Modeling air travel behavior

Af Valdemar Warburg

Polyteknisk Midtvejsprojekt
10 point

Center for Trafik og Transport
Danmarks Tekniske Universitet

FORORD

Dette midtvejsprojekt på 10 point blev påbegyndt i januar 2005 ved University of Texas at Austin, hvor jeg læste et år ved Transportation Engineering-uddannelsen som led i udvekslingsprogrammet Global Engineering Education Exchange. Da jeg har estimeret modellerne og skrevet det meste af midtvejsprojektet i USA, er det skrevet på engelsk.

Jeg vil først og fremmest takke min vejleder, professor Otto Anker Nielsen, CTT for mange gode kommentarer til arbejdet samt for vejledning til litteratursøgning. Jeg vil også gerne takke min amerikanske vejleder, Dr. Chandra Bhat ved University of Texas for at have givet mig en virkelig god introduktion til diskrete valgmodeller gennem sine forelæsinger og for hjælp til de første dele af projektet. Desuden takkes Tom Adler og Gregory Spitz fra Resource Systems Group for udlån af data fra deres flypassagerundersøgelse. Også tak til Muriel Beser, Transek AB for at sende mig sin Ph.D-afhandling, der var meget relevant for mit projekt.

Nogle af resultaterne fra midtvejsprojektet er blevet til en artikel, *Modeling demographic and unobserved heterogeneity in air passengers' sensitivity to service attributes in itinerary choice* med Chandra Bhat og Tom Adler som medforfattere. Jeg skal fremlægge resultaterne mundtligt ved *Transportation Research Board Annual Meeting* i Washington, DC i januar 2006, og artiklen bliver trykt i foråret 2006 i tidskriftet *Transportation Research Record*.

Lyngby, den 24. oktober 2005

Valdemar Warburg

ABSTRACT

Modeling passengers' flight choice behavior is valuable to understanding the increasingly competitive airline market and predicting air travel demands. This report estimates standard and mixed multinomial logit models of itinerary choice for business travel, based on a stated preference survey conducted in 2001.

Previous work on air travel behavior modeling has almost exclusively been confined to studying either airport or airline choice. However, two recent papers have expanded the study area to cover air itinerary choice, recognizing the fact that the average air traveler focuses on all the attributes of getting from A to B. This report models itinerary choice as well as accommodating observed and unobserved variations in parameters, due to demographic and trip-related differences. This combination is new to the literature.

In this report, the random parameters in the mixed logit model take a Normal distribution. Although this distribution can cause counter-intuitive signs of the coefficient, it was preferred to the lognormal distribution, because of better convergence properties, comparisons with earlier work, and disappointing results with the lognormal distribution.

The dataset consists of 119 business passengers and 521 non-business passengers, who each report their most recent domestic flight. A computer then generates 10 itinerary alternatives with the same origin-destination, such that the respondent has to make 10 binary choices between the actual flight and the hypothetical flights. Many demographic variables are included in the dataset.

The multinomial logit results show that there is significant differences the preferences of business and non-business passengers. In general, business passengers are more time sensitive, and non-business passengers more fare sensitive. In addition, the business passengers can be split up into many demographic groups, each with its own preferences. Among demographics, gender and income level have the most noticeable effects on sensitivity to service attributes in itinerary choice behavior, but frequent flyer membership, employment status, travel frequency, and group travel also emerge as important determinants. Some of these variables are also affecting choice in the non-business segment.

The mixed logit models show in addition that there is significant residual heterogeneity due to unobserved factors even after accommodating sensitivity variations due to demographic and trip-related factors. Consequently, substitution rates for each service attribute show substantial variations in the willingness-to-pay among observationally identical business passengers. The results suggest that observed demographic and trip related differences get incorrectly manifested as unobserved heterogeneity in a random coefficients mixed logit model that ignores demographic and trip-related characteristics of travelers.

Willingness-to-pay values are somewhat close to the values in earlier work. However, there is significant random variation around the mean values of on time performance, access time, and number of connections. When weighting the substitution rates by

demographic shares to obtain a "typical" business traveler, the variation becomes narrower, and thus the WTP values more accurate.

Overall, the results from the study should aid in targeting and adjusting service products to specific passenger groups. An example is presented on how the results can be applied in an airline's pricing strategy. The study also facilitates a clear understanding of the air travel market, which should help air carriers design appropriate pricing schemes and better predict passenger demands when introducing new routes and services. Finally, the analysis can also assist in airport terminal planning and ground access through more accurate prediction of air travel demand.

TABLE OF CONTENTS

1 INTRODUCTION

Since deregulation and privatization of airlines in the United States, Europe, and Asia, airlines have been allowed to compete on a market basis. This has changed the industry dramatically. Many of the legacy carriers that once were owned by governments are struggling to make ends meet, and new low cost carriers are benefiting from the price sensitive customers. Due to this competitive airline market, predicting the travel choices of air passengers has become increasingly important.

This report uses a web-based Stated Preference survey to examine the flight itinerary choice behavior of business air travelers. The objectives of this report are:

- To identify the factors that affect passengers' choice of flight itineraries
- Examine observed and unobserved demographic and socio-economic variation among passengers
- Calculate substitution rates for different attributes of itineraries that airlines can use in determining traffic demands and pricing

There are several dimensions characterizing air travel choice behavior after a traveler has decided to travel to a particular destination from a particular origin. These include the origin and destination airports in multi-airport regions, the airline carrier choice, the desired departure and arrival times, fare, aircraft types, and airport access mode choice, among others. This report covers all the important factors.

One approach to modeling itinerary choices is to construct revealed preference (RP) data by using actual passenger loads. However, this method has the disadvantage that the analyst lacks information about the demographic variations across the decision makers. This issue is discussed in further detail in chapter 2. Therefore, stated preference data (SP) are increasingly being used by analysts to calculate trade-off values of passengers. Consequently, SP data is used in this report.

The survey used was conducted by Resource Systems Group in spring, 2001, prior to the events of 9/11. It represents one of the most comprehensive stated preference design experiments conducted in the air travel behavior field. The current report uses the data source and considers a wide range of air travel service characteristics, trip-related information, demographic attributes of the traveler, and interactions of these variables to model air itinerary choice behavior. The emphasis is on accommodating the different sensitivities across travelers to air service characteristics based on the trip-related and demographic attributes of the traveler. In addition, the report accommodates unobserved sensitivity variations across individuals in the business segment using a mixed multinomial logit model.

The ultimate goal is to be able to predict, as accurately as possible, what itineraries different groups of air passengers will choose, when booking a flight. The survey respondents reveal through the choices they make their willingness-to-pay for different service attributes. This information is extremely valuable for airlines to obtain.

1.1 Report structure

The following is an outline of the report with descriptions of the different chapters.

Chapter 1: Introduction
Background, methodology, and outline of the report.

Chapter 2: Literature review
A thorough presentation of the previous literature within air travel behavior modeling, with particular emphasis on two recent, relevant papers on the use of complex discrete choice models in the air travel field.

Chapter 3: Methodology
An overview and introduction to discrete choice models, particularly the multinomial logit model and the mixed logit model.

Chapter 4: Data
Description of the dataset from Resource Systems Group, which is used in this report. Important sample variables are discussed.

Chapter 5: Multinomial logit results
Discussion of the empirical results. First, statistical tests of market segmentation between business and non-business are performed. Next, the logit model results for the two segments are presented.

Chapter 6: Mixed logit results
Description of the results obtained from the mixed logit estimation simulation of the business segment.

Chapter 7: Trade-off analysis
Calculation of substitution rates of different parameters and an example of how these values can be applied in the airline industry.

Chapter 8: Conclusion
The final chapter concludes the report by summarizing the findings.

Appendix A: Survey instrument
List of the survey questions for the variables that are used in this report.

2 LITERATURE REVIEW

2.1 Background studies

Several studies have examined different aspects of air travel choice behavior. For example, Skinner (1976), Harvey (1987), Ashford and Benchemam (1987), Ozoka and Ashford (1989), Innes and Doucet (1990), Thompson and Caves (1993), Windle and Dresner (1995), Basar and Bhat (2004), Hess and Polak (2005), and Pathomsiri and Haghani (2005) all model airport choice in multi-airport regions. Other studies have modeled airport choice along with other dimensions of travel. For instance, Ndoh *et al.* (1990) examine passenger route choice and airport choice; Furiuchi and Koppelman (1994) study destination choice and airport choice; Pels *et al.* (2001) analyze airline and airport choice; and Pels *et al.* (2003) and Hess and Polak (2005) model the three dimensions of airport access mode choice, airline choice, and airport choice. A few studies have also focused on air traveler choices other than airport choice. These include Proussaloglou and Koppelman (1999), Chin (2002), Yoo and Ashford (1996), and Algers and Beser (2001), all of whom examine airline choice.

A majority of the studies discussed above have used a simple multinomial logit model of choice to examine air traveler behavior. A few studies, such as Furiuchi and Koppelman (1994), Ndoh *et al.* (1990), and Pels *et al.* (2001), have used a nested logit model to model multidimensional or spatial choices in air travel behavior. But it has been only recently that studies have attempted to consider such important behavioral issues as consideration effects in air travel choices (see Basar and Bhat, 2003) and variations in sensitivity across individuals due to unobserved factors (see Hess and Polak, 2004; 2005 and Pathomsiri and Haghani, 2005). These studies use a mixing structure over the multinomial logit kernel, either in the form of a discrete distribution (leading to a latent class model as in Basar and Bhat, 2003) or in the form of a continuous distribution (leading to the mixed multinomial logit model as in Hess and Polak, 2005 and Pathomsiri and Haghani, 2005). While being important methodological contributions, the above three studies (along with the rest of the studies discussed earlier) have been rather limited in their perspective of the choices that characterize air travel decisions. Specifically, the Basar and Bhat (2003), Hess and Polak (2004), and Pathomsiri and Haghani (2005) studies focus exclusively on airport choice, while the Hess and Polak (2005) study confines its attention to the choices of airport, airline, and access mode.

An additional issue with earlier studies that accommodate unobserved taste variations is that they do not adequately accommodate *observed* taste variations. As emphasized in Bhat (2003), it is critical to first accommodate systematic variations in as comprehensive a way as possible, so one is able to explain differences in sensitivity based on tangible, observed, attributes that can be used for targeting and marketing of air service improvements by air carriers and airport management. The introduction of unobserved taste variation should not be in lieu of observed taste variation, but only to recognize the inevitable presence of unobserved factors affecting sensitivities, even after the most comprehensive control for observed factors.

2.1.1 Revealed preference vs. stated preference data

Some of the studies have used revealed preference (RP) data by drawing on actual passenger loads, interpreting the shares as probabilities, and defining the choice set as all the itinerary combinations that are available to the decision maker. While RP data represents actual itinerary choices, and therefore provide important information about preferences in a real choice environment, it is unlikely that decision makers consider all itinerary combinations in their choice decisions. Another limitation of revealed preference data is the inability to obtain precise estimates of the sensitivity to various air service measures due to the (typically) limited variation in, and high correlation among, service attributes (see Gunn *et al.*, 1992; Swait and Louviere, 1993; Hensher, 1999; Bhat and Sardesai, 2004). Consequently, trade-off analysis is difficult using revealed preference data.

The limitations of revealed preference data for trade-off analysis has led to the widespread use of stated preference (SP) data (self-stated preferences for market products and services) in the marketing and travel demand fields, separately or in conjunction with revealed preference data, to analyze consumers' evaluation of multi-attributed products and services. The primary applications of stated preference data in the marketing field have focused on new product or concept evaluation, design of effective product repositioning strategies, market share projections under hypothetical market scenarios, and product pricing analysis (see Whitaker *et al.*, 2005 for a review). Those in the transportation field have addressed issues such as value of travel time savings, effectiveness of ride/sharing incentives and auto use disincentives in reducing traffic congestion, and ridership forecasts on a new, or substantially improved, transportation service (see Louviere *et al.*, 2002, Brownstone *et al.*, 2003, and Bhat and Sardesai, 2004). The limitations of SP data are associated with its reliability: respondents may be indifferent, uninterested, and careless in a survey, may consider only a few attributes rather than all attributes, may impute values for omitted variables, and may use the survey to express an opinion about the survey context rather than provide information about usage of a new product. However, some of these issues have been resolved through advances in the design of stated preference experiments in the past decade (see, for example, Huber and Zwerina, 1996; Carson *et al.*, 1994; Hensher, 2004). Further recent research suggests that respondents can comprehend and evaluate even seemingly complex choice scenarios, thereby offering the potential to increase the realism and information gained from stated choice experiments (see Hensher and Louviere, 1999; Hensher, 2004).

2.2 Two recent studies of importance

Two recent studies, and the ones most pertinent to this report in the context of addressing the limitations discussed, are Coldren and Koppelman (2005) and Adler *et al.* (2005). Both these studies consider the whole suite of choices (origin and destination, airports in multi-airport regions, airline, fare, departure and arrival times, airport type, and number of connections) using *itineraries* as the alternatives in their discrete models. This shift of focus from evaluating a few isolated air travel choice dimensions to analyzing the multi-dimensional set of choices implicit in selecting an itinerary is a significant one in the literature. After all, travelers choose from among various itineraries rather than choosing an airport or an airline. In the next two sections, the Coldren and Koppelman and Adler studies are discussed in more detail.

2.2.1 The Coldren and Koppelman study

This study uses air travel itinerary share data to estimate aggregate hybrid ordered generalized extreme value (OGEV) models to capture inter-itinerary competition. The data are based on detailed records of individual-booked itineraries obtained from a compilation of computer reservation systems (CRS). These bookings data are complemented with air carrier schedule information obtained from the Official Airline Guide (OAG Worldwide Limited, 2001) and average fares by carrier across all itineraries for each airport pair. The authors use the itinerary building engine of a major carrier to generate the set of feasible itineraries by airport pair, and obtain the share of each itinerary for each airport pair by merging the generated set of feasible itineraries with the bookings data from the CRS. These itinerary shares are modeled as a function of several service characteristics, including itinerary level of service indicators (nonstop, direct, single-connect or double-connect), connection quality, carrier attributes, aircraft type, and departure time.

The Coldren and Koppelman study is possibly the most comprehensive published effort that models itinerary choice using actual revealed preference bookings data. The data preparation in the research is a demanding exercise and should serve as a reference basis for data compilation in future revealed preference studies of itinerary choice. In addition, the authors use an ordered-generalized extreme value structure among the itineraries to accommodate the higher sensitivity between itineraries, which are "proximate" in departure time. Further, Coldren and Koppelman also consider the higher degree of sensitivity across itineraries sharing a common carrier and a common level-of-service indicator. The overall model takes the form of an Ordered Generalized Extreme Value Nested Logit (OGEV-NL) structure. The results show evidence of higher sensitivity among itineraries along the time, carrier, and level-of-service dimensions, as well as proximate covariance in departure time choice.

Overall, the Coldren and Koppelman study is an important contribution to the literature. However, there are three limitations of the study. First, the bookings data do not include individual demographics (gender, income, employment, *etc.*) and individual travel characteristics (group travel, frequency of travel, trip purpose, *etc.*) and as a result, only aggregate share models can be estimated with these data. Such share models cannot accommodate sensitivity variations to service attributes based on individual demographic and travel characteristics. Second, the fare data for all itineraries between an airport pair vary only by carrier, since itinerary-level fare data were not available to the authors. This limited fare variation among itineraries introduces additional error and potential biases in the estimation of willingness-to-pay. Third, it is unlikely that individuals consider all possible itineraries between airport pairs before making their choice. While Coldren and

Koppelman do not use individual-level models, their share model is based on the premise that individuals consider all itineraries within a pre-defined window when making their booking.

2.2.2 The Adler et al. study

Unlike the Coldren and Koppelman study that focused on the better representation of the competitiveness structure (sensitivity) across itineraries, the Adler *et al.* (2005) study was motivated by a need to better understand the trade-offs in the many service characteristics in an increasingly option-laden airline industry. For example, low-fare airlines are positioning themselves in the market by flying out of more remote airports, flying circuitous routes with several transfers, and providing "no-frills" service. At the same time, the "legacy" airlines are re-positioning themselves through route and schedule re-alignments, pay-for-food services, and varying other service attributes such as seat spacing. Clearly, an understanding of the tradeoffs that individuals use in their itinerary choices becomes critical to airline managers in such an environment.

The Adler *et al.* study uses a 2003 internet-based stated preference survey that collected detailed information on the most recent paid domestic air trip of about 600 individuals. The web-based survey, which is annually conducted by Resource Systems Group, Inc., also obtained information from respondents on their preferred ticketing, airport, and airline alternatives, and implemented a stated choice experiment customized to the attributes of the respondent's reported trip. Specifically, a heuristic programmed into the survey software generates a "realistic" itinerary alternative for the respondent's reported trip. Ten such itinerary alternatives are constructed based on a fractional factorial experimental design and presented as alternatives to the actual reported itinerary in ten separate stated choice experiments for each individual. The attributes characterizing the itineraries in the stated choice experiments include airline carrier, airport, access/egress time, flight times, connections, fare, the time difference between the desired arrival time at destination and the scheduled arrival time of the itinerary, aircraft type, and on-time performance.

The authors use a mixed multinomial logit model to capture the sensitivity variations to the service attributes mentioned earlier. They find statistically significant, and expected, effects of all the service characteristics and significant effects from frequent flyer program status. The inclusion of unobserved sensitivity variations using the mixed logit model provides a statistically superior fit compared to the standard multinomial logit model. The authors conclude that "...mixed logit provides additional insights that can be used to further target air service alternatives".

The Adler *et al.* study is, like the Coldren and Koppelman study, an important contribution to the aviation demand literature. The stated preference design in the research reduces correlations among service attributes and facilitates an accurate trade-off analysis. But a limitation of the Adler study is that, like the Coldren and Koppelman study, it does not incorporate the full effects of demographics and trip characteristics on the sensitivity to service attributes.

2.3 The current report

The current report contributes to the itinerary choice models in the literature by examining the influence of service characteristics using data from a spring 2001 internet-based revealed and stated preference survey such as in the Adler *et al.* study. A mixed logit model is used to allow random taste variations in the sensitivity to service characteristics. However, in addition, this study examines taste variations due to a comprehensive set of demographic and trip characteristics of individuals. These characteristics are available in the data collected by Adler *et al.*, but were not explored in detail previously. The taste differences between various demographic and travel groups are highlighted and discussed. To focus the analysis, this study examines mixed logit models only for business travelers. However, standard multinomial logit models have been estimated for both segments, and the results are discussed in the report.

3 METHODOLOGY

This chapter provides a brief background of discrete choice modeling in general, as well as the two models used in this report: the multinomial logit model and the mixed logit model.

3.1 Discrete choice models

A discrete choice model is a common term for a mathematical model that can describe and predict individuals' choices from a finite choice set. Four elements describe the choice process, which is illustrated in Figure 1:

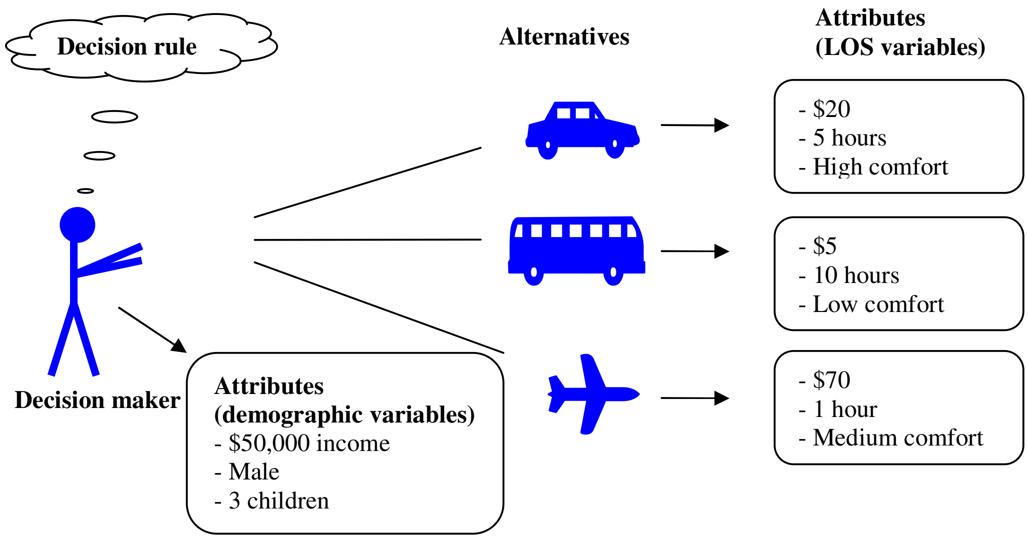

Figure 1. An illustration of the four elements in the choice process. In this case, there are three alternatives: car, bus, and air, and each of these alternatives are characterized by three attributes: price, time, and comfort.

- Decision maker: This is the individual, household, or entity that is making the choice.
- Alternatives: These are the choices available from the choice set. A universal choice set means that every decision maker has access to the same choices. For example, if we are studying mode choice between car, bus, and air, and all the individuals in our dataset has access to those three modes, we have a universal choice set. However, some individuals may only have access to some of the alternatives (e.g. car and bus) and some may have access to other alternatives (e.g. bicycle). In that case, there is a separate *feasible choice set* for each individual. There are two requirements for the choice set (Train, 2003): The alternatives have to be *mutually exclusive* (one cannot choose more than one alternative) and *collectively exhaustive* (all possible/feasible alternatives are included, meaning that the individual necessarily has to choose one alternative).
- Attributes: The attributes of the alternatives, such as price and time, are called *alternative specific variables*. Choices are often based on the values of the attributes in-

stead of the alternatives themselves. They are also called *level-of-service (LOS) variables,* because they describe the service level of each alternative. Additionally, the attributes of the decision maker, such as gender and income are called *individual specific* or *demographic variables*. These attributes play a major role in the choice, because different demographic groups have different preferences.

- Decision rule: This is the principle that the individual follows when choosing his/her alternative. The most commonly used approach is utility maximization, which will be discussed in section 3.1.1.

3.1.1 Utility maximization

Utility maximization introduces a scalar index of value, by placing rates on each of the attributes of the alternatives. By summing up the rates of an alternative, the *representative utility*, V is found. The representative utility V_{nj}, that decision maker n obtains from alternative j is a function of the attributes of the alternatives (LOS variables), x_{nj} and the attributes of the decision maker (demographic variables), s_n:

$$V_{nj} = V(x_{nj}, s_n) \tag{1}$$

However, as analysts, we cannot observe all the factors that affect the choice. This is why *random utility models* (RUMs) include an error term, ε_{nj}, which denotes the randomness of the utility function – in other words: all the unobserved factors are captured in ε_{nj}. The actual utility, U_{nj} then becomes:

$$U_{nj} = V_{nj} + \varepsilon_{nj} \tag{2}$$

The probability that the decision maker n chooses alternative i, is the probability that the utility of i is greater than the utilities of all the other alternatives j:

$$\begin{aligned}
\Pr_{ni} &= \Pr(U_{ni} > U_{nj} \quad \forall j \neq i) \\
&= \Pr(V_{ni} + \varepsilon_{ni} > V_{nj} + \varepsilon_{nj} \quad \forall j \neq i) \\
&= \Pr(\varepsilon_{nj} - \varepsilon_{ni} < V_{ni} - V_{nj} \quad \forall j \neq i)
\end{aligned} \tag{3}$$

It is now clear that if we know the distribution of ε_n, then the probability (3) can be calculated. The distribution of the error term is in fact the feature that distinguishes the two most widely used models, *logit* and *probit*. In this chapter, only the logit model will be discussed.

3.2 The multinomial logit model

In the multinomial logit model, the error term is characterized by having a Gumbel distribution, as shown in (4) and plotted in Figure 2:

$$f(\varepsilon) = \exp(-e^{-\varepsilon} - \varepsilon) \tag{4}$$

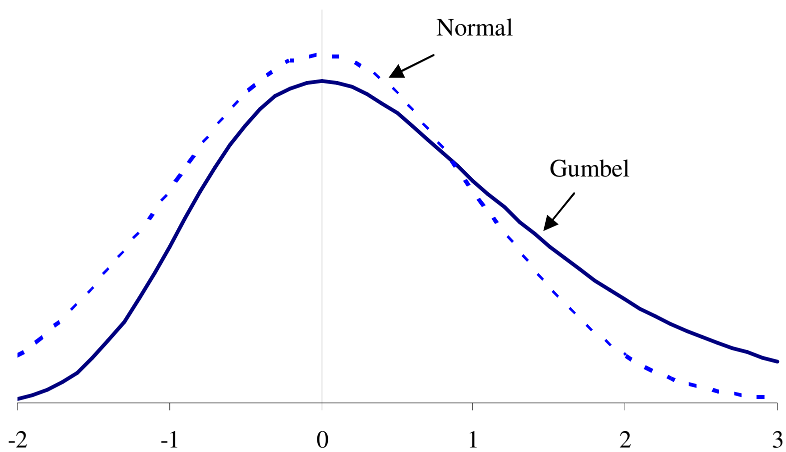

Figure 2. The Gumbel distribution is a very good approximation of the Normal distribution, and it has better integration properties.

The Gumbel distribution is selected because it closely approximates the Normal distribution and because it (unlike the Normal distribution) produces a model for which the probability of choice can be calculated without using numerical integration and simulation. This advantage leads to the general expression for the probability of choosing an alternative i from a set of J alternatives:

$$\Pr(i) = \frac{\exp(V_i)}{\sum_{j=1}^{J} \exp(V_j)} \tag{5}$$

where V_i denotes the representative utility for alternative i.

3.2.1 Maximum likelihood estimation

The concept of maximum likelihood is that the analyst wants to estimate the parameters that maximize the probability that all the individuals n in the sample would choose the alternative that they have actually chosen. These optimal parameters should then lead to the best prediction. The likelihood function is defined as this probability:

$$l = \prod_{\forall i} \prod_{\forall n} \Pr_{ni}{}^{\delta_{ni}} \tag{6}$$

where

$$\delta_{ni} = \begin{cases} 1, & \text{if } i \text{ is chosen by individual } n \\ 0, & \text{otherwise} \end{cases}$$

$\Pr_{ni}$ = probability that individual n chooses alternative i

Because the log of a function yields the same maxima and often is more convenient to differentiate, we maximize the log likelihood function instead:

$$L = \ln(l) = \sum_{\forall i} \sum_{\forall n} \delta_{ni} \ln P_{ni} \tag{7}$$

3.2.2 MNL example

In order to understand how the multinomial logit model is applied in this air travel dataset, and to introduce some of the expressions used later in the results, we can look at an example, in which the choice set consists of only two itineraries: one with airline 1 and one with airline 2. For simplicity reasons, let the only attributes of the two itineraries be fare (f), time (t), and number of connections (c). Figure 3 shows the two itineraries that the respondents can choose between.

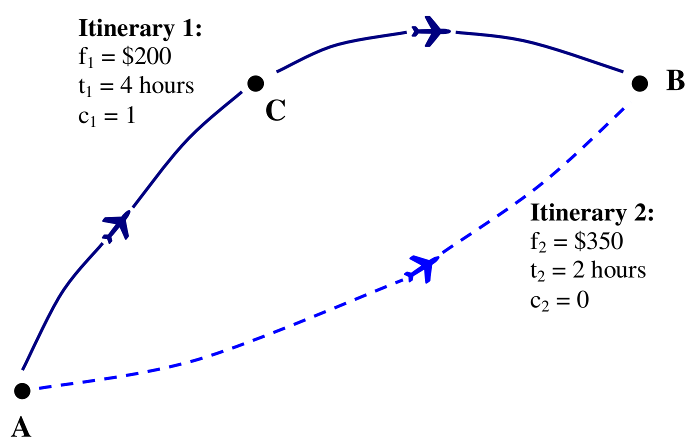

Figure 3. In this hypothetical situation, the decision maker has two itineraries to choose from, in order to get from A to B. The only decision variables are airfare, flight time, and number of connections.

The utility V_i of each itinerary i is given by:

$$V_i = \beta_{fare} \cdot f_i + \beta_{time} \cdot t_i + \beta_{connection} \cdot c_i \tag{8}$$

The goal is to estimate the β's from a dataset that includes the respondents' actual choices and the attributes of the alternatives in the choice set. A negative β implies that the attribute is a disutility to the respondent, while a positive β implies that the respondents prefer more of this attribute. In this simple case, all three parameters (fare, price, and connections) are expected to be negative. Let us assume that we have estimated the β's to be:

$$\begin{aligned}
\beta_{fare} &= -0.01\ \$^{-1} \\
\beta_{time} &= -0.60\ \text{hr}^{-1} \\
\beta_{connection} &= -0.45\ \text{conn.}^{-1}
\end{aligned}$$

(8) then becomes:

$$V_i = -0.01 \cdot f_i - 0.6 \cdot t_i - 0.45 \cdot c_i \tag{9}$$

and the utilities of each of the alternatives are:

$$\begin{aligned}
V_1 &= -0.01\,\$^{-1} \cdot \$200 - 0.60\,\text{hr}^{-1} \cdot 4\text{ h} - 0.45\,\text{conn.}^{-1} \cdot 1\,\text{conn.} = -4.85 \\
V_2 &= -0.01\,\$^{-1} \cdot \$350 - 0.60\,\text{hr}^{-1} \cdot 2\text{ h} - 0.45\,\text{conn.}^{-1} \cdot 0\,\text{conn.} = -4.70
\end{aligned}$$

It is clear that itinerary 2 has a slightly higher utility than itinerary 1, so we will expect the probability of choosing itinerary 2 to be higher than that of itinerary 1. This probability can be calculated by using (5):

$$\Pr(2) = \frac{\exp(V_2)}{\exp(V_1) + \exp(V_2)} = \frac{\exp(-4.70)}{\exp(-4.85) + \exp(-4.70)} = 0.537$$

Thus, there is a 53.7% probability of choosing itinerary 2, or in other words, out of 100 passengers, around 54 are expected to choose itinerary 2. We can conclude, that the savings in time and connections of itinerary 2 makes up for the extra cost. In effect, we can calculate the actual willingness-to-pay rates for the attributes by dividing the β-coefficients of the service attributes by the β-coefficient of the fare variable:

$$\begin{aligned}
WTP_{time} &= \frac{\beta_{time}}{\beta_{fare}} = \frac{-0.60\ \text{hr}^{-1}}{-0.01\ \$^{-1}} = \$60 \text{ per hr.} \\
WTP_{conn.} &= \frac{\beta_{conn.}}{\beta_{fare}} = \frac{-0.45\ \text{conn.}^{-1}}{-0.01\ \$^{-1}} = \$45 \text{ per connection}
\end{aligned}$$

The *WTP* values demonstrate that the average decision maker feels that one hour of saved flying time is worth $60, and that one saved connection is worth $45.

3.2.3 *t*-statistics

β estimates are always associated with uncertainty, a sample error, because they are estimated from a sample of the total population. Therefore, a *t*-statistic is provided with the estimation results. A *t*-statistic is the test statistic of the null-hypothesis $H_0 : \beta = \beta_0 = 0$, that the β estimate in question is equal to zero, meaning that the corresponding attribute does not have any impact on itinerary choice:

$$t\text{ - statistic} = \frac{\beta - \beta_0}{SD_\beta} = \frac{\beta}{SD_\beta} \tag{10}$$

The *t*-statistic follows the Student-t distribution, and for samples larger than 150 records, the critical t-values of selected confidence levels are shown in Table 1:

Table 1. Selected critical t-values and confidence levels for large sample sizes (>150)

Critical *t*-value (two-tailed test)	1.00	1.30	1.65	1.96	2.58
Confidence level	68%	81%	90%	95%	99%

The selection of critical *t*-values depends on how certain the analyst wants to be about the impact of the β-estimates. Usually, a *t*-value below 1.65 is considered insignificant (confidence level < 90%), but if the analyst has strong reasons to believe that a variable should be included in the model, it can be reasonable to keep variables with lower *t*-statistics.

In this report, *t*-statistics above 1.00 have been accepted in the models, *only* if they have the expected sign *and* are deemed important determinants in the itinerary choice process. This relatively small critical *t*-value has been allowed, because of the rather small sample size (119 passengers = 1190 records).

3.3 The mixed logit model

The mixed logit model generalizes the multinomial logit (MNL) model by allowing a mixing distribution over the multinomial logit kernel. Each individual has a specific model formulation, in which the β-parameters vary by taking a distribution around a mean value (Sørensen, 2003). In other words, the mixed logit model recognizes the fact that people have different preferences for reasons that are not necessarily possible to capture by their demographic differences, and must therefore be analyzed as random terms.

For repeated choice data from the same respondent, as in the current stated choice experiment, the mixed logit structure takes the form shown below:

$$P_n(\theta) = \int_{-\infty}^{\infty} L_n(\beta) f(\beta \mid \theta) d(\beta) \tag{11}$$

where $$L_n(\beta) = \prod_i \prod_t \left(\frac{e^{\beta' x_{nit}}}{\sum_j e^{\beta' x_{njt}}} \right)^{\delta_{nit}} \qquad (12)$$

and:

P_n = the unconditional probability of the sequence of observed choices of the individual over the alternatives i (i = 1 or 2 for the data in this report) and choice occasions t (t = 1,2,...,10 for the data in this report)

x_{nit} = a vector of variables specific to individual n, alternative i, and choice occasion t

β = parameters which are random realizations from a density function $f(.)$,

$$\delta_{nit} = \begin{cases} 1, & \text{if } i \text{ is chosen by individual } n \text{ on choice occasion } t \\ 0, & \text{otherwise} \end{cases}$$

θ = vector of underlying moment parameters characterizing $f(.)$.

3.3.1 Distribution of $f(.)$

Choosing an appropriate distribution for $f(.)$ can impact the simulation speed greatly and will also have consequences for interpretation of the results. It is a subject that is widely discussed in the literature on discrete choice models. Most analysts specify $f(.)$ with the Normal or the Lognormal distribution, but also triangular and uniform distributions have been used by some researchers (Train, 2003).

The major reason for using a bounded distribution such as the lognormal is that it prevents the possibility of unexpected parameter signs, since the distribution only contains positive values. This is particularly an issue in estimations of the value of time. A time coefficient in a transportation choice model is usually expected to have a negative sign (since travel time poses a disutility to decision makers), but if the distribution allows this coefficient to become positive, for some respondents, the model is suggesting that these respondents *favor* more travel time. Therefore, the lognormal distribution is preferred by some researchers.

However, the lognormal distribution has some disadvantages. Its long right tail can lead to some unrealistically high standard deviations. Additionally, some researchers have found that during simulation, the lognormal distribution has problems with slow convergence or any convergence at all (Hess *et al.*, 2004; Algers *et al.*, 1998). Furthermore, all distributions with a fixed bound at zero constrain the coefficients to be a particular sign. Hence, it does not allow the model to produce unexpected coefficient signs, even if this would result in a better model fit. For example, a model result with a positive coefficient for travel time may be an indication that the dataset is infected or that the model is misspecified. These errors may not be captured, if the coefficients are lognormally distributed.

For the estimations in this report, a Normal mixing distribution for $f(.)$ is considered. This has been done because of several reasons:

- Preliminary simulation runs of the mixed logit model in GAUSS failed to converge, when using the lognormal distribution, but worked well with the Normal distribution.
- It was preferred to compare the trade-off values, including standard deviations, with the results obtained from the Adler *et al.* (2005) study, because similar datasets were used in the two studies. Adler *et al.* (2005) used Normally distributed parameters, and they obtained unexpected results with the lognormal distribution.
- As noted above, the lognormal distribution allows the possibility of extreme standard deviations, and it was judged that the symmetrical Normal distribution would give a better, tighter representation of the randomness in passengers' choice parameters.

One of the disadvantages of using the Normal distribution, is that calculating values of time becomes more complicated. The ratio of two Normally distributed parameters is Cauchy distributed, which has an undefined mean and variance (Nielsen and Jovicic, 2003). This problem does not arise when using lognormal distributions, as the ratio then becomes lognormally distributed as well. In this project, the fare coefficient was fixed for reasons discussed in chapter 6, which eliminated the ratio problem.

3.3.2 Simulation

The mixed logit model requires the evaluation of analytically intractable multidimensional integrals in the classical estimation approach. The approximation of these integrals is undertaken using simulation techniques that entail the evaluation of the integrand at a number of draws taken from the multivariate Normal distribution and computing the average of the resulting integrand values across the different draws. Rather than taking random draws from the multivariate Normal distribution, Bhat (2001) proposed and introduced, in 1999, a simulation approach using Quasi-Monte Carlo (QMC) sequences for estimating discrete choice models with analytically intractable likelihood functions. Bhat demonstrated that the estimation of the mixed logit model using Halton draws (a type of QMC sequence) is much more efficient than using the usual random draws - computational time decreased by as much as 90%. Sivakumar *et al.* (2005), most recently, compared a whole range of different QMC sequences in terms of their ability to recover the underlying parameters of the mixing distribution accurately and precisely, and found that a certain kind of scrambled Faure sequence performs best. However, generating the Faure sequence is relatively tedious, and so the Halton sequence is used in this project. Halton sequences have the advantage that they give a very uniform coverage.

All estimations are undertaken using LIMDEP and the GAUSS matrix programming language. Gradients of the likelihood functions with respect to the parameters were coded for the optimization procedure.

4 DATA

4.1 Data source

The sample used in this report is drawn from a 2001 online survey of 621 air travelers, of which 119 were business travelers, conducted by the company Resource Systems Group (RSG, 2001). Respondents were selected from an online consumer panel and screened to include only those individuals who had made a recent paid domestic U.S. air trip. They were compensated for participating in the 30-45 minute web survey and the resulting response rate was just over 60%.

The respondents were asked to provide a wide range of characteristics of their latest US domestic flight, as well as basic demographic data about themselves. After obtaining the information, an internally coded heuristic in the survey software generated ten hypothetical sets of alternative itineraries based on an experimental design and presented each alternative itinerary along with the revealed choice itinerary in a series of 10 binary choice exercises to the respondent (see Figure 4).

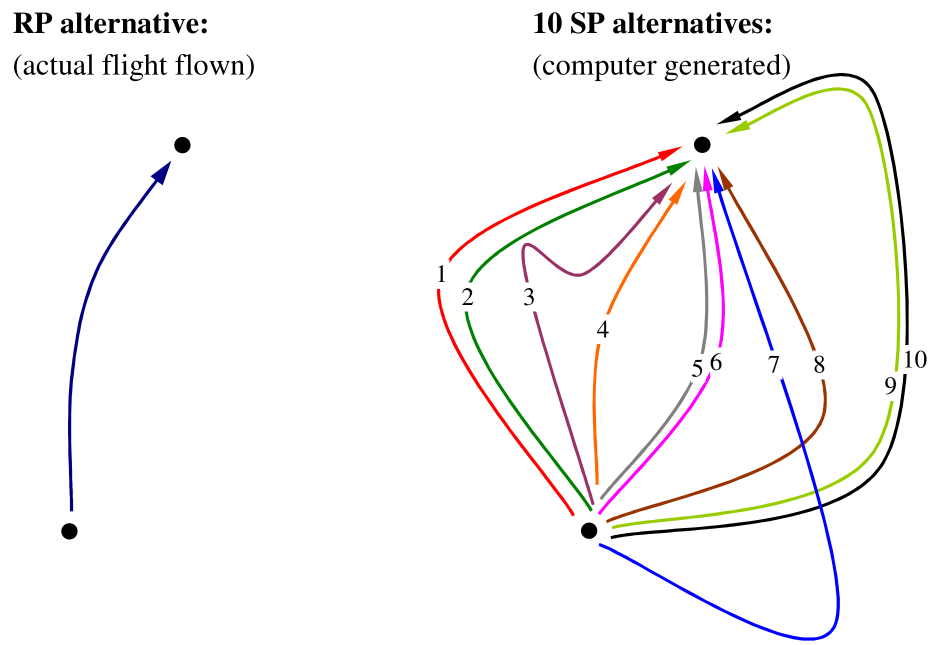

Figure 4. An illustration of the ten generated SP itineraries, resulting in ten different binary choices between each SP alternative and the RP alternative. Some of the SP alternatives may follow the same route as the RP alternative, but have different service amenities, thus distinguishing them from the RP alternative.

The generated alternative itineraries included the same origin and destination city as in the RP alternative, but with varying departure airports, fares, number of connections etc. The respondent had the choice of choosing his/her revealed choice itinerary or the alternative (SP) itinerary in each exercise. The precise definition of, and possible levels for, each attribute in the stated preference experiment is presented in Table 2.

Table 2. Definition of attributes characterizing an itinerary and possible levels of each attribute		
Attribute	**Definition**	**Possible levels of attribute**
Fare	Cost of itinerary (return fare)	Varied based on design around the price paid by respondent for the RP itinerary
Flight time	Total departure to arrival gate times	Varied based on design, while maintaining realistic flight times
On-time performance	Percentage of times the flight itinerary is on time	Varied in 10% intervals between 50% and 90%
Access time	Time of travel from respondent supplied trip end location and airport location at departure end	--
Connections	The number of connections in traveling from the origin airport to the destination airport	Non-stop, one stop, and two stops
Schedule time difference	Time representing difference between itinerary's arrival time and desired arrival time	Varied based on design
Aircraft type	Equipment used in itinerary	Propeller, regional jet, standard jet (single aisle), widebody jet (double aisle), and additional 12 combinations of these for "connecting flight" itineraries
Airline	Airline carrier in itinerary	The 28 most common US domestic airlines operating scheduled commercial service in 2001
Airport	Departure airport	All airports deemed "reasonable" from airport database and respondent-identified list of airports

Hence, there is no universal choice set, because each traveler has different flight itineraries to choose from, so each individual has a different choice set available to him/her. A similar situation could be a model of the choice of which shopping center to shop at. If observations are taken in many different cities, each individual will have access to different shopping centers, thus a different choice set. One could argue that the universal alternatives among the respondents are the revealed preference (RP) alternative, i.e. the flight that was actually flown, and the stated preference (SP) alternative, i.e. the computer generated itinerary. However, in each choice, these two alternatives represent different itineraries, so they do not result in a universal choice set.

Essentially, there are 10 choices per individual, i.e. 6210 choices made in total. Such data, in which numerous choices are made by each decision maker, is called *panel data* (Train, 2003). Each choice situation becomes a separate observation. The mixed logit specification in (11) treats the coefficients as varying over respondents but being constant over the 10 choice situations for each respondent. There is no variation within persons, because it is reasonable to assume that each respondent's preferences do not change over the course of answering the 10 SP questions. The multinomial logit (MNL) model, however, does not allow the coefficients to vary across respondents, so it assumes that everyone has the same preferences, aside from the differences that can be *observed* directly from demographic variation.

Table 3 shows a fragment of the air travel data - the first six variables (out of around 160). The hypothetical flight fares were generated such that they varied around the fare of the flown itinerary (RSG, 2001). The column "Choice" denotes whether the RP alternative (1) or the SP alternative (2) was chosen in the different SP exercises.

Table 3. Section of the survey results datasheet									
ID	**Choice**	**Fare**		**Connections**		**Airline**		**Airport**	
		RP alt.	*SP alt.*	*RP alt.*	*SP alt.*	*RP alt.*	*SP alt.*	*RP alt.*	*SP alt.*
1	1	$230	$115	0	0	American	Continental	IAD	BWI
1	1	$230	$345	0	2	American	Southwest	IAD	BWI
1	2	$230	$172	1	2	American	Jetblue	IAD	DCA
1	2	$230	$115	0	2	American	Continental	IAD	DCA
1	1	$230	$172	0	1	American	American	IAD	IAD
1	1	$230	$345	0	1	American	Airtran	IAD	BWI
1	1	$230	$115	0	1	American	United	IAD	IAD
1	2	$230	$115	0	1	American	Delta	IAD	BWI
1	1	$230	$345	0	1	American	Delta	IAD	DCA
1	2	$230	$115	0	0	American	Southwest	IAD	BWI
2	1	$250	$125	0	0	US Airways	United	ORD	MDW
2	1	$250	$375	0	1	US Airways	US Airways	ORD	ORD
2	2	$250	$375	0	0	US Airways	American	ORD	MDW

4.2 Survey description and sample statistics

This section describes the most important data characteristics relevant to the current report. For a detailed list of the survey results, see appendix A.

4.2.1 Market shares

As mentioned, there is no universal choice set, but rather generated choice sets for each respondent based on his or her RP alternative. In 70% of the stated choice questions, the respondents chose their RP alternative. This indicates the presence of inertia, which will be discussed in section 5.2.10. However, when the airline in the SP alternative was the same as in the RP alternative, respondents remained with their RP alternative in only 62% of the cases. This suggests that airline loyalty plays a role in itinerary choice.

4.2.2 Airline carriers and preferences

The respondents were asked to rank airlines in order of preference, assuming equal prices, by stating the first, second, third, and least favorite airline, from a list of the 28 most common US airlines at the time of the survey. These include:

- AirTran
- America West
- America West Express
- American Airlines
- American Eagle
- Alaska Airlines
- Continental
- Continental Express
- Delta
- Delta Connection
- Delta Express
- Delta Shuttle
- Frontier
- Jet Blue
- Metrojet
- Midway Airlines
- Midwest Express
- Northwest
- Northwest Airlink
- Reno Air
- Southwest
- TWA
- United
- United Express
- United Shuttle
- US Airways
- US Airways Express
- US Airways Shuttle

Three dummy variables were created: each indicate whether an itinerary is flown with the respondent's first preferred airline, second preferred airline, or third preferred airline, respectively (For connecting flights, it is assumed that all flight legs are flown with the same airline. The least preferred airline is chosen as the base. In the SP questions, the computer generates alternative itineraries with one of the four ranked airlines only. This means that the ranking of the airlines of the alternative itineraries is always known. However, the ranking of the airline in the RP alternative may not always be known. For example, if the respondent flew on an airline, that was not the first, second, third, or least preferred, we cannot know the ranking of that airline, and by the definition of the dummy variables, it will be treated as the base (the least preferred airline), even though it might be the 8th preferred, for example. This situation is prevalent for 101 respondents (16% of the sample), so it does create some source of error. Essentially, it means that airlines ranked 4 through 28 are all considered ranked as number 28, the *least preferred*. A justification for this is that there may be ambiguity when ranking a large amount of airlines, so only the first couple of rankings can be considered credible. Many passengers may have an idea about what their favorite airport is, and maybe their second and third favorite airport, but after that, it could become unclear.

The respondents' ranking of airlines based on their perception of the quality and service includes the most dominant carriers in the market. In particular, Delta received the highest ranking by 20% of the respondents, followed by Southwest and United (both 11%) and American (10%). The airlines chosen in the RP alternative are consistent with the most popular ranked airlines. Hence, the most frequently chosen airlines were Delta (18%), Southwest (15%), United (12%), and Northwest (10%). While the relatively small sample size means that these are not completely reflective of the actual pre-9/11 business market shares, the sample includes a sufficient representation to support modeling of itinerