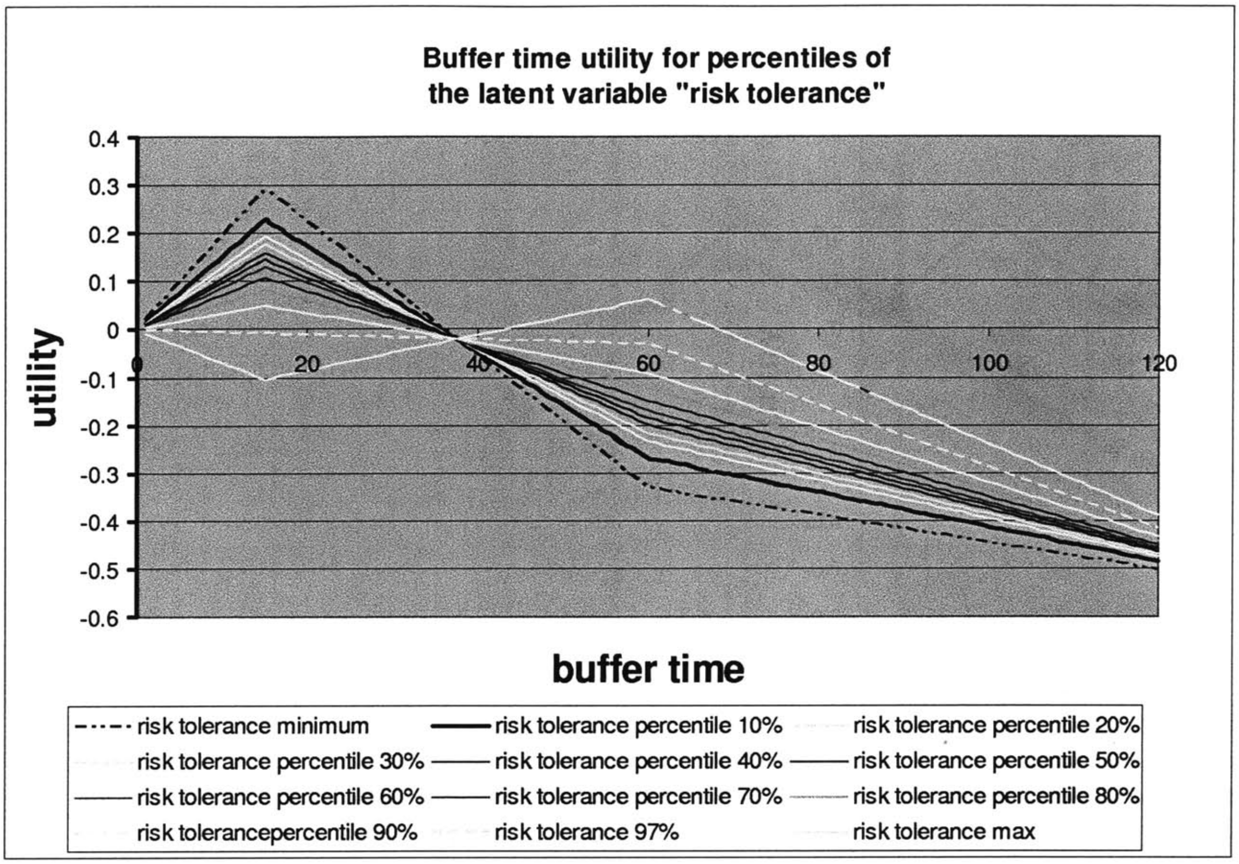
Figure 6-20: Buffer Time Utility for Different Risk Tolerance Levels



The graph for the median risk tolerance is similar in shape to the one based on the choice model alone. The risk tolerant curve shows monotonic behavior, meaning that starting at a buffer time of zero, any increase in buffer time reduces utility for that individual.

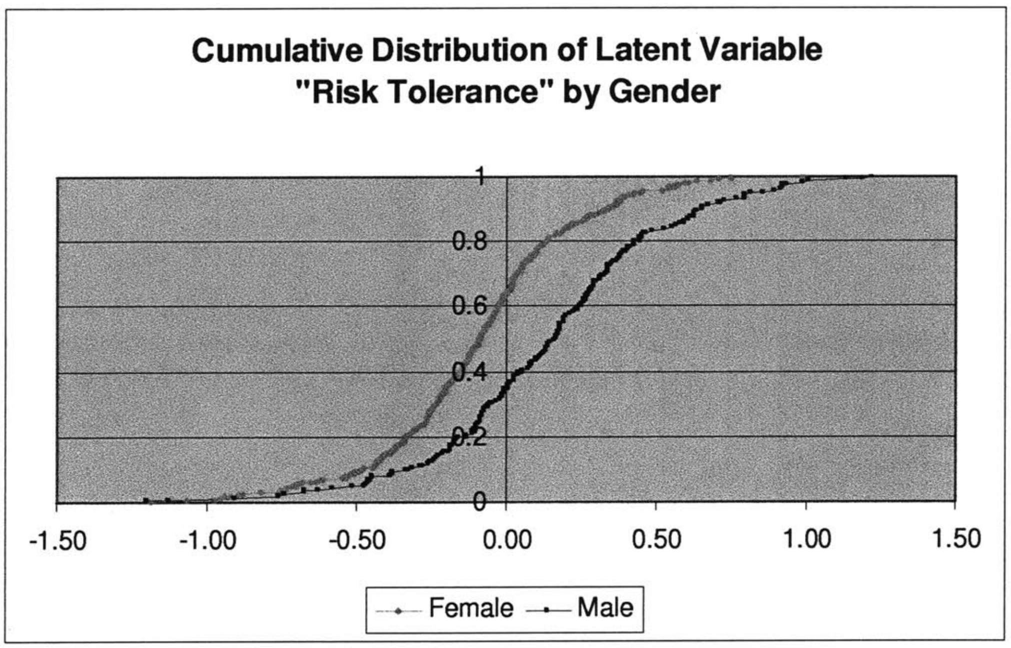
Figure 6-21 shows buffer time curves for different percentiles of the risk tolerance variable distribution. As expected, individuals with low risk tolerance have increasing utility for the first 15 minutes of buffer time, decreasing thereafter. Individuals with higher levels of risk tolerance have lower gains of utility by increased buffer times in the first 15 minutes. The maximum level of risk tolerance shows an unexpected curve, with increasing utility between minutes 15 and 60. For this outlier, the model does not show the expected results. Up to a percentile level of 97%, results are as expected. Further research with an extended dataset would have to be done to capture this outlier.

Figure 6-21: Buffer Time Utility for Percentiles of "Risk Tolerance" Distribution



Since these findings cannot be generalized because of the non-representative sampling of the survey, we also analyze subsets of the sampling population. In Figure 6-22, the cumulative distribution of the latent variable risk tolerance is shown for female and male respondents. On average, male respondents are more risk tolerant with a median of +0.15, and a median for the female subset of -0.09. 8.8 % of males have a risk tolerance value of .703 or higher, showing monotonic buffer time utility. For the female subset, 1.1 % have a risk tolerance of .703 or higher.

Figure 6-22: Cumulative Distribution of Risk Tolerance by Gender



Summarizing, the model interacting buffer time with the latent variable "risk tolerance" shows the expected results for latent variable values up to the 97% percentile. With a high risk tolerance, utility of buffer time is negative from minute zero. However, individuals with lower risk tolerance have an increasing utility for buffer time from 0 to 15 minutes, and decreasing utility thereafter.

6.3.2.2 Rush Aversion

In the second integrated model, the latent variable rush aversion is interacted with buffer time. In addition, risk tolerance and trust are interacted with number of connections. The results are displayed in Table 6-4 and Figures 6-23 and 6-24. The integrated choice model shows that being rush averse increases the utility of added buffer time in the first 15 minutes. This is what would be expected for a rush averse individual. Correspondingly, negative rush aversion reduces the utility of buffer time.

Table 6-4: Choice Model with Interactions (ICLV Model): Rush Aversion and Buffer Time

	Coefficient	Standard Error	T-test
Alternative specific constant current trip	1.093	0.078	13.9
ln (fare)	-4.028	0.148	-27.2
Elite frequent flyer program member	0.463	0.208	2.2
Standard frequent flyer program member	0.213	0.103	2.1
Ontime performance in percentage points	0.0115	0.003	4.5
Preferred airlines	0.157	0.069	2.3
First airport choice	0.805	0.10	8.0
Second airport choice	0.306	0.093	3.3
Number of connections	-0.418	0.132	-3.2
Night Departure	-0.389	0.157	-2.5
Access time in min	-0.00436	0.001	-6.5
Elapsed time in min	-0.00319	0.000	-11.7
Minimum connecting time in min	-0.00656	0.003	-2.1
Buffertime < 15 min in min	0.0113	0.005	2.4
Buffertime 15-59 min in min	-0.00397	0.002	-1.9
Buffertime > 60 min in min	-0.00141	0.002	-0.9
Interactions:			
Buffer time < 15 min in min and rush aversion	0.0193	0.006	3.5
Buffer time 15-59 min in min and rush aversion	-0.00671	0.003	-1.9
Buffer time > 60 min in min and rush aversion	0.00117	0.001	0.9
Number of connections and risk tolerance	0.107	0.065	1.7
Number of connections and trust	0.0720	0.072	1.0

Number of Observations: 4664
Final log-likelihood: -1653.51

Figure 6-23: Measurement Equation of ICLV model with Rush Aversion

$\hat{\kappa}+\hat{\mathbf{A}}=$	+ Risk Tolerance (t-stats in parentheses)	Rush Aversion	Trust	R^2		
3.52(162.6)	−0.326(−16.7)	0.467(21.4)	0	0.30	I like to take my time when connecting between flights	I1
2.52(108.9)	−0.153(−6.6)	0.107(4.9)	−0.340(−12.9)	0.12	It's hard for me to find my way through airports	I2
3.76(172.5)	0.547(26.5)	0	0	0.28	Given two itineraries that only differ in connecting time, I always choose the one with shorter connecting time	I3
2.65(127.0)	0	0.965(39.0)	0	0.66	I don't think time at airports is wasted because I can shop, eat, or work at airports	I4
2.71(126.3)	0.673(28.0)	0	0	0.36	I'm willing to accept the risk of a missed connection if this gets me to my destination earlier most of the time	I5
1.89(64.1)	0.358(15.1)	0.121(6.3)	−0.0664(−2.7)	0.12	I usually arrive at the check-in counter just before the check-in deadline	I6
3.77(194.1)	0	0.214(12.7)	−0.539(−21.2)	0.37	Airlines sometimes underestimate the time needed to connect between flights	I8
3.19(139.8)	0	0.192(8.5)	0.134(5.6)	0.04	It is the passenger's responsibility to plan for a sufficient transfer time when booking a connecting itinerary	I9
3.17(153.9)	0.762(31.9)	−0.114(−6.1)	0	0.50	I don't mind being rushed at a connecting airport if this means I'll arrive at my final destination earlier	I10
2.98(163.1)	0.241(12.3)	0	0.477(−19.0)	0.29	Airlines only sell connections that they expect passengers could make	I11
3.51(171.4)	−0.424(−20.2)	0.281(14.5)	0	0.26	I try to avoid short connections because of the risk of either me or my luggage missing the connecting flight	I12
2.76(132.0)	0	0.924(37.1)	0	0.63	I enjoy having extra time at airports	I13
4.01(212.5)	−0.212(−13.1)	0	0	0.07	I make sure that the planned connecting time is adequate for me when booking a connecting itinerary	I14

Figure 6-24: Structural Equation of ICLV Model with Rush Aversion

$$\hat{\Lambda}' = \begin{bmatrix}
-0.241(-4.9) & 0.0992(2.4) & -0.186(-3.2) & Female \\
0.201(3.0) & -0.259(-4.2) & -0.856(-10.4) & mMissed \\
0.235(3.3) & 0.506(8.3) & 0.236(3.0) & Elite \\
-0.0944(-2.0) & 0.00990(0.2) & 0.167(2.5) & Two\ to\ five\ trips \\
0.272(4.1) & 0.130(2.1) & 0.0350(0.4) & More\ than\ five\ trips \\
-0.0259(-0.2) & 0.440(5.3) & 0.458(4.8) & Business \\
-0.0118(-0.2) & 0.0369(0.8) & 0.0377(0.6) & Vacation \\
0.0126(0.3) & 0.0805(1.8) & 0.333(5.2) & Online \\
-0.245(-2.2) & -0.193(-2.2) & -0.451(-4.0) & Company \\
-0.0197(-0.3) & 0.0385(0.7) & 0.102(1.4) & Alone \\
-0.058(-1.0) & -0.163(-3.1) & -0.0692(-1.0) & Couple \\
0.291(4.6) & 0.0279(0.5) & -0.0850(-1.1) & Max\ three\ nights \\
0.202(3.6) & -0.142(-2.7) & -0.0166(-0.3) & Four\ to\ seven\ nights \\
-0.054(-0.9) & 0.108(2.0) & -0.0257(-0.3) & Employed \\
0.346(2.7) & 0.0853(0.7) & 0.255(1.7) & Student \\
-0.0150(-1.1) & 0.00569(0.4) & -0.0591(-3.7) & Age < 30 \\
-0.00500(-1.1) & 0.00279(0.7) & 0.00475(0.9) & Age 30\ to\ 50 \\
-0.0366(-6.5) & -0.00792(-1.7) & -0.00506(-0.7) & Age\ 50+ \\
-0.00339(-0.5) & -0.00559(-1.0) & 0.0453(4.5) & Income < 30K \\
-0.000978(-0.5) & 0.0264(1.5) & 0.000231(0.1) & Income\ 30\ to\ 75K \\
-0.000254(-0.2) & -0.0121(-9.4) & -0.00109(-0.7) & Income > 75K \\
-0.182(-3.3) & 0.135(2.8) & -0.00628(0.1) & Bags \\
Risk\ Tolerance & Rush\ Aversion & Trust &
\end{bmatrix}$$

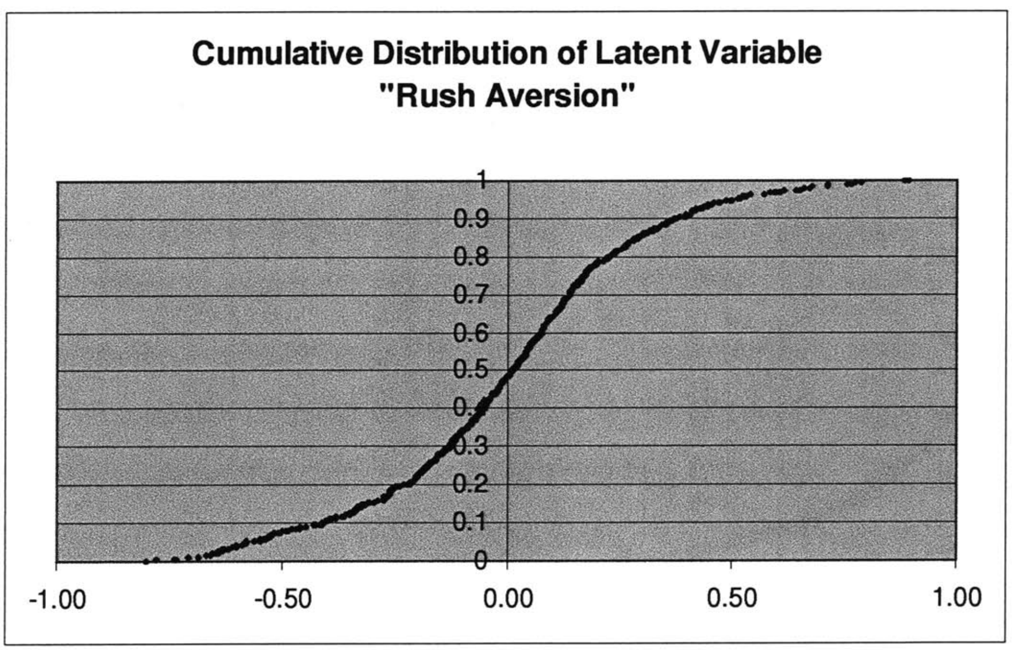
Sq. Mult. Corr.	0.13	0.07	0.14

(t-stats in parentheses)

The coefficient value for the interaction of "number of connections" and risk is also as expected. The coefficient value for "number of connections" alone is negative, meaning that one or more connections reduce utility. For a risk tolerant individual, the interaction value for risk tolerance and number of connections reduces the disutility of a connection. This intuitively makes sense since a risk tolerant person would see a connection as less of a burden. A similar effect can be seen in the interaction between "number of connections" and trust. This coefficient is positive as well, showing that a positive trust value reduces the disutility of a connection.

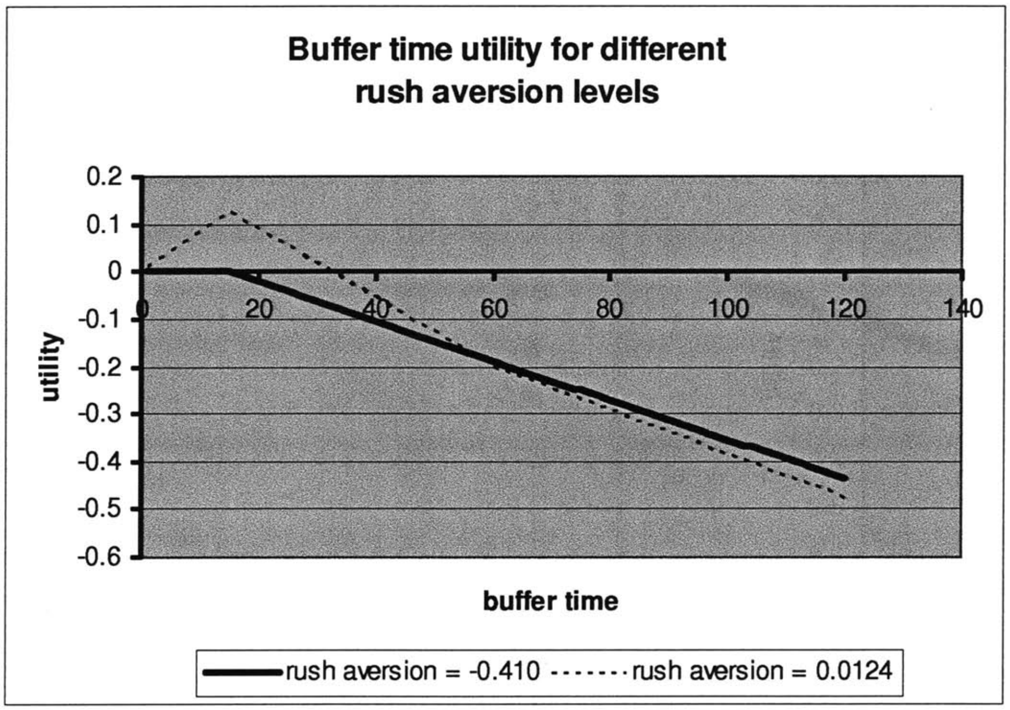
In Figure 6-25, the distribution of the rush aversion variable is presented. The median of the latent variable "rush aversion" is 0.0124, with a standard deviation of 0.303. Correspondingly to the calculation described in section 6.3.2.1, the maximum value allowed for a monotonic buffer time utility for rush aversion is -.0410. Of all respondents, 9.9% have rush aversion levels lower than -0.410, leading to monotonic buffer time curves.

Figure 6-25: Cumulative Distribution of Rush Aversion Variable



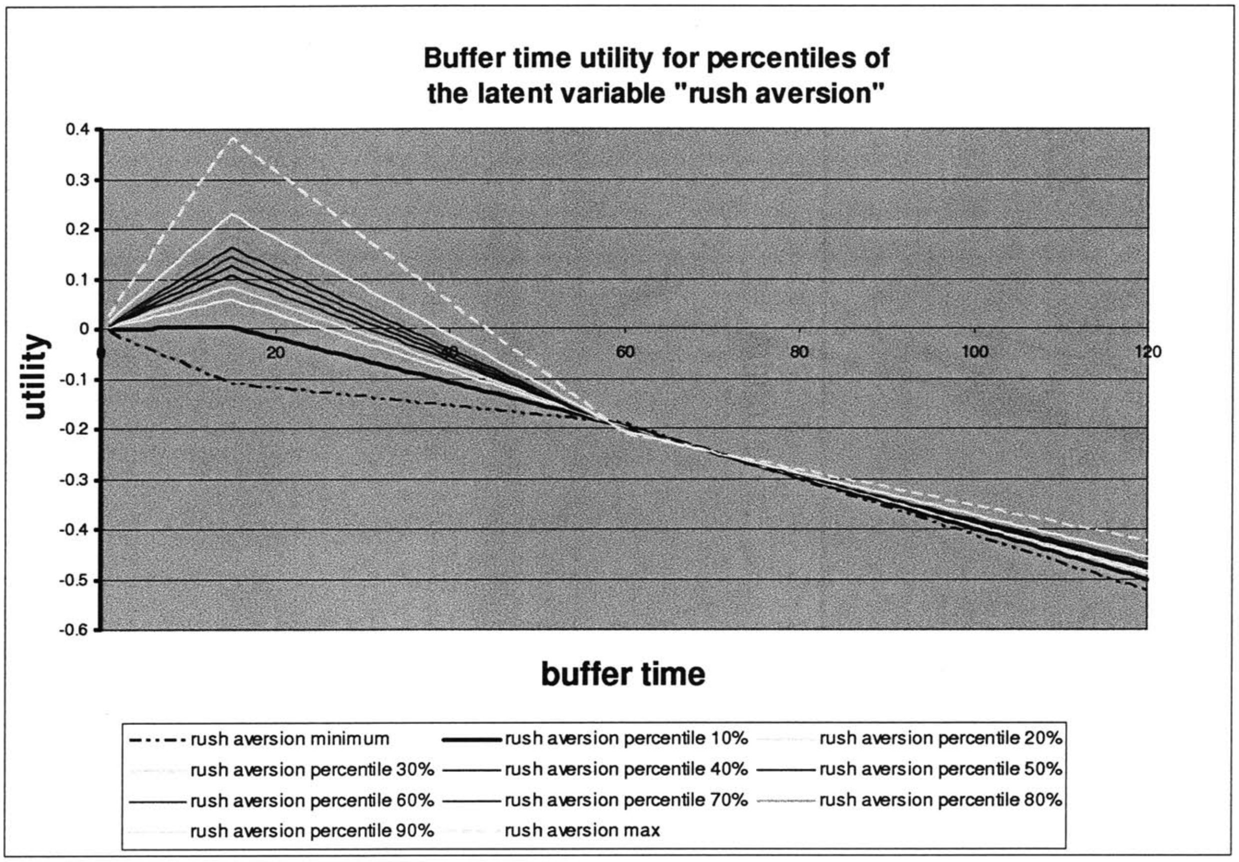
In Figure 6-26, two buffer time utility curves are presented, one with the median value for rush aversion, the other one with the rush aversion value of -0.410. Any value equal or lower than this value would lead to monotonic buffer time curves. Any value higher than -0.410 leads to non-monotonic buffer time curves, i.e. the utility increases for the first fifteen minutes. The rush averse individual's curve is similar to that of the choice model alone.

Figure 6-26: Buffer time Utility for Different Rush Aversion Levels



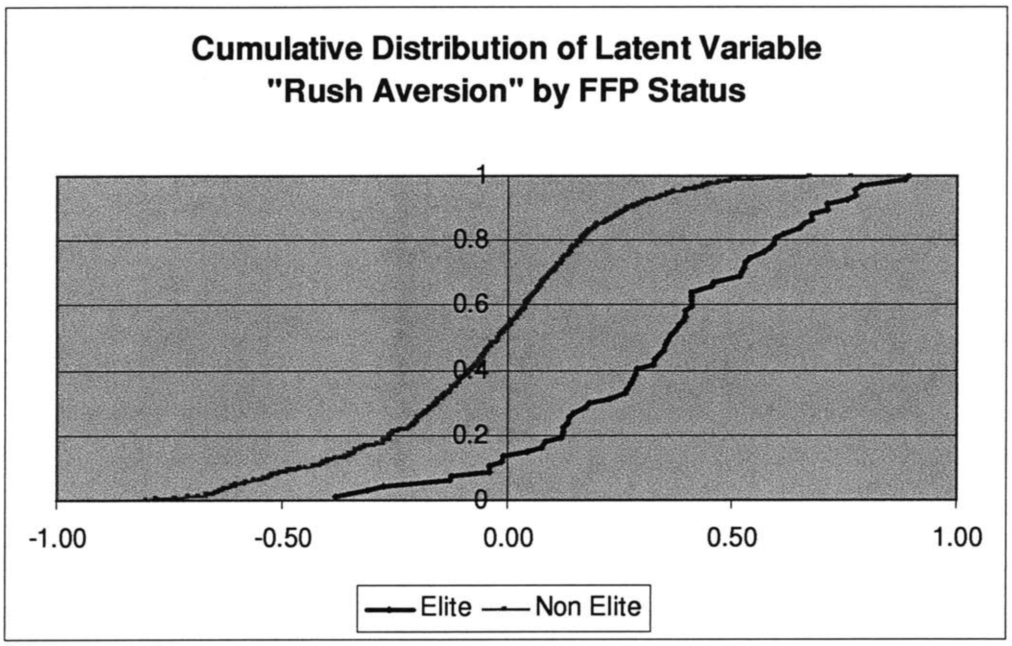
In Figure 6-27, different percentiles for the latent variable "rush aversion" are presented. It can be seen that the individuals with maximum rush aversion has the highest utility gain for buffer time in the first 15 minutes. Individuals with a lower level of rush aversion, for instance the 90th percentile of the latent variable distribution, accordingly have a lower utility gain for buffer time. The 9.9% of the respondents with the lowest rush aversion have a monotonic buffer time curve, as can be seen for the individual with the minimum rush aversion. At the 20% percentile, there is positive buffer time utility for 26 minutes, at the maximum rush aversion, for 44 minutes. Relating this to the case study in section 2.4 where individuals voluntarily added buffer time of 65 to 90 minutes to the minimum connecting time of 60 minutes (or added elapsed time of 50 to 60 minutes to the alternative itinerary), a hypothesis may be that individuals would have added a lower buffer time if that option had been available, but still opted to add the mentioned buffer time because it has a higher utility for them than not adding buffer time.

Figure 6-27: Buffer Time Utility for Percentiles of "Rush Aversion" Distribution



Since the non-probabilistic sampling of the survey does not allow to generalize the share of respondents with non-monotonic buffer time curves, we analyze two subsets of the sample, one being respondents who are an Elite member in any frequent flyer program, and the other subset those are not. Figure 6-28 shows the cumulative distribution of the latent variable "rush aversion" for these two groups. The median value for "Elites" is + 0.36, while the median value for Non-Elites is -0.02. Based on the previous discussion, individuals have a monotonic buffer time curve if their rush aversion value is -0.410 or lower. For Elites, non of the individuals have a smaller value, meaning that Elites all have a non-monotonic buffer curve in regards to rush aversion. For Non-Elites, 11.2% exhibit a monotonic buffer time curve.

Figure 6-28: Cumulative Distribution of Latent Variable "Rush Aversion" by Frequent Flyer Program (FFP) Status



In summary, the results show that heterogeneity can be better understood through interaction of buffer time with the latent construct of "rush aversion". Depending on the level of rush aversion of an individual, the buffer time utility may be increasing or decreasing for the first fifteen minutes of buffer time. Thus, introducing the interaction of buffer time and the latent variable "rush aversion" adds to the explanatory power of the model.

6.3.2.3 Trust in Airlines' Scheduling Reliability

The third integrated model interacts trust with buffer time. The results of the model are presented in Table 6-5 and Figures 6-29 and 6-30. Similar to the preceeding models, it was also tried to interact number of connections with the remaining two latent variables, however, this model did not converge. For the interaction of buffer time and "trust in airlines' scheduling reliability", the estimation results are as expected: the coefficient for the interaction of buffer time and trust is negative, meaning that a positive value for trust reduces the positive effect of buffer time in the first fifteen minutes. Thus, an individual with a high level of trust has fewer or no gains in utility through adding buffer time in the first fifteen minutes.

Table 6-5: Choice Model with Interactions (ICLV Model): Trust and Buffer Time

	Coefficient	Standard Error	T-test
Alternative specific constant current trip	1.103	0.078	14.1
In (fare)	-4.026	0.149	-27.1
Elite frequent flyer program member	0.510	0.210	2.4
Standard frequent flyer program member	0.189	0.103	1.8
Ontime performance in percentage points	0.0116	0.003	4.5
Preferred airlines	0.140	0.07	2.0
First airport choice	0.811	0.10	8.1
Second airport choice	0.312	0.093	3.3
Number of connections	-0.432	0.131	-3.3
Night Departure	-0.423	0.158	-2.7
Access time in min	-0.00438	0.001	-6.5
Elapsed time in min	-0.00313	0.000	-11.5
Minimum connecting time in min	-0.00621	0.003	-2.0
Buffertime < 15 min in min	0.0113	0.005	2.9
Buffertime 15-59 min in min	-0.00335	0.002	-1.5
Buffertime > 60 min in min	-0.00156	0.002	-1.0
Interactions:			
Buffer time < 15 min in min and trust	-0.0219	0.006	-3.1
Buffer time 15-59 min in min and trust	0.00817	0.003	2.0
Buffer time > 60 min in min and trust	0.00131	0.001	0.8

Number of Observations: 4664
Final log-likelihood: -1649.12

Figure 6-29: Measurement Equation of ICLV Model with Trust

$$\hat{\mathbf{A}} = \begin{pmatrix} 3.52(162.6) \\ 2.52(108.9) \\ 3.76(172.5) \\ 2.65(127.0) \\ 2.71(126.3) \\ 1.89(64.1) \\ 3.77(194.1) \\ 3.19(139.8) \\ 3.17(153.9) \\ 2.98(163.1) \\ 3.51(171.4) \\ 2.76(132.0) \\ 4.01(212.5) \end{pmatrix} + \begin{pmatrix} -0.327(-16.4) & 0.466(21.2) & 0 \\ -0.153(-6.6) & 0.106(4.8) & -0.343(-13.2) \\ 0.547(26.5) & 0 & 0 \\ 0 & 0.966(38.6) & 0 \\ 0.673(28.0) & 0 & 0 \\ 0.358(15.1) & 0.121(6.3) & -0.0683(-2.7) \\ 0 & 0.213(12.5) & -0.534(-21.4) \\ 0 & 0.193(8.8) & 0.133(5.5) \\ 0.762(31.9) & -0.113(-6.3) & 0 \\ 0.241(12.7) & 0 & 0.477(-19.0) \\ -0.425(-20.2) & 0.280(14.7) & 0 \\ 0 & 0.924(37.1) & 0 \\ -0.213(-13.3) & 0 & 0 \end{pmatrix}$$

Risk Tolerance | Rush Aversion | Trust

(t-stats in parentheses)

R^2	Statement	
0.30	I like to take my time when connecting between flights	I1
0.12	It's hard for me to find my way through airports	I2
0.28	Given two itineraries that only differ in connecting time, I always choose the one with shorter connecting time	I3
0.66	I don't think time at airports is wasted because I can shop, eat, or work at airports	I4
0.36	I'm willing to accept the risk of a missed connection if this gets me to my destination earlier most of the time	I5
0.12	I usually arrive at the check-in counter just before the check-in deadline	I6
0.37	Airlines sometimes underestimate the time needed to connect between flights	I8
0.04	It is the passenger's responsibility to plan for a sufficient transfer time when booking a connecting itinerary	I9
0.50	I don't mind being rushed at a connecting airport if this means I'll arrive at my final destination earlier	I10
0.29	Airlines only sell connections that they expect passengers could make	I11
0.26	I try to avoid short connections because of the risk of either me or my luggage missing the connecting flight	I12
0.63	I enjoy having extra time at airports	I13
0.07	I make sure that the planned connecting time is adequate for me when booking a connecting itinerary	I14

Figure 6-30: Structural Equation of ICLV Model with Trust

$$\hat{\Lambda}' =$$

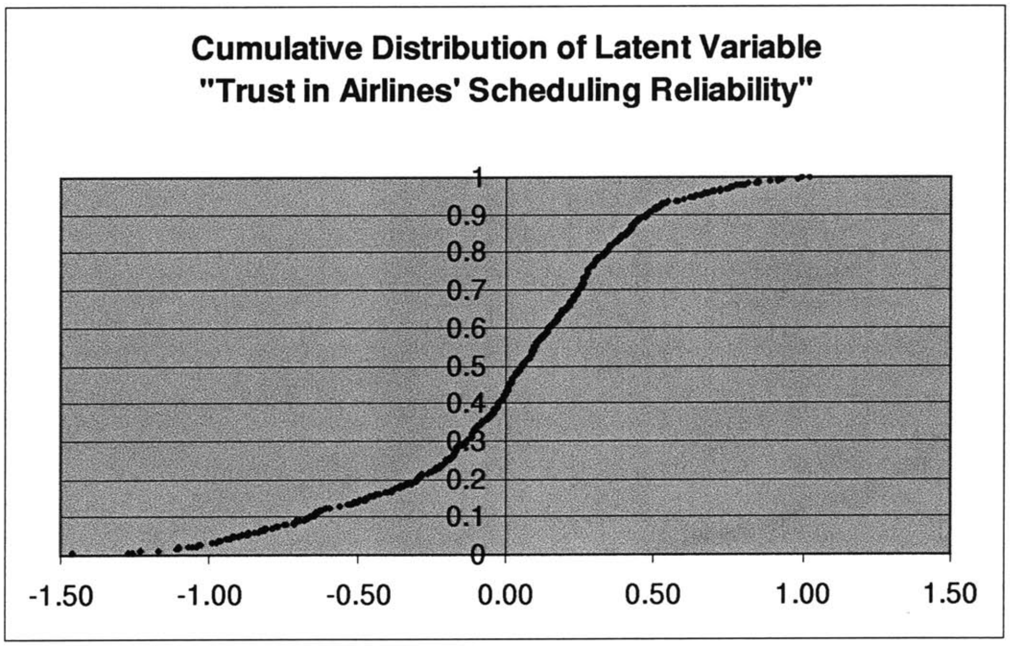
	Risk Tolerance	*Rush Aversion*	*Trust*
Female	−0.240(−4.9)	0.0992(2.4)	−0.189(−3.2)
Missed	0.201(3.0)	−0.258(−4.2)	−0.859(−10.3)
Elite	0.235(3.3)	0.505(8.3)	0.227(2.9)
Two to five trips	−0.0944(−2.0)	0.0112(0.2)	0.167(2.5)
More than five trips	0.273(4.1)	0.129(2.1)	0.0307(0.4)
Business	−0.0261(−0.3)	0.442(5.3)	0.468(4.9)
Vacation	−0.0106(−0.2)	0.0382(0.8)	0.0398(0.6)
Online	0.0135(0.3)	0.0808(1.8)	0.334(5.2)
Company	−0.245(−2.2)	−0.194(−2.3)	−0.458(−4.2)
Alone	−0.0203(−0.3)	0.0379(0.7)	0.0997(1.4)
Couple	−0.0583(−1.0)	−0.165(−3.1)	−0.0745(−1.1)
Max three nights	0.291(4.6)	0.0273(0.5)	−0.0863(−1.1)
Four to seven nights	0.201(3.7)	−0.143(−2.8)	−0.0215(−0.3)
Employed	−0.054(−0.9)	0.106(2.0)	−0.0314(−0.4)
Student	0.342(2.7)	0.0847(0.7)	0.264(1.8)
Age < 30	−0.0154(−1.2)	0.00528(0.4)	−0.0597(−3.7)
*Age*30 *to* 50	−0.00490(−1.2)	0.00283(0.7)	0.00480(1.0)
Age 50+	−0.0367(−6.1)	−0.00797(−1.6)	−0.00487(−0.6)
Income < 30*K*	−0.00349(−0.5)	−0.00585(−1.0)	0.0451(4.5)
*Income*30 *to* 75*K*	−0.000959(−0.5)	0.0278(1.4)	0.000579(0.3)
Income > 75*K*	−0.000264(−0.3)	−0.0121(−9.4)	−0.00123(−0.6)
Bags	−0.183(−3.3)	0.134(2.7)	−0.00890(0.1)
Sq. Mult. Corr.	0.13	0.07	0.14

(t-stats in parentheses)

In Figure 6-31, the distribution of the latent variable "trust in airlines' scheduling reliability" is presented. The median of the latent variable is 0.0551, with a standard deviation of 0.439.

For a monotonic buffer time utility, a minimum value of 0.372 is required for trust. Of the respondents, 17.3% have a trust value of at least 0.372. Lower values of trust lead to an increase of utility with increasing buffer time up to 15 minutes.

Figure 6-31: Trust Latent Variable Distribution



In Figure 6-32, two different buffer time utilities are presented. The curve based on the median trust value is similar to that of the choice model. The curve with a trust value of 0.372 remains at a utility level of zero for the first fifteen minutes. Any trust value higher than 0.372 would produce disutility from the first minute for every additional minute of buffer time.

Figure 6-32: Buffer time Utility for Different Trust Levels

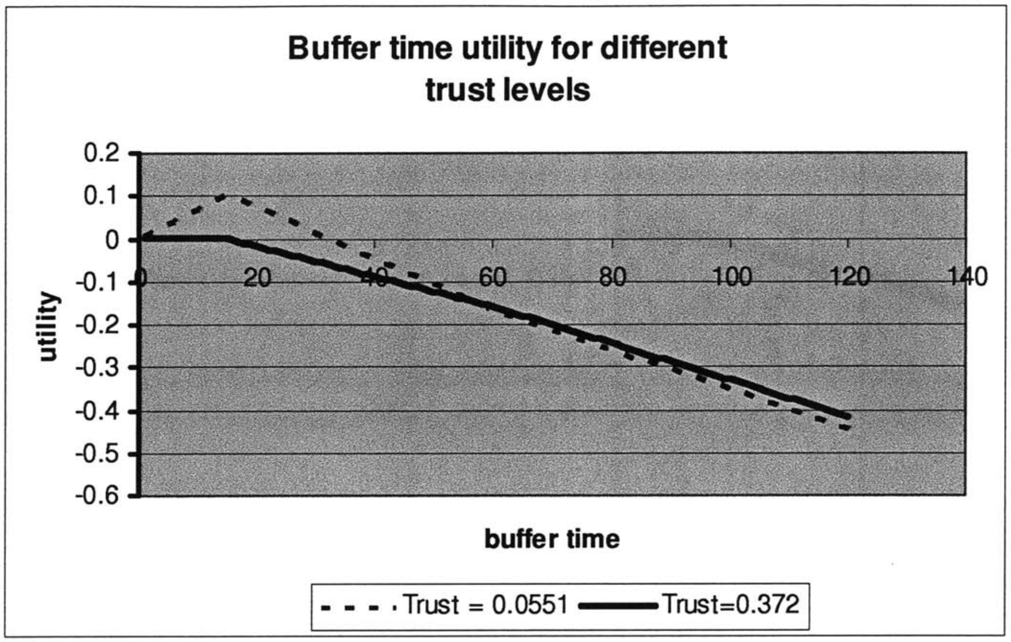
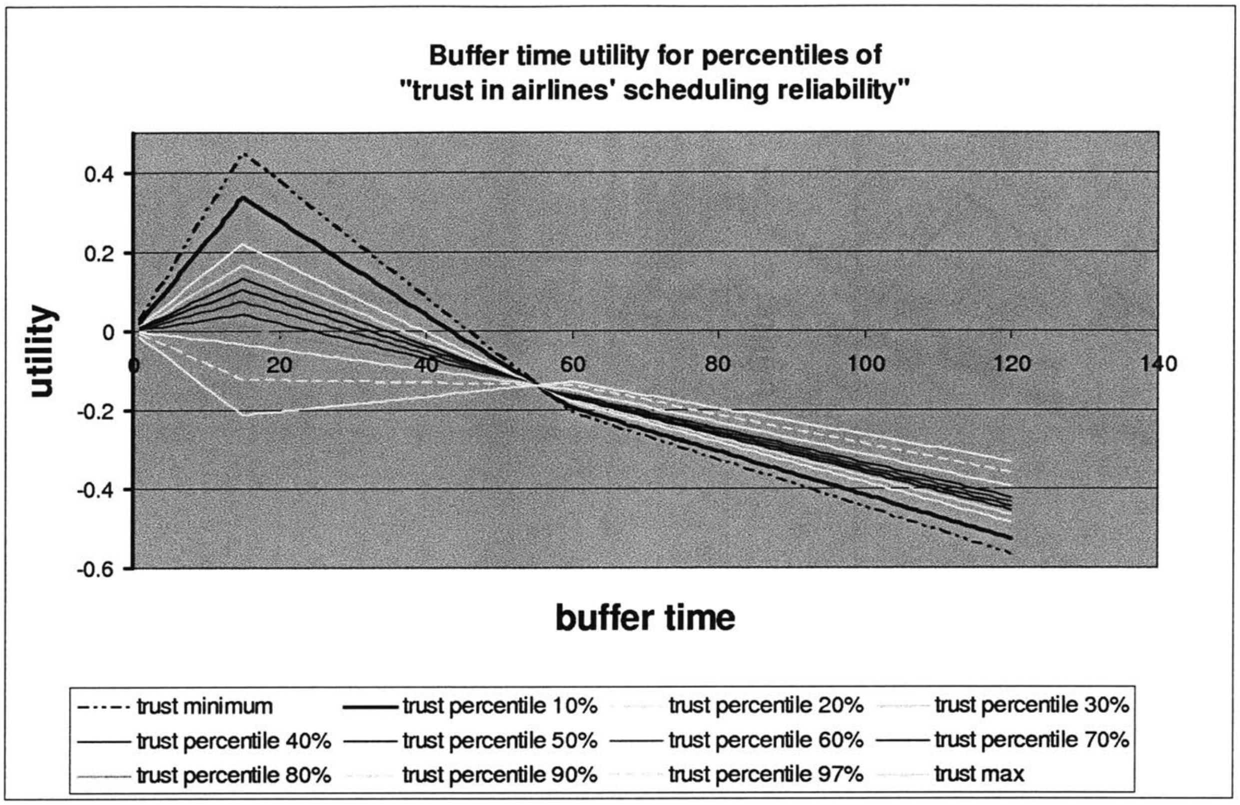


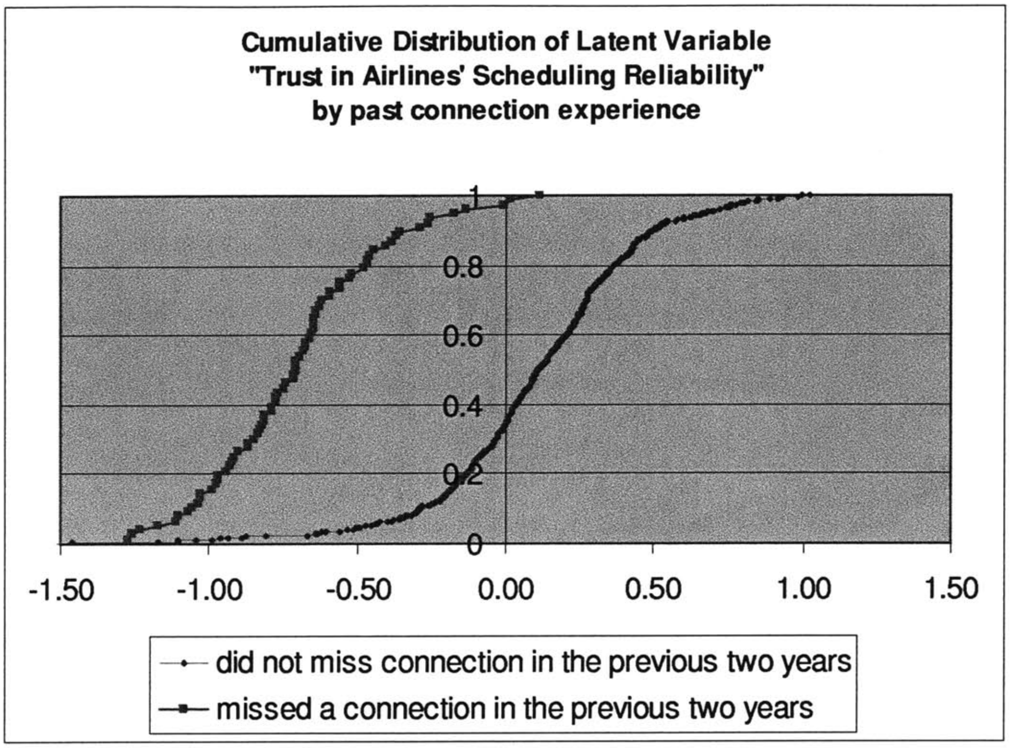
Figure 6-33 shows the graphs for different percentiles of the latent variable distribution. As expected, an individual with a minimum level of trust in airlines' scheduling reliability would have the highest increase in utility for the first fifteen minutes of buffer time, and decreasing utility thereafter. An individual with a 10^{th} percentile trust in airlines scheduling reliability has a slightly lower peak of utility than the individual with minimum trust. On the graph, the 90^{th} percentile trust level (high trust in airlines' scheduling reliability) has decreasing utility for buffer time from minute 0, as expected. The maximum level of trust shows decreasing utility from minute 0 as expected, however an unexpected increase for the segment of minutes 15 to 60. With an increased dataset, this could be improved.

Figure 6-33: Buffer Time Utility for Percentiles of "Trust" Distribution



In order to make a conditional statement on the non-monotonicity of buffer time, we segment the survey population in two parts: Those that have missed a connection in the previous two years and those that have not. Figure 6-34 shows the distributions for both segments. The distribution of respondents who have not missed a connection in the previous two years has a median of +0.11, while the median for those respondents who did miss a connection is -0.71. This intuitively is understandable since the missed connection will have had an effect on the individual's trust in an airline's scheduling reliability. Based on earlier discussion, any trust value of 0.372 or higher leads to a monotonic buffer time curve, i.e. a disutility of buffer time from the first minute. This is the case for 19.9% of respondents who have not missed a connection in the previous two years. For those that missed a connection, none have a trust value higher than 0.372, i.e. all of those that missed a connection display a non-monotonic buffer time curve.

Figure 6-34: Cumulative Distribution of Latent Variable "Trust" by Past Connection Experience



In summary, like the previous two models, the integrated choice and latent variable model with interaction of buffer time and "trust in airlines' scheduling reliability shows that the integration of latent constructs can add to the explanatory power and can explain heterogeneity. Depending on the level of trust an individual has in airlines' scheduling reliability, his utility of additional buffer time in an itinerary may increase or decrease.

6.3.2.4 Summary of ICLV Model Results

Three models were developed and estimated. The first one interacts buffer time with the latent variable "risk tolerance", the second model interacts buffer time with the latent variable "rush aversion", and the third one interacts buffer time with the latent variable "trust in airlines' scheduling reliability. For further research, based on a larger data sample, all latent variables should be interacted with buffer time to achieve a fully integrated model. The separate models in general show the expected results, i.e. a higher risk tolerance leads to lower utility of buffer time, a higher rush aversion leads to increased utility of buffer time, and a higher trust in airlines' scheduling reliability leads to a lower utility of buffer time. The models show that the latent constructs of "risk tolerance", "rush aversion", and "trust in an airlines' scheduling reliability can capture heterogeneity in the model. In a classic choice model, the decision process is a "black box", linking observed input to observed output. The integrated choice and latent variable model allows us to better understand the decision process itself and attitudes that influence the decision. The advantage of the integrated choice and latent variable model (for instance vs. capturing heterogeneity with random parameters) is that it gives us more information on the latent constructs itself (as a function of socio-demographics and attributes). The model can also be used in application without repeating the rating exercise, thereby giving airlines or others an easy tool to integrate attitudes based on socio-demographics and attributes into their applications.

6.4 Summary

The choice model results demonstrate the non-monotonicity of buffer time. The best fit model has cut off values for the piecewise linear specification of buffer time at 15 and 60 minutes. In the first fifteen minutes, utility increases when adding buffer time. In the second and third segment, buffer time utility decreases. The added average willingness to pay for 15 minutes of buffer time beyond the minimum connecting time is $15.

The Multiple Indicator Multiple Causes Model uses indicators from the rating exercise as manifestations of the latent variables risk, rush and trust. These are then explained with socio-demographics of the individuals. The squared multiple correlation values of the latent variable model are 0.12, 0.08 and 0.16 respectively for risk tolerance, rush aversion and trust. These values are rather low, either leading to the conclusion that more socio-demographic and trip information needs to be collected or that the latent variables in the model are not well explainable by any socio-demographics and trip characteristics. In the latter case, this would mean that individuals with very similar socio-demographics and trip characteristics could still have very different attitudes toward risk and rush and different levels of trust in airlines.

The latent variables are then interacted with attributes of the alternative to better explain behavior. It can be seen that risk tolerance and trust in airlines' scheduling reliability reduces the additional buffer time utility, up to the point that the buffertime utility is monotonistic. Rush aversion increases the additional utility of buffer time in the first 15 minutes, showing the added value of increased buffer time for rush averse individuals.

The results show that the hypothesis that, on average, slight increases of connecting time beyond the minimum connecting time, do not reduce revenue potential is valid.

7 Conclusions

7.1 Summary

Traditionally, it is assumed that airline passengers have a preference to minimize elapsed travel time and thus scheduled connecting times. In this thesis, we present a novel hypothesis about passenger utility in respect to connecting time. We hypothesize that passengers' utility might increase with increasing scheduled connecting time close to the minimum connecting times of airports, contrary to the traditional belief. We assume that attitudes toward rush and the risk of misconnection as well as the perceptions of an airline's scheduling reliability lead to an "n-shaped" utility function with respect to connecting times.

A case study based on airline booking data reveals that up to 25% of passengers in a given itinerary pair voluntarily choose the longer connection. To ensure the validity of results, booking data is compared to availability data on the booking class level. Only itinerary pairs where passengers would have had the same choice of fare class and price for the alternative connection were included. The results of the case study indicate that the phenomenon of avoiding short connecting times exists.

A survey was designed to incorporate effects toward connecting time as well as to collect information on attitudes, i.e. risk tolerance, rush aversion and trust into airline's scheduling reliability. The descriptive statistics demonstrate that 56% of survey respondents somewhat or strongly agree to the statement that they try to avoid short connections due to the risk of misconnection. 59% of respondents somewhat or strongly agree to the statement "I like to take my time when connecting at airports." Regarding trust into airline's schedule reliability, 67% of respondents somewhat or strongly agree that airlines sometime underestimate the time needed to connect. This demonstrates that attitudes toward risk, rush and trust can influence passenger itinerary choice.

Based on the stated preference data, a choice model was estimated that quantifies the disutility of short connecting times. On average, the willingness to pay for a connection with a buffer time of 15 min is $15 higher than for a connection scheduled with minimum connecting time. In order to capture attitudes toward risk, rush and the perception of airlines' scheduling reliability, rating questions were presented to the survey respondents. Responses were used as indicators for these latent variables. The Multiple Indicator Multiple Causes Model is a tool to explain the extracted attitudes with trip characteristics and socio-demographics. The results show that specific trip characteristics and socio-demographics, such as gender, elite status and having missed a connection in the past two years, influence attitudes toward risk, rush, and influence the level of trust into an airline's scheduling reliability. A relatively higher risk tolerance is observed for passengers who are Elite members of a frequent flyer program vs. non-members, for respondents who have missed a connection in the past two years vs. those who have not, for passengers with more than five trips/year, for short trips (maximum three nights) vs. longer trips, and for students vs. other employment types. A relatively lower risk tolerance is observed for women vs. men, for company paid trips vs. other payment sources, for passengers 50 years and older, and for passengers checking bags vs. those who do not check bags. A relatively higher level of rush aversion is observed for females, for Elite status passengers, for business travelers, and for those passengers with checked bags. A lower level of rush aversion is observed for those customers having missed a connection in the past two years, and those with a household income larger than $75K. Trust in airlines' scheduling reliability is relatively higher for Elite status passengers, for business travelers, and for customers who booked online. Trust into airlines' scheduling reliability is relatively lower for women, for those who have missed a connection in the past two years and for passengers traveling on a company paid ticket.

The Integrated Choice and Latent Variable Model (ICLV) combines the choice and latent variable model. By interacting the latent variables with attributes of the alternative, it can be shown that the latent variables influence choice behavior. Risk averse and rush averse individuals prefer longer connecting times than the minimum feasible connecting time.

The findings refute a long standing and non-contested assumption that it is advantageous to minimize connecting times for passengers in order to maximize revenue (see e.g. Mayer & Sinai, 2003).

The described framework, incorporating attitudes into a choice model, and the insights regarding the passenger transfer process can be applied to airline networks as well as other public transportation networks and itinerary choice environments.

7.2 Implications for Airlines

Traditionally, airlines tried to minimize connecting times in order to increase their revenue potential. In the discussion of depeaking of airline timetables, one of the concerns voiced was the increase in average connection times (Flint, 2002). Depeaking reduces the average ground times of aircraft, but increases the average connection time of passengers. Airlines were able to calculate the operational savings due to depeaking, but did not have the tools to calculate the demand side effects. Contrary to the traditional belief, however, the results of this thesis show that minimizing connecting times reduces the willingness to pay for an itinerary. Therefore, scheduling a connection at minimum connecting time reduces revenue potential as well as increases the operational costs at the airport.

Airlines have a number of options to improve their profitability based on these results. The key is to align the preferences of customers with the operational cost structure of the airline.

Fixed timetable

Without changing the timetable, a number of measures could be taken to increase profitability. Ideally, a segmentation of passengers based on their trip and socio-demographic characteristics would provide the base for offering risk and rush averse passengers longer connections in a given timetable. For instance, if an elderly woman with luggage to be checked would book a trip, the airline could proactively offer a longer connection to this passenger. Currently, travel agents sometimes perform this role, but with the increasing shift to web bookings, airlines need to provide this kind of advice themselves. By moving passengers who prefer longer connecting times or who are indifferent regarding connecting times to a longer connection, the airline reduces its expected value of irregularity costs, since a short connection has a higher probability of producing additional costs such as ramp direct transfer (direct bus shuttles from gate to gate), misconnection follow up costs (meals, hotels, etc.), and long-term customer retainment costs.

By actively offering longer connections and thus reducing demand for flights into banks, airlines could also potentially reduce the number of different sized aircraft in their fleet, i.e. increase their fleet commonality. For example, an airline may now fly an hourly service into its hub. Into the international banks, it may use a 270 seat aircraft to offer short connections, while in the off-peak times it may use a 150 seat aircraft. If the airline is able to shift those passengers who do not discount longer connecting times into off peak flights, it could use e.g. an 180 seat aircraft all day on that route for the hourly service. This would reduce their operational costs by reducing the fleet complexity and reducing training, maintenance and other costs for the aircraft that previously was specifically in service to provide flights into the banks.

In case no socio-demographic information is available on the person booking and traveling, airlines could still benefit from the described effect. First, airlines should rethink their display ranking on their web sites. Sorting by elapsed time (assuming that passengers have such a preference order) neither maximizes revenue potential nor reduces operational costs. If there are passengers who only look at the first line or who accept the suggested itinerary, the airline is selling an itinerary that creates a burden operationally, even if the passenger potentially would have been willing to accept a longer connection. Thus, airlines could avoid this effect by changing their default sortation criterion.

If a passenger books a very short connection, more airlines could give a warning like that displayed on ITA software's website (see Appendix A, Figure A-2): "The layover in Chicago has little room for delay, and for this route a missed connection would likely be very inconvenient". Simultaneously, the airline could offer an alternative connection with longer connecting time. This would allow the passenger to be aware of the short connection time, and based on his preferences and attitudes toward risk and rush, the passenger could choose between the two itineraries. In a second step, it could be tested if the willingness to pay for a longer connection could be turned into actual revenue of the airline, by adding e.g. $10 to the fare of the longer connection.

This would, of course, be traded off with the smaller number of passengers who would switch and thus the smaller operational cost savings. The airline would also have to evalute the probability that passengers would change to a different airline when seeing the warning about the short connecting time.

As an example, let us assume that the minimum connecting time is 45 minutes at the hub. The airline offers two alternative connections, one with a 45 minute connection and one with a 60 minute connection. Passengers that show non-monotonistic behavior regarding buffer time will be offered the 60 minute connection instead of the 45 minute connection.

Based on the integrated choice and latent variable model results, we can approximatively calculate the additional willingness to pay of passengers. Using the rush aversion interaction model, we assume that all passengers with rush aversion higher than the indifference level (-.410) will be offered the longer connection. Summing up the willingness to pay for elapsed time, buffer time and buffer time interaction with rush aversion yields a willingness to pay for the additional buffer time of $11.75 per passenger. Assuming an airline with annual passengers of 50M, 38% of these connecting, 10% of connecting passengers having a 0 min buffer time connection and as calculated above 90% of passengers showing non-monotonic behavior leads to an annual potential revenue increase of 20M$ as a result of increasing buffer time for these passengers. This calculation should serve as indicative for further research, using the stochastic specification of the model and the distribution of the input variables to simulate different scenarios. On the cost side, there are additional profit increasing effects of this measure as a result of reduced misconnections. Misconnections lead to disruption costs (accommodation, food and beverage, passenger goodwill). Misconnections also lead to a higher level of no shows in overbooking algorithms. Overbookings always incur a potential risk of oversales (i.e. passengers not being able to get a seat), which represents an additional cost to the airline.

Adapting the timetable

Depeaking the timetable leads to a slight increase in connecting times. Operationally, depeaking increases the utilization of resources (gates, ground equipment, baggage handling system, personnel) and thus reduces overall costs. Classical timetable design is based on minimizing the connecting times for the most important origin destination flows (based on revenue or number of passengers). In case of irregularities, this also leads to high misconnection costs. Relaxing this assumption would benefit risk and rush averse passengers and reduce operational costs.

When connecting times increase slightly, risk and rush averse passengers are better serviced than in the previous timetable, aimed at minimizing connecting times. In the traditional hub timetable, a passenger who wants to avoid a very short connecting time has to take a later flight, usually increasing the connecting time substantially. If however, connecting times are only increased by e.g. 15 minutes, this serves the risk and rush averse passenger while reducing potential misconnection costs of the airline. In addition, risk and rush averse passengers who might have moved to another airline can be retained as customers. (If, for instance, the connecting time with airline A is 30 minutes, connecting time with airline B is 45 minutes, the risk averse passenger may have chosen airline B because of the longer connecting time).

For example, airlines can choose to change their timetable and shift one of the flights such that the connecting time is increased from 45 to 60 minutes for all passengers. Such a strategy, based on the same assumptions as in the "fixed timetable example", would yield an average benefit of $10.25 per passenger or an annual benefit of $19M. On the cost side, in addition to the effects described above, there are the benefits from depeaking in terms of more efficient use of resources.

Increasing the connecting time for all passengers, however, also has negative effects. Those passengers who are risk tolerant and do not avoid rush are offered a service that is not as good as before. In this case, the airline has to trade off these effects and assess at the net effects. Based on our results, the net revenue effect of adding a 15 minute buffer time to the minimum connecting time is positive.

In either case of a fixed or an adapted timetable, the result of longer connecting times is a reduction of costs of misconnection, a reduction of the need for ad hoc services such as ramp direct service or rebooking services, and a reduction of outbound, passenger connection related, aircraft delays. The overall effect on the airline's bottom line, both on the revenue as well as cost side, would have to be evaluated by the airline before implementing these measures.

7.3 Contributions

The contributions of this thesis are as follows:

Based on airline booking data, it is demonstrated that the phenomenon of avoiding short connecting times exists. In the sample market, up to 25% of passengers voluntarily chose a longer connection, with all else equal. This shows that minimizing elapsed time does not always maximize the utility for individuals in a real market situation.

The non-monotonicity of connecting time utility is demonstrated based on survey data. This refutes the traditional assumption in the literature that short connections are better than longer connections and that individuals discount longer connections. Previous studies have based this on value of time aspects and have ignored other factors such as risk or rush aversion in the assessment of connecting time. This is the first research specifically studying connecting time utility, thus expanding the state of the art in airline itinerary choice modeling and enriching the explanatory power of such models.

Attitudes toward risk and rush and trust into airlines' scheduling reliability were quantified and linked to passenger trip charactistics and socio-demographics in order to explain taste variation in airline itinerary choice. It is demonstrated that e.g. gender, elite status and having missed a connection in the past significantly influence attitudes toward rush and risk and trust in airlines' scheduling reliability and thus result in different preferences regarding connecting time. Previously, there has been no research in airline itinerary choice that incorporates latent variables in model estimation.

As part of this dissertation, a framework to systematically evaluate the impact of connecting time on airline itinerary choice is developed and implemented. A stated preference experiment was designed and executed to incorporate these effects. In addition, a rating exercise was designed to capture latent attitudes toward rush, risk and trust into airlines' scheduling reliability. Explicitly incorporating these unobservable constructs into the choice model leads to a behaviorally more realistic representation of choice. This framework can be applied in future airline itinerary choice studies.

7.4 Future Research Directions

As with any research, this dissertation provides a wealth of opportunity for future research directions.

Applying the framework to additional attitudes: This is the first time in the literature that attitudes are integrated in an airline choice model. It would be advantageous to expand the framework to include additional attitudes and to link them to other attributes in the choice model. For instance, attitudes toward safety or security could be linked with airline choice or attitudes toward convenience with airport choice.

Quality of dataset: The benefit of the developed framework relies heavily on the quality of the data set. For future research, applying the framework to a larger dataset would allow for further insights regarding attitudes in passenger airline choice. For an airline, it would be useful to apply the framework to its own passengers, and to conduct a survey amongst its customers, giving the airline a better understanding of their passengers' needs and preferences. In addition, performing a survey at an airlines' hub would allow an airline to not only survey its own passengers, but also those of other airlines departing from their hub airport.

Integration in market share models of airlines: To our knowledge, the market share models currently in use at major airlines do not specifically treat the length of connecting time or include attitudes in their choice model. Attitudes are important to explain the unobserved individual heterogeneity amongst the survey population. Including these effects would provide for a richer and more realisitic representation of choice. Since market share models usually rely on revealed preference data, including attitudes would require an expansion of the framework to integrate revealed and stated preference data while including attitudes in the choice model. For airlines, this would be a major step in improving their market share models and their forecasting possibilities.

Using findings to enhance distribution channel displays: Given the results of this thesis, it would be advantageous to expand the capabilities of airline web sites by offering customers more choices. If socio-demographics of the customer are known, specific itineraries could be recommended, similar to recommendations on e.g. amazon.com. If the customer is unknown, the web site could be improved by providing an interactive dialogue and offering an alternative itinerary to the customer after he has made his choice. Similar to the offerings regarding price ("there is a cheaper flight if you're willing to change your schedule or move to an adjacent airport"), the website could offer a connection with longer connecting time, thereby giving customers a better choice. This could be combined with a sell up mode. Further research is needed to determine the best tradeoff between revenue from sell up and reduction in operational costs by extending the connecting time for this individual.

Incorporating attitudes in choice models provides for a more realistic representation of choice, and there is a wealth of opportunity in the airline choice world to apply these models, to improve customer satisfaction, and ultimately to improve profitability of airlines using these models.

Appendix A:

Treatment of Elapsed Time in Distribution Channels

Figure A-1: American Airlines Alert " Connecting time either short or long"

Returning												
Select	Carrier	Flight #	Departing City	Departing Date & Time	Arriving City	Arriving Date & Time	Aircraft Type	Cabin	AA Flight Miles	Meals	Travel Time
○	AMERICAN AIRLINES	1840	PDX Portland	11/27/2006 01:15 PM	ORD Chicago	11/27/2006 07:15 PM	M83	Economy View Seats	1730	$	8 hr 25 min
	AMERICAN AIRLINES	818	ORD Chicago	11/27/2006 09:30 PM	BOS Boston	11/28/2006 12:40 AM	M83	Economy View Seats	867	N/A	
○	AMERICAN AIRLINES	1398	PDX Portland	11/27/2006 12:00 PM	DFW Dallas/ Fort Worth	11/27/2006 05:40 PM	M83	Economy View Seats	1616	$	7 hr 55 min
	AMERICAN AIRLINES	1436	DFW Dallas/ Fort Worth	11/27/2006 06:20 PM	BOS Boston	11/27/2006 10:55 PM	M80	Economy View Seats	1562	$	
	Alert: Connecting time either short or long.										

Figure A-2: ITA Software Connection Warning

E-mail this itinerary | Back to search results | Modify search | Log out | Help | Comments | ITA Software (careers)

$660.20 in US Dollars
1 adult @ $660.20

Show booking details

This ticket is non-refundable.

Changes to this ticket will incur a penalty fee.

Airport legend:
BOS: Boston Logan
ORD: Chicago O'Hare
BOI: Boise Gowen Field
DEN: Denver Int'l

Boston, MA to Boise, ID: 2295 miles — 7 hrs 1 min

United Airlines Flight UA545 on a Boeing 757 (jet) in coach class
Departs **Boston, MA** (BOS) — Fri, Sept 29 — **6:00p** — 2 hrs 39 min
Arrives **Chicago, IL** (ORD) — **7:39p**

Layover in Chicago — 41 min

United Airlines Flight UA627 on an Airbus A320 (jet) in coach class (movie)
Departs **Chicago, IL** (ORD) — Fri, Sept 29 — **8:20p** — 3 hrs 41 min
Arrives **Boise, ID** (BOI) — **11:01p**

Note: The layover in Chicago (ORD) has relatively little room for delays, and for this route a missed connection would likely be very inconvenient.

Boise, ID to Boston, MA: 2395 miles — 6 hrs 29 min

United Airlines Flight UA112 on a Boeing 737 (jet) in coach class
Departs **Boise, ID** (BOI) — Fri, Oct 6 — **7:39a** — 1 hr 43 min
Arrives **Denver, CO** (DEN) — **9:22a**

Layover in Denver — 59 min

United Airlines Flight UA1192 on a Boeing 757 (jet) in coach class (food for purchase, movie)
Departs **Denver, CO** (DEN) — Fri, Oct 6 — **10:21a** — 3 hrs 47 min
Arrives **Boston, MA** (BOS) — **4:08p**

Table A-1: Elapsed Time and Connecting Time in Airline Distribution Channels

		Airlines' web sites							Online Travel Agencies				Metacrawlers			Opaque search engines	
		AA	CO	DL	NW	UA	US	WN	Expedia	Orbitz	Travelocity	ITA Software	Kayak	Mobissimo	Sidestep	Hotwire	Priceline
Default first screen	SD-Schedule driven search FD-fare driven search	SD	FD	No choice (FD)	FD	FD	SD	No choice (FD)	No choice (FD)	No choice (FD)	No choice (FD)	No choice (FD)	No choice (FD)	No choice (FD)	No choice (FD)	No choice (FD)	No choice (FD)
Selection of itineraries	D-per direction R-Return	SD:D FD:R	D	D	SD:D FD:R	SD:D FD:R	D	D	D	R	D	R	R	R	R	R	D
Display of search results	Displays separate flights with times when connecting	yes	yes	no	yes	yes	no	no	no	yes	no	no	no	no	no	no	no
	Displays connecting time	no	yes	yes	no	no	no	no	no	yes	no	no	no	no	no	no	no
	Displays elapsed time	yes	yes	yes	no	no	yes	yes	no	yes	no	yes	no	no	no	no	yes
	Sortation criteria	FD: 1.P 2. DD SD: 1.#C 2.DD	FD: ? SD: 1. #C 2.Dep	1. #C 2. Dep 3. ET	FD: 1. P SD: 1.Dep	FD: 1.P SD: 1.DD	SD 1.#C 2.Dep FD: 1.P 2.#C 3. Dep	1.Dep	1.price	1.price	1.price 2.DD	1.price	1.price	1.price	1.price	1.price	1.price
	Results sortable by elapsed time	no	yes	yes	yes	no-FD yes-SD	yes	no	yes	yes	yes	yes	yes	no	no	no	yes
	Other sortation options for user	FD: Dep, Arr	P, Dep, Arr, Miles	P, Dep, Arr	P, Dep	SD: Dep	P, Dep, Arr	none	P, Dep, Arr, DD	Dep	Airline, Dep, Arr	Airline, #C, Dep	Airline, Dep, Arr	Airline, Dep, Arr, #C, Provider	Airline, Dep, #C		Dep, Arr
	Warnings w/ respect to connecting time	see Note 1	none	none	none	none	none	none	non	none	none	see Note 2	none	none	none	none	none

Search date: March 18, 2006

Legend:
AA: American Airlines
CO: Continental Airlines
DL: Delta Airlines
NW: Northwest Airlines
UA: United Airlines
US: US Airlines
WN: Southwest Airlines

D – per direction
R – Return flight

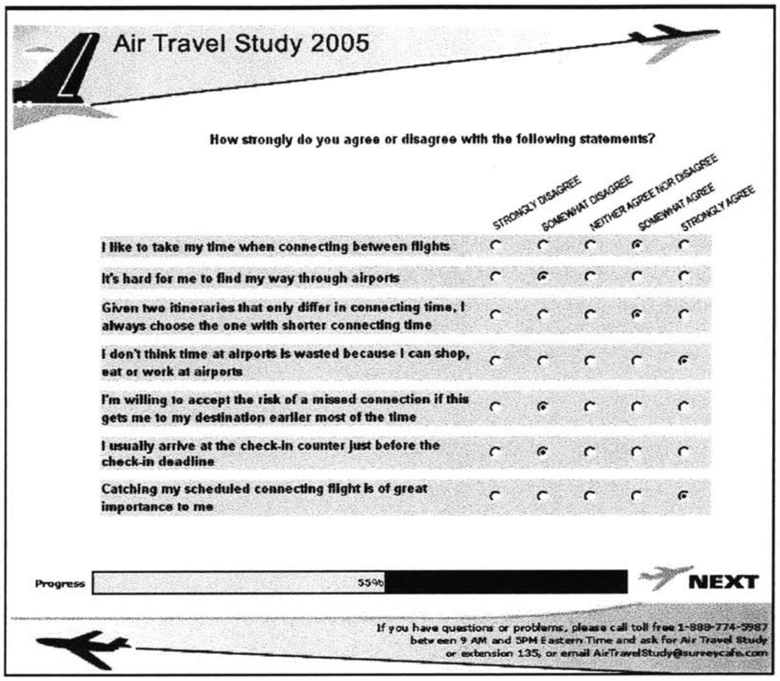
FD: Fare driven
SD: Schedule driven

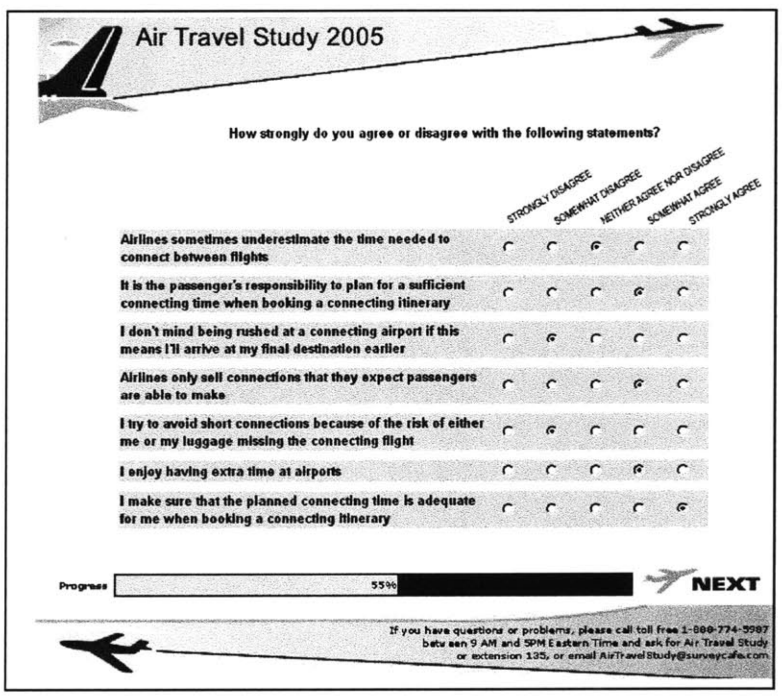
P: Price
Dep: Departure Time
Arr: Arrival Time
DD: Deviation from requested departure time
ET: elapsed time

Note 1: Connecting Time either long or short
Note 2: "tight connection/long layover"

Appendix B: Survey Design

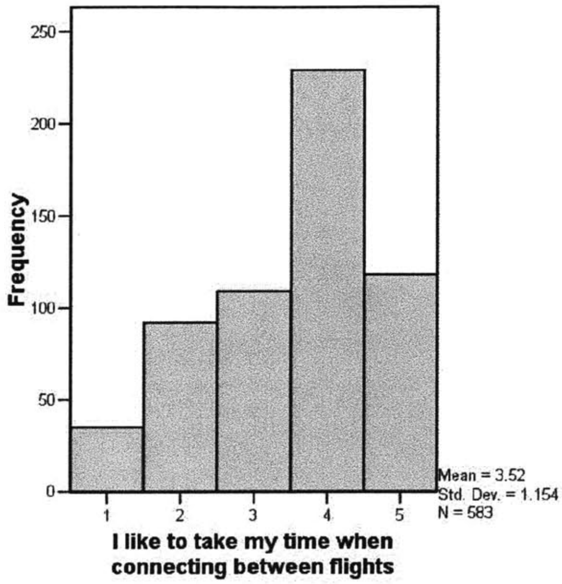
Figure B-1: Screen Shots of Rating Exercise





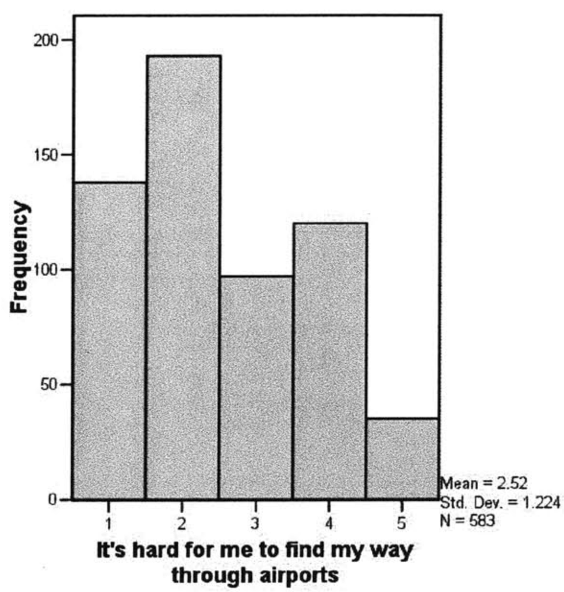
Appendix C: Summary Statistics

Figure C-1: Rating Question 1



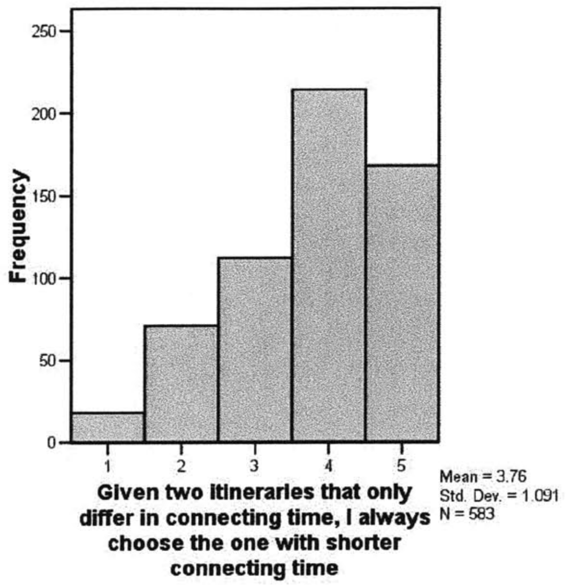
1-strongly disagree
2-somewhat disagree
3-neither agree or disagree
4-somewhat agree
5-strongly agree

Figure C-2: Rating Question 2



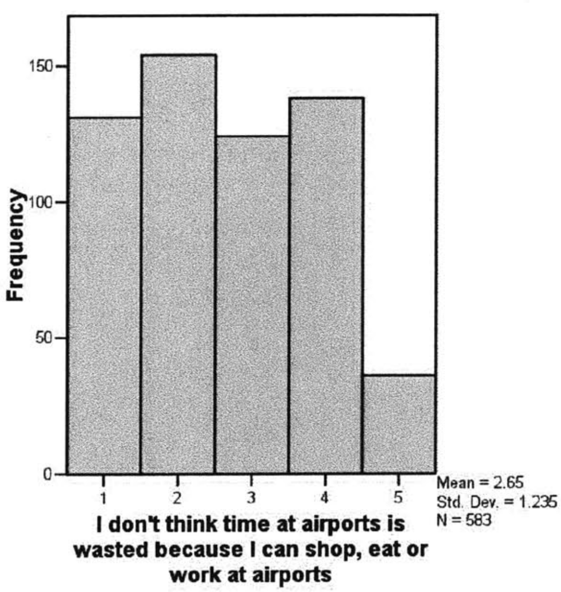
1-strongly disagree
2-somewhat disagree
3-neither agree or disagree
4-somewhat agree
5-strongly agree

Figure C-3: Rating Question 3



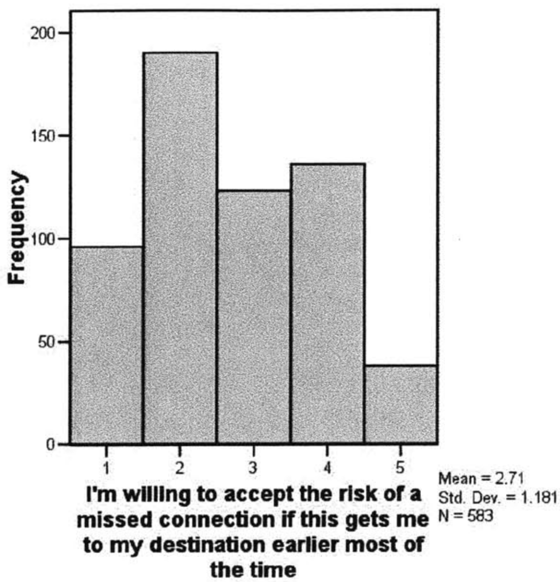
1-strongly disagree
2-somewhat disagree
3-neither agree or disagree
4-somewhat agree
5-strongly agree

Figure C-4: Rating Question 4



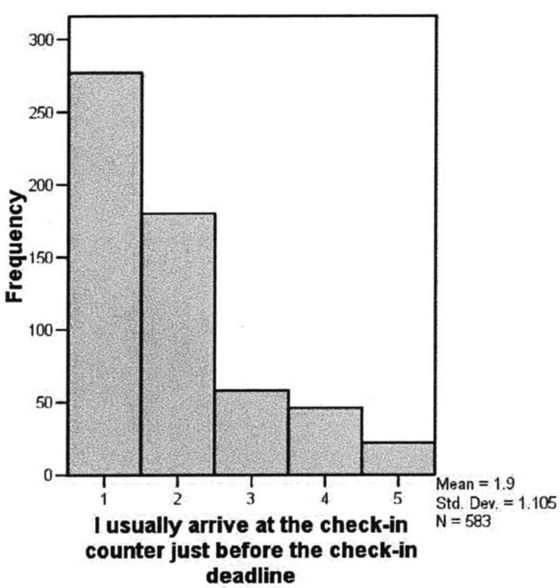
1-strongly disagree
2-somewhat disagree
3-neither agree or disagree
4-somewhat agree
5-strongly agree

Figure C-5: Rating Question 5



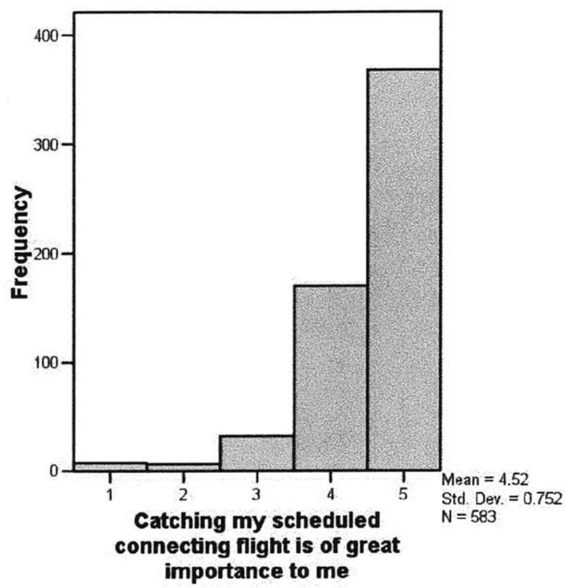
1-strongly disagree
2-somewhat disagree
3-neither agree or disagree
4-somewhat agree
5-strongly agree

Figure C-6: Rating Question 6



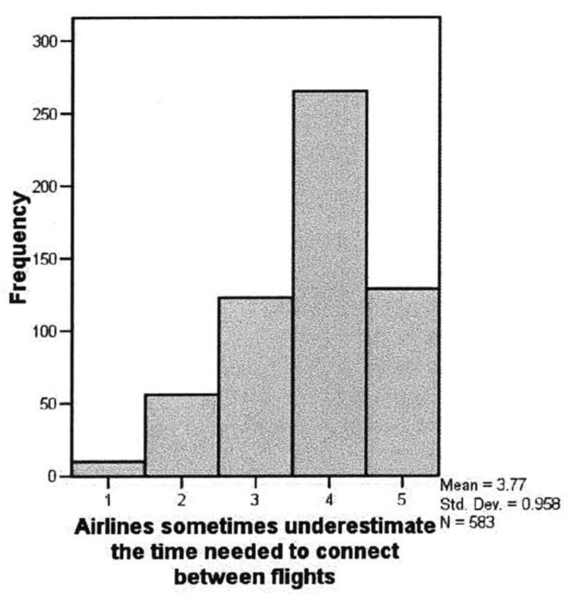
1-strongly disagree
2-somewhat disagree
3-neither agree or disagree
4-somewhat agree
5-strongly agree

Figure C-7: Rating Question 7



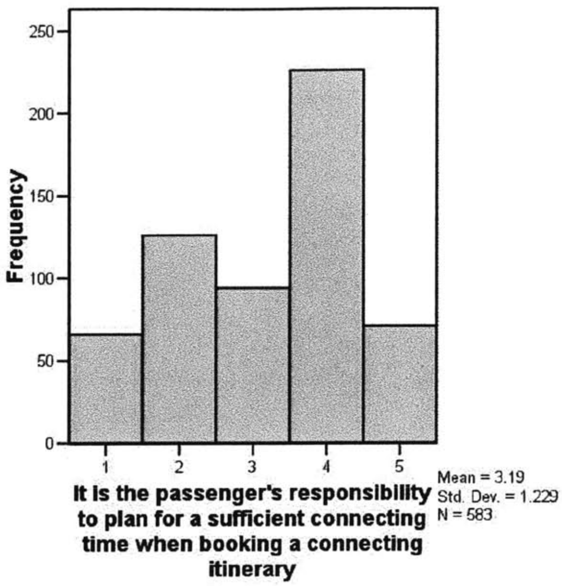
1-strongly disagree
2-somewhat disagree
3-neither agree or disagree
4-somewhat agree
5-strongly agree

Figure C-8: Rating Question 8



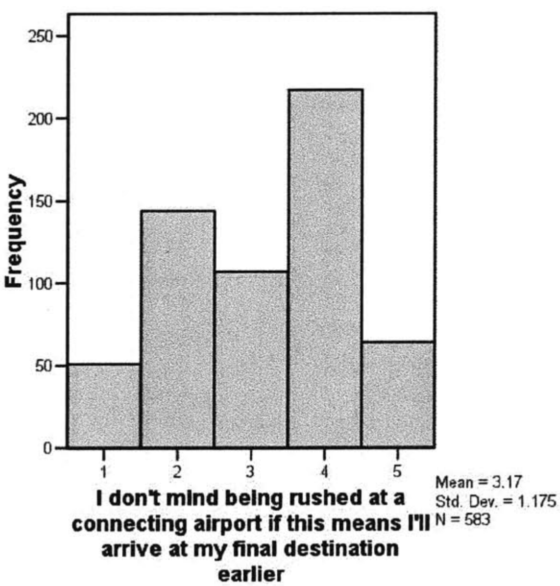
1-strongly disagree
2-somewhat disagree
3-neither agree or disagree
4-somewhat agree
5-strongly agree

Figure C-9: Rating Question 9



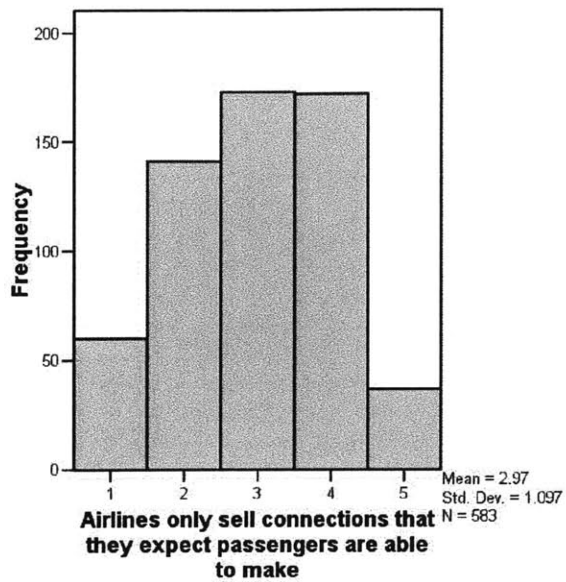
1-strongly disagree
2-somewhat disagree
3-neither agree or disagree
4-somewhat agree
5-strongly agree

Figure C-10: Rating Question 10



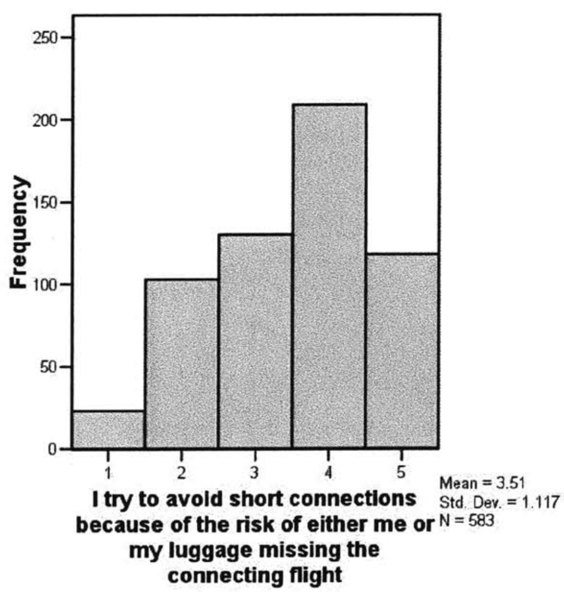
1-strongly disagree
2-somewhat disagree
3-neither agree or disagree
4-somewhat agree
5-strongly agree

Figure C-11: Rating Question 11



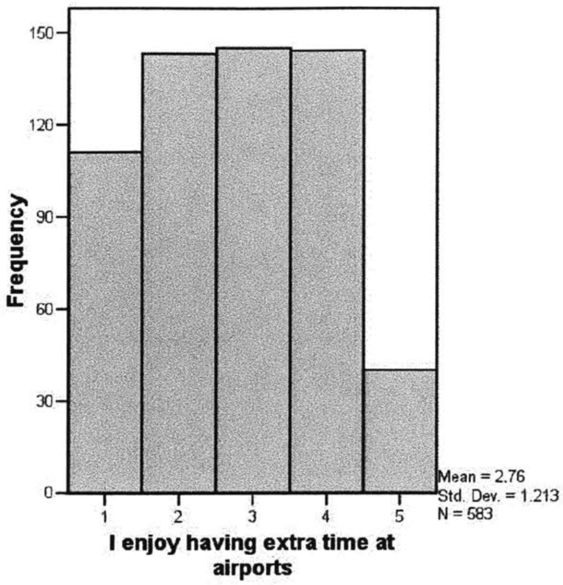
1-strongly disagree
2-somewhat disagree
3-neither agree or disagree
4-somewhat agree
5-strongly agree

Figure C-12: Rating Question 12



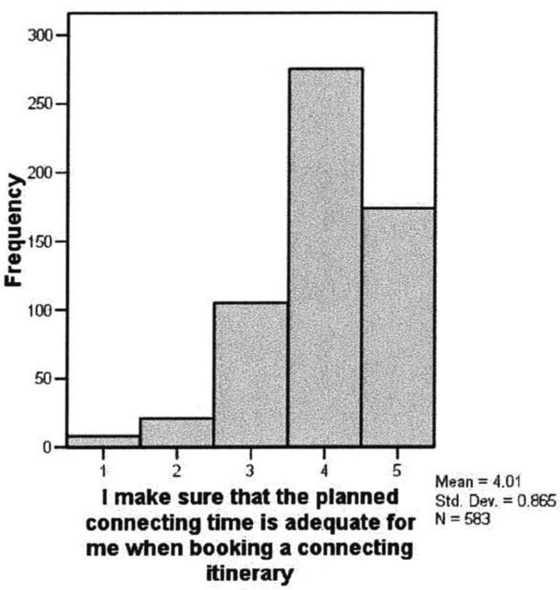
1-strongly disagree
2-somewhat disagree
3-neither agree or disagree
4-somewhat agree
5-strongly agree

Figure C-13: Rating Question 13



- 1-strongly disagree
- 2-somewhat disagree
- 3-neither agree or disagree
- 4-somewhat agree
- 5-strongly agree

Figure C-14: Rating Question 14



- 1-strongly disagree
- 2-somewhat disagree
- 3-neither agree or disagree
- 4-somewhat agree
- 5-strongly agree

Appendix D: Stability of Parameter Estimates

Table D-1: Stability of Multiple Indicators Multiple Causes Model

Measurement Equation (13 equations, one per row)

		Intercept				Risk			
No. of Halton draws		150 draws	150 draws	500 draws		150 draws	150 draws	500 draws	
Starting Value		0.1	0.5	0.5		0.1	0.5	0.5	
Equation	Rating Q.	Est	Est	Est	Std Err	Est	Est	Est	Std Err
1	1	3.52	3.52	3.52	0.022	-0.330	-0.329	-0.328	0.019
2	2	2.52	2.52	2.52	0.023	-0.158	-0.159	-0.154	0.023
3	3	3.76	3.76	3.76	0.022	0.548	0.550	0.547	0.021
4	4	2.65	2.65	2.65	0.021				
5	5	2.71	2.71	2.71	0.021	0.676	0.677	0.672	0.024
6	6	1.90	1.89	1.90	0.029	0.360	0.361	0.357	0.024
7	8	3.77	3.77	3.77	0.019				
8	9	3.19	3.19	3.19	0.023				
9	10	3.17	3.17	3.17	0.021	0.765	0.769	0.763	0.024
10	11	2.98	2.97	2.98	0.018	0.248	0.250	0.242	0.019
11	12	3.51	3.51	3.51	0.020	-0.429	-0.427	-0.425	0.021
12	13	2.76	2.76	2.76	0.021				
13	14	4.01	4.01	4.01	0.019	-0.213	-0.213	-0.211	0.016

		Rush				Trust				Disturbance Parameter			
No. of Halton draws		150 draws	150 draws	500 draws		150 draws	150 draws	500 draws		150 draws	150 draws	500 draws	
Starting Value		0.1	0.5	0.5		0.1	0.5	0.5		0.1	0.5	0.5	
Equation	Rating Q.	Est	Est	Est	Std Err	Est	Est	Est	Std Err	Est	Est	Est	Std Err
1	1	0.470	0.473	0.467	0.022					0.966	0.965	0.965	0.014
2	2	0.109	0.108	0.107	0.022	-0.351	-0.340	-0.339	0.026	1.141	1.145	1.145	0.021
3	3									0.923	0.923	0.920	0.014
4	4	0.968	0.965	0.967	0.025					0.716	0.719	0.708	0.018
5	5									0.940	0.941	0.938	0.015
6	6	0.122	0.121	0.120	0.019	-0.0666	-0.0646	-0.0655	0.024	1.033	1.033	1.033	0.016
7	8	0.212	0.215	0.214	0.017	-0.502	-0.513	-0.539	0.025	0.767	0.758	0.736	0.017
8	9	0.194	0.192	0.192	0.022	0.137	0.140	0.136	0.024	1.200	1.200	1.200	0.023
9	10	-0.118	-0.119	-0.115	0.018					0.832	0.831	0.826	0.015
10	11					0.472	0.473	0.477	0.025	0.923	0.923	0.920	0.017
11	12	0.285	0.286	0.282	0.019					0.961	0.962	0.961	0.014
12	13	0.931	0.928	0.924	0.025					0.732	0.734	0.732	0.016
13	14									0.833	0.833	0.833	0.010

Structural Equation (3 equations, 1 per column)

	Risk				Rush				Trust			
	150 draws	150 draws	500 draws		150 draws	150 draws	500 draws		150 draws	150 draws	500 draws	
	0.1	0	0		0.1	0	0		0.1	0	0	
	Est	Est	Est	Std Err	Est	Est	Est	Std Err	Est	Est	Est	Std Err
female	-0.253	-0.249	-0.245	0.049	0.0984	0.0936	0.0967	0.042	-0.198	-0.200	-0.187	0.058
missed	0.203	0.197	0.194	0.066	-0.259	-0.259	-0.257	0.061	-0.877	-0.882	-0.859	0.081
elite	0.236	0.222	0.231	0.071	0.511	0.504	0.507	0.060	0.210	0.217	0.229	0.078
two to five trips	-0.098	-0.103	-0.098	0.048	0.0117	0.0151	0.0116	0.046	0.168	0.169	0.168	0.065
more than five trips	0.280	0.274	0.272	0.067	0.114	0.120	0.121	0.061	0.0376	0.0271	0.0353	0.087
business	-0.0340	-0.0204	-0.0201	0.10	0.416	0.409	0.438	0.083	0.493	0.494	0.460	0.094
vacation	-0.0168	-0.0100	-0.0154	0.055	0.0463	0.0453	0.0389	0.046	0.0507	0.0494	0.0430	0.061
online	0.0199	0.0145	0.0150	0.048	0.0848	0.0866	0.0845	0.043	0.363	0.3491	0.3329	0.064
company	-0.243	-0.259	-0.260	0.11	-0.152	-0.167	-0.193	0.086	-0.439	-0.455	-0.450	0.111
alone	-0.00666	-0.00207	-0.00950	0.061	0.0225	0.0368	0.0274	0.052	0.113	0.101	0.106	0.070
couple	-0.0590	-0.0567	-0.0634	0.059	-0.159	-0.155	-0.163	0.054	-0.0747	-0.0730	-0.0666	0.067
max three nights	0.300	0.296	0.293	0.063	0.0286	0.0506	0.0280	0.060	-0.0889	-0.0977	-0.0845	0.077
four to seven nights	0.211	0.213	0.208	0.055	-0.150	-0.121	-0.145	0.052	-0.0217	-0.0342	-0.0182	0.065
employed	-0.0602	-0.0708	-0.0565	0.058	0.120	0.110	0.114	0.053	-0.0358	-0.0275	-0.0294	0.074
student	0.322	0.326	0.339	0.13	0.106	0.084	0.092	0.11	0.286	0.264	0.243	0.147
age < 30	-0.0167	-0.0153	-0.0151	0.013	0.00695	0.00578	0.00650	0.013	-0.0602	-0.0581	-0.0597	0.016
age 30 to 50	-0.00447	-0.00501	-0.00497	0.0040	0.00253	0.00219	0.00245	0.0040	0.00480	0.00368	0.00497	0.005
age 50+	-0.0362	-0.0360	-0.0361	0.0060	-0.00812	-0.00794	-0.00831	0.0050	-0.00409	-0.00381	-0.00540	0.008
income < 30	-0.00492	-0.00355	-0.00341	0.0070	-0.00567	-0.00715	-0.00533	0.0060	0.0485	0.0480	0.0450	0.010
income 30 to 75	-0.000159	-0.000301	-0.000758	0.0020	0.00227	0.00278	0.00226	0.0020	0.000536	0.0002959	0.000419	0.0020
income > 75	-0.000467	-0.000585	-0.000288	0.0010	-0.0120	-0.0118	-0.0119	0.0010	-0.00132	-0.00137	-0.00123	0.0020
bags	-0.166	-0.161	-0.177	0.055	0.122	0.128	0.126	0.049	0.0103	-0.0043	-0.0005	0.062

Table D-2: Stability of Integrated Choice and Latent Variable Model (Interaction with Risk) – Choice Model

Choice Model

	150	500	
Parameter	Est	Est	Std Err
Alternative specific constant current trip	1.102	1.103	0.078
ln (fare)	-4.008	-4.008	0.148
Elite frequent flyer program member	0.516	0.516	0.208
Standard frequent flyer program member	0.199	0.198	0.103
Ontime performance	0.0114	0.0115	0.003
Preferred airlines	0.147	0.147	0.069
First airport choice	0.791	0.791	0.10
Second airport choice	0.312	0.311	0.093
Number of connections	-0.368	-0.367	0.132
Night Departure	-0.425	-0.424	0.157
Access time	-0.00443	-0.00443	0.001
Elapsed time	-0.00318	-0.00318	0.000
Minimum Connecting Time	-0.00714	-0.00716	0.003
Buffertime < 15 min in min	0.0126	0.0126	0.005
Buffertime 15-59 min in min	-0.00404	-0.00399	0.002
Buffertime > 60 min in min	-0.00145	-0.00145	0.002
Interactions:			
Buffertime < 15 min in min and risk tolerance	-0.0134	-0.0134	0.005
Buffertime 15-59 min in min and risk tolerance	0.00895	0.00891	0.004
Buffertime > 60 min in min and risk tolerance	-0.00251	-0.00237	0.002
Number of connections and rush aversion	0.0450	0.0406	0.062
Number of connections and trust	0.0844	0.0889	0.072
Log-likelihood	-1653.57	-1653.65	

Table D-3: Stability of Integrated Choice and Latent Variable Model (Interaction with Risk) – Latent Variable Model

Measurement Equation (13 equations, one per row)

		Intercept			Risk			Rush			Trust			Disturbance Parameter		
No. of Halton draws		150 draws	500 draws		150 draws	500 draws		150 draws	500 draws		150 draws	500 draws		150 draws	500 draws	
Equation	Rating Q.	Est	Est	Std Err	Est	Est	Std Err	Est	Est	Std Err	Est	Est	Std Err	Est	Est	Std Err
1	1	3.52	3.52	0.022	-0.329	-0.327	0.02	0.473	0.466	0.022				0.965	0.965	0.014
2	2	2.52	2.52	0.023	-0.159	-0.153	0.023	0.107	0.106	0.022	-0.340	-0.340	0.026	1.144	1.145	0.021
3	3	3.76	3.76	0.022	0.551	0.548	0.021							0.922	0.92	0.014
4	4	2.65	2.65	0.021				0.965	0.965	0.025				0.719	0.708	0.018
5	5	2.71	2.71	0.021	0.676	0.670	0.024							0.943	0.94	0.015
6	6	1.89	1.90	0.03	0.360	0.357	0.024	0.121	0.121	0.019	-0.0644	-0.0665	0.025	1.034	1.033	0.016
7	8	3.77	3.77	0.019				0.215	0.214	0.017	-0.513	-0.540	0.025	0.758	0.736	0.018
8	9	3.19	3.19	0.023				0.192	0.192	0.022	0.140	0.134	0.024	1.2	1.2	0.023
9	10	3.17	3.17	0.021	0.770	0.764	0.024	-0.118	-0.112	0.018				0.829	0.826	0.015
10	11	2.97	2.98	0.018	0.250	0.241	0.02				0.473	0.476	0.025	0.923	0.921	0.017
11	12	3.51	3.51	0.02	-0.427	-0.425	0.021	0.285	0.280	0.019				0.962	0.961	0.014
12	13	2.76	2.76	0.021				0.929	0.924	0.025				0.734	0.731	0.016
13	14	4.01	4.01	0.019	-0.213	-0.212	0.016							0.833	0.833	0.01

Structural Equation (3 equations, 1 per column)

	Risk			Rush			Trust		
	150 draws	500 draws		150 draws	500 draws		150 draws	500 draws	
	Est	Est	Std Err	Est	Est	Std Err	Est	Est	Std Err
female	-0.249	-0.241	0.049	0.0933	0.0985	0.042	-0.202	-0.186	0.059
missed	0.198	0.202	0.066	-0.259	-0.258	0.061	-0.881	-0.855	0.082
elite	0.220	0.233	0.071	0.504	0.506	0.061	0.219	0.238	0.079
two to five trips	-0.102	-0.092	0.048	0.0146	0.0107	0.046	0.168	0.167	0.066
more than five trips	0.274	0.273	0.067	0.119	0.129	0.061	0.0262	0.0348	0.087
business	-0.0192	-0.0248	0.104	0.409	0.442	0.083	0.494	0.458	0.094
vacation	-0.0111	-0.0113	0.055	0.0449	0.0376	0.046	0.0476	0.0374	0.062
online	0.0154	0.0145	0.049	0.0866	0.0813	0.044	0.349	0.333	0.064
company	-0.260	-0.245	0.11	-0.167	-0.194	0.087	-0.456	-0.451	0.112
alone	-0.00485	-0.02320	0.061	0.0376	0.0374	0.052	0.102	0.102	0.07
couple	-0.0598	-0.0616	0.059	-0.155	-0.165	0.053	-0.0719	-0.0691	0.068
max three nights	0.294	0.288	0.063	0.0504	0.0273	0.06	-0.0969	-0.0843	0.077
four to seven nights	0.211	0.199	0.055	-0.121	-0.142	0.052	-0.0338	-0.0164	0.065
employed	-0.0725	-0.0564	0.058	0.109	0.107	0.053	-0.0264	-0.0260	0.074
student	0.326	0.343	0.126	0.0838	0.0851	0.115	0.268	0.255	0.149
age < 30	-0.0157	-0.0156	0.014	0.00592	0.00528	0.013	-0.058	-0.059	0.016
age 30 to 50	-0.00489	-0.00481	0.004	0.00212	0.00287	0.004	0.00351	0.00473	0.005
age 50+	-0.0359	-0.0366	0.006	-0.00787	-0.00801	0.005	-0.00363	-0.00506	0.008
income < 30	-0.00361	-0.00357	0.007	-0.00720	-0.00577	0.006	0.0481	0.0453	0.01
income 30 to 75	-0.000209	-0.000865	0.002	0.00282	0.00273	0.002	0.000254	0.000240	0.002
income > 75	-0.000615	-0.000279	0.001	-0.0118	-0.0121	0.001	-0.00135	-0.00109	0.002
bags	-0.162	-0.184	0.055	0.128	0.134	0.049	-0.00200	-0.00585	0.062

Table D-4: Stability of Integrated Choice and Latent Variable Model (Interaction with Rush) – Choice Model

Choice Model

	150	500	
Parameter	Est	Est	Std Err
Alternative specific constant current trip	1.093	1.093	0.078
ln (fare)	-4.028	-4.028	0.148
Elite frequent flyer program member	0.464	0.463	0.208
Standard frequent flyer program member	0.213	0.213	0.103
Ontime performance	0.0114	0.0115	0.003
Preferred airlines	0.157	0.157	0.069
First airport choice	0.805	0.805	0.10
Second airport choice	0.306	0.306	0.093
Number of connections	-0.417	-0.418	0.132
Night Departure	-0.389	-0.389	0.157
Access time	-0.00436	-0.00436	0.001
Elapsed time	-0.00319	-0.00318	0.000
Minimum Connecting Time	-0.00655	-0.00656	0.003
Buffertime < 15 min in min	0.0113	0.0113	0.005
Buffertime 15-59 min in min	-0.00396	-0.00397	0.002
Buffertime > 60 min in min	-0.00141	-0.00141	0.002
Interactions:			
Buffertime < 15 min in min and rush aversion	0.0198	0.0193	0.006
Buffertime 15-59 min in min and rush aversion	-0.0072	-0.00671	0.003
Buffertime > 60 min in min and rush aversion	0.0012	0.00117	0.001
Number of connections and risk tolerance	0.102	0.107	0.065
Number of connections and trust	0.073	0.072	0.072
Log-likelihood	-1653.33	-1653.51	

Table D-5: Stability of Integrated Choice and Latent Variable Model (Interaction with Rush) – Latent Variable Model

Measurement Equation (13 equations, one per row)

		Intercept			Risk			Rush			Trust			Disturbance Parameter		
No. of Halton draws		150 draws	500 draws		150 draws	500 draws		150 draws	500 draws		150 draws	500 draws		150 draws	500 draws	
Equation	Rating Q.	Est	Est	Std Err	Est	Est	Std Err	Est	Est	Std Err	Est	Est	Std Err	Est	Est	Std Err
1	1	3.52	3.52	0.022	-0.328	-0.326	0.020	0.474	0.467	0.022				0.9661	0.966	0.014
2	2	2.52	2.52	0.023	-0.159	-0.153	0.023	0.108	0.107	0.022	-0.340	-0.340	0.026	1.145	1.146	0.021
3	3	3.76	3.76	0.022	0.550	0.547	0.021							0.9243	0.921	0.014
4	4	2.65	2.65	0.021				0.965	0.965	0.025				0.7208	0.710	0.018
5	5	2.71	2.71	0.021	0.678	0.6725	0.024							0.9416	0.939	0.015
6	6	1.89	1.89	0.030	0.361	0.358	0.024	0.121	0.121	0.019	-0.0644	-0.0664	0.025	1.0340	1.0340	0.016
7	8	3.77	3.77	0.019				0.215	0.214	0.017	-0.513	-0.539	0.025	0.759	0.738	0.018
8	9	3.19	3.19	0.023				0.192	0.192	0.023	0.140	0.134	0.024	1.201	1.201	0.023
9	10	3.17	3.17	0.021	0.768	0.762	0.024	-0.120	-0.114	0.018				0.832	0.829	0.015
10	11	2.97	2.98	0.018	0.249	0.241	0.020				0.473	0.477	0.025	0.924	0.922	0.017
11	12	3.51	3.51	0.020	-0.426	-0.424	0.021	0.287	0.281	0.019				0.963	0.962	0.014
12	13	2.76	2.76	0.021				0.929	0.924	0.025				0.735	0.732	0.016
13	14	4.01	4.01	0.019	-0.213	-0.212	0.016							0.834	0.834	0.010

Structural Equation (3 equations, 1 per column)

	Risk			Rush			Trust		
	150 draws	500 draws		150 draws	500 draws		150 draws	500 draws	
	Est	Est	Std Err	Est	Est	Std Err	Est	Est	Std Err
female	-0.250	-0.241	0.049	0.0941	0.0992	0.042	-0.201	-0.186	0.059
missed	0.198	0.201	0.066	-0.259	-0.259	0.061	-0.882	-0.856	0.082
elite	0.223	0.235	0.071	0.503	0.506	0.061	0.217	0.236	0.079
two to five trips	-0.103	-0.094	0.048	0.0142	0.00990	0.046	0.169	0.167	0.066
more than five trips	0.275	0.272	0.067	0.122	0.130	0.061	0.0271	0.0350	0.087
business	-0.0202	-0.0259	0.104	0.407	0.440	0.083	0.494	0.458	0.094
vacation	-0.0105	-0.0118	0.055	0.0441	0.0369	0.046	0.0482	0.0377	0.062
online	0.0144	0.0126	0.049	0.0868	0.0805	0.044	0.349	0.333	0.064
company	-0.260	-0.245	0.11	-0.166	-0.193	0.087	-0.456	-0.451	0.112
alone	-0.00173	-0.0197	0.061	0.0377	0.0385	0.052	0.102	0.102	0.070
couple	-0.0562	-0.0575	0.059	-0.152	-0.163	0.054	-0.0721	-0.0692	0.068
max three nights	0.297	0.291	0.063	0.0511	0.0279	0.06	-0.0976	-0.0850	0.077
four to seven nights	0.214	0.202	0.055	-0.120	-0.142	0.052	-0.0341	-0.0166	0.065
employed	-0.0699	-0.0540	0.058	0.111	0.108	0.053	-0.0267	-0.0257	0.074
student	0.328	0.346	0.127	0.0848	0.0853	0.116	0.268	0.255	0.149
age < 30	-0.0150	-0.0150	0.014	0.00620	0.00569	0.013	-0.058	-0.0591	0.016
age 30 to 50	-0.00512	-0.00500	0.004	0.00214	0.00279	0.004	0.00354	0.00475	0.0050
age 50+	-0.0358	-0.0366	0.006	-0.00796	-0.00792	0.005	-0.00363	-0.0051	0.0080
income < 30	-0.00347	-0.00339	0.007	-0.00691	-0.00559	0.006	0.0481	0.0453	0.0100
income 30 to 75	-0.000319	-0.000978	0.002	0.00266	0.00264	0.002	0.000271	0.000231	0.0020
income > 75	-0.000587	-0.000254	0.001	-0.0117	-0.0121	0.001	-0.00136	-0.00109	0.0020
bags	-0.160	-0.182	0.055	0.130	0.135	0.049	-0.00248	-0.00628	0.062

Table D-6: Stability of Integrated Choice and Latent Variable Model (Interaction with Trust) – Choice Model

Choice Model

	150	500	
Parameter	Est	Est	Std Err
Alternative specific constant current trip	1.102	1.103	0.078
ln (fare)	-4.0270	-4.0260	0.15
Elite frequent flyer program member	0.507	0.510	0.21
Standard frequent flyer program member	0.188	0.189	0.10
Ontime performance	0.0115	0.0116	0.003
Preferred airlines	0.142	0.140	0.069
First airport choice	0.811	0.811	0.10
Second airport choice	0.313	0.312	0.093
Number of connections	-0.432	-0.432	0.13
Night Departure	-0.421	-0.424	0.16
Access time	-0.00440	-0.00438	0.001
Elapsed time	-0.00313	-0.00313	0.000
Minimum Connecting Time	-0.00626	-0.00621	0.003
Buffertime < 15 min in min	0.0131	0.0131	0.005
Buffertime 15-59 min in min	-0.00336	-0.00335	0.002
Buffertime > 60 min in min	-0.00154	-0.00156	0.002
Interactions:			
Buffertime < 15 min in min and trust	-0.0223	-0.0219	0.007
Buffertime 15-59 min in min and trust	0.00801	0.00817	0.004
Buffertime > 60 min in min and trust	0.00130	0.00131	0.002
Log-likelihood	-1648.97	-1649.12	

Table D-7: Stability of Integrated Choice and Latent Variable Model (Interaction with Trust) – Latent Variable Model

Measurement Equation (13 equations, one per row)

		Intercept			Risk			Rush			Trust			Disturbance Parameter		
No. of Halton draws		150 draws	500 draws		150 draws	500 draws		150 draws	500 draws		150 draws	500 draws		150 draws	500 draws	
Equation	Rating Q.	Est	Est	Std Err	Est	Est	Std Err	Est	Est	Std Err	Est	Est	Std Err	Est	Est	Std Err
1	1	3.52	3.52	0.022	-0.329	-0.327	0.02	0.473	0.466	0.022				0.966	0.966	0.014
2	2	2.52	2.52	0.023	-0.159	-0.153	0.023	0.107	0.106	0.022	-0.342	-0.343	0.026	1.145	1.145	0.021
3	3	3.76	3.76	0.022	0.550	0.547	0.021							0.924	0.922	0.014
4	4	2.65	2.65	0.021				0.965	0.966	0.025				0.720	0.709	0.018
5	5	2.71	2.71	0.021	0.677	0.672	0.024							0.942	0.939	0.015
6	6	1.89	1.89	0.030	0.361	0.358	0.024	0.121	0.121	0.019	-0.0660	-0.0683	0.025	1.0340	1.0340	0.016
7	8	3.77	3.77	0.019				0.213	0.213	0.017	-0.508	-0.534	0.025	0.763	0.742	0.017
8	9	3.19	3.19	0.023				0.192	0.193	0.022	0.140	0.133	0.024	1.201	1.201	0.023
9	10	3.17	3.17	0.021	0.768	0.7619	0.024	-0.119	-0.113	0.018				0.832	0.829	0.015
10	11	2.98	2.98	0.018	0.249	0.241	0.019				0.474	0.477	0.025	0.923	0.921	0.017
11	12	3.51	3.51	0.020	-0.426	-0.425	0.021	0.286	0.280	0.019				0.963	0.962	0.014
12	13	2.76	2.76	0.021				0.929	0.924	0.025				0.735	0.733	0.016
13	14	4.01	4.01	0.019	-0.213	-0.213	0.016							0.834	0.834	0.010

Structural Equation (3 equations, 1 per column)

	Risk			Rush			Trust		
	150 draws	500 draws		150 draws	500 draws		150 draws	500 draws	
	Est	Est	Std Err	Est	Est	Std Err	Est	Est	Std Err
female	-0.249	-0.240	0.049	0.093	0.099	0.042	-0.205	-0.189	0.059
missed	0.197	0.201	0.066	-0.259	-0.258	0.061	-0.884	-0.859	0.083
elite	0.224	0.235	0.071	0.504	0.505	0.061	0.210	0.227	0.079
two to five trips	-0.103	-0.094	0.048	0.0148	0.0112	0.046	0.169	0.167	0.066
more than five trips	0.274	0.273	0.067	0.119	0.129	0.061	0.0193	0.0307	0.087
business	-0.0200	-0.0261	0.104	0.409	0.442	0.083	0.505	0.468	0.095
vacation	-0.0103	-0.0106	0.055	0.0444	0.0382	0.046	0.0505	0.0398	0.062
online	0.0149	0.0135	0.049	0.0864	0.0808	0.044	0.348	0.334	0.064
company	-0.260	-0.245	0.11	-0.166	-0.194	0.086	-0.462	-0.458	0.11
alone	-0.00294	-0.0203	0.061	0.0362	0.0379	0.052	0.100	0.0997	0.07
couple	-0.0571	-0.0583	0.059	-0.155	-0.165	0.053	-0.0774	-0.0745	0.068
max three nights	0.296	0.291	0.063	0.0511	0.0273	0.06	-0.0979	-0.0863	0.077
four to seven nights	0.213	0.201	0.055	-0.121	-0.143	0.052	-0.0384	-0.0215	0.065
employed	-0.0702	-0.0540	0.058	0.109	0.106	0.053	-0.0333	-0.0314	0.074
student	0.325	0.342	0.127	0.0846	0.0847	0.115	0.277	0.2641	0.15
age < 30	-0.0153	-0.0154	0.013	0.00586	0.00528	0.013	-0.0583	-0.0597	0.016
age 30 to 50	-0.00500	-0.00490	0.004	0.00212	0.00283	0.004	0.00351	0.0048	0.005
age 50+	-0.0361	-0.0367	0.006	-0.00786	-0.00797	0.005	-0.00331	-0.00487	0.008
income < 30	-0.00354	-0.00349	0.007	-0.00715	-0.00585	0.006	0.0477	0.04512	0.01
income 30 to 75	-0.000348	-0.000959	0.002	0.00279	0.00278	0.002	0.000687	0.000579	0.002
income > 75	-0.000570	-0.000264	0.001	-0.0118	-0.0121	0.001	-0.00151	-0.00123	0.002
bags	-0.161	-0.183	0.055	0.128	0.134	0.049	-0.00626	-0.00890	0.063

Bibliography

Adler, T., C. Falzarano and G. Spitz (2005), "Modeling Service Trade-offs in Air Itinerary Choices", paper presented at the 44th Annual Meeting of the Transportation Research Board, Washington, DC.

Air Canada (2007), Management's Discussion and Analysis, Quarter 2, 2007. http://www.aircanada.com/en/about/investor/documents/2007_MDA_q2.pdf, accessed Aug 10, 2007.

Alamdari, F. E. and I. G. Black (1992). "Passengers Choice of Airline under Competition - the Use of the Logit Model." Transport Reviews 12(2): 153-170.

Airline Business (2002). "American cuts operations yet again", in: Airline Business, Sep. 2002, p. 15.

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

Ben-Akiva, M. and B. Boccara (1987), Integrated Framework for travel behavior analysis, Presented at IATBR conference, Aix-en-Provence, France.

Ben-Akiva, M., D. McFadden, T. Gärling, D. Gopinath, J. Walker, D. Bolduc, A. Börsch-Supan, P. Delquié, O. Larichev, T. Morikawa, A. Polydoropoulou and V. Rao (1999) "Extended Framework for Modeling Choice Behavior", Marketing Letters 10(3):187-203.

Ben-Akiva, M., D. McFadden, K. Train, J.Walker, C. Bhat, M. Bierlaire, D. Bolduc, A. Boersch-Supan, D. Brownstone, D. Bunch, A. Daly, A. De Palma, D. Gopinath, A. Karlstrom and M. Munizagam (2002a) "Hybrid Choice Models: Progress and Challenges", Marketing Letters 13(3): 163-175.

Ben-Akiva, M., J. Walker, A. Bernardino, D. Gopinath, T. Morikawa, A. Polydoropoulou (2002b) "Integration of Choice and Latent Variable Models", in (H. Mahmassani, Ed.) In Perpetual Motion: Travel Behaviour Research Opportunities and Application Challenges, Elsevier Science, 431-470.

Bernardino, A. (1996). Telecommuting: Modeling the Employer's and the Employee's Decision Making Process. Ph.D. dissertation. Massachusetts Institute of Technology.

Boeing Commercial Airplane Group (1993). "Decision Window, Path Preference Methodology, Time Mode." AMADWM-1. Handout for MIT 16.75 Airline Management Class.

Bolduc, D. and A. Giroux (2005). The Integrated Choice and Latent Variable Model - Handout to accompany estimation software, Working paper.

Bolduc, D., N. Boucher, R. Alvarez-Daziano (2008) "Hybrid Choice Modeling of New Technologies for Car Choice in Canada", Transportation Research Record No. 2082: pp. 63-71.

Bollen, K. (1989), Structural Equations with Latent Variables, New York.

Bradley, M. A. (1998), Behavioural models of airport choice and air route choice, in J. de D. Ortuzar, D. Hensher & S. R. Jara-Diaz, eds, "Travel behaviour research: updating the state of play (IATBR 94)", Elsevier, Oxford, pp. 141–159.

Cambridge Systematics (Inc.) (1986),Customer Preference and Behavior Project Report, prepared for the Electric Power Research Institute.

Churchill, G. and D. Iacobucci (2002), Marketing Research – Methodological Foundations, 8th edition, Mason, OH.

Coldren, G. and F. Koppelman (2003), "Modeling the competition among air travel itinerary shares: GEV model development, paper presented at the 8th IATBR Conference, Switzerland, 10-15 August 2003.

Coldren, G. and F. Koppelman (2005), "Modeling the competition among air-travel itinerary shares: Gev model development", Transportation Research Part A: Policy and Practice 39(4), 345–365.

Coldren, G., F. Koppelman, K. Kasturirangan and A. Mukherjee (2003), "Modeling aggregate air-travel itinerary shares: logit model development at a major US airline", Journal of Air Transport Management 9(6), 361–369.

Department of Transportation (2003), Regulation 14CFR Part 255, 4a(2), http://www.access.gpo.gov/nara/cfr/waisidx_03/14cfr255_03.html.

Department of Transportation (2004), Final Rule, DOT 14CFR Part 255, RIN 2105-AC65, http://dmses.dot.gov/docimages/pdf88/263280_web.pdf.

European Council (1999), Council Regulation (EC) No. 323/1999, Annex 1.

Flint, P. (2002). No Peaking. Air Transport World, 39,11, p. 22-27.

Goetz, A. and C. Sutton (1997), "The Geography of Deregulation in the U.S. airline industry", Annals of the Association of American Geographers, 87, 2, pp. 238-263.

Hess, S., T.Adler and J. Polak (2005), "Computing willingness-to-pay indicators for air-travellers from SP survey data", Paper presented at Air Transport Research Conference, Rio de Janeiro.

Holloway, S. (2003), Straight and Level: Practical Airline Economics, Hampshire.

House of Lords (1998), European Communities Committee, 32nd report, 17 Nov 1998, Part 4, No. 51, http://www.parliament.the-stationery-office.co.uk/pa/ld199798/ldselect/ldeucom/156/15606.htm.

Jöreskog, K. (1973), A general method for estimating a linear structural equation system. In A.S. Golderberger & O.D. Duncan (eds.), Structural equation models in the social sciences, New York, p. 85-112.

Kanafani, A. and E. Sadoulet (1975), 'The partitioning of long haul air traffic - a study in multinomial choice', Transportation Research 11(1), 1–8.

Kanafani, A., and A. Ghobrial (1985), "Airline hubbing – some implications for airport Economics", Transportation Research Part A, 19, pp. 15-27.

Keesling, J.W. (1972), Maximum Likelihood approaches to causal flow analysis. Ph.D. dissertation, University of Chicago.

Kline, R. (2005), Principles and Practice of Structural Equation Modeling, 2nd edition, New York.

Mayer, C. and T. Sinai (2003): Network Effects, Congestion Externalities, and Air Traffic Delays: Or why not all Delays are Evil, American Economic Review, 93 (4), 1194-1215.

McFadden, D. (1986). The choice theory approach to market research. Marketing Science, 5, 275-297.

Morikawa, T., M. Ben-Akiva, D. McFadden (2002), "Discrete Choice Models Incorporating Revealed Preferences and Psychometric Data", Advances in Econometrics, Vol 16, 29-55.

Nason, S. D. (1980), 'Analyzing ticket-choice decisions of air travellers', TransportationResearch Record 768, 20–24.

Ndoh, N. N., D.E. Pitfield and R.R.Caves (1990), Air transportation passenger routechoice: a Nested Multinomial Logit analysis, in M. M. Fisher, P. Nijkamp & Y. Y.Papageorgiou, eds, 'Spatial Choices and Processes', Elsevier Science Publishers B.V.,North Holland, 49–365.

Nielsen/NetRatings (2005): "Online Travel Purchases split evenly between travel agencies and suppliers' web sites", Press Release June 21, 2005, http://www.nielsennetratings.com/pr/pr_050621.pdf, accessed June 30, 2005.

OAG (2004), Official Airline Guides, April 2004, Dunstable.

Polydoropoulou, A. (1997). Modeling User Response to Advanced Travelers Information Systems (ATIS). Ph.D. dissertation, Massachusetts Institute of Technology.

Proussaloglou, K. and F. Koppelman (1995). "Air Carrier Demand - an Analysis of Market Share Determinants." Transportation 22(4): 371-388.

Proussaloglou, K. and F. Koppelman (1999). "The Choice of Carrier, Flight and Fare Class." Journal of Air Transport Management 5(4): 193-201.

Schröder, A. (2004), Preiseffekte bei Passagierprognosen im Luftverkehr, unpublished "Diplomarbeit" (Masters Thesis), Technische Universität Clausthal.

Schumacker, R. and R. Lomax (2004), A Beginner's Guide to Structural Equation Modeling, 2nd edition, Mahwah.

Soncrant, C. and C. Hoppenstad (1998), "Airline Schedule Evaluation", in: Butler, G. and M. Keller (eds.), Handbook of Airline Marketing, New York, 601-612.

Smith, B. et al. (1998). "Airline Planning and Marketing Decision Support: A Review of Current Practices and Future Trends", in: Butler, G. and M. Keller (eds.), Handbook of Airline Marketing, New York,117-130.

Taubmann, H. (2004), "Chaos auf dem Monitor", Presentation held at "FVW Kongress", 14 October 2004.

Walker, J. (2001). Extended Discrete Choice Models: Integrated Framework, Flexible Error Structures, and Latent Variables. Ph.D. dissertation, Massachusetts Institute of Technology.

Walker, J. and M. Ben-Akiva (2002). "Generalized Random Utility Model." Mathematical Social Sciences 43 (3): 303-343.

Wiley, D. (1973). The identification problem for structural equation models with umeasured variables. In A.S. Golderberger & O.D. Duncan (eds.)., Structural equation models in the social sciences, New York, p. 69-83.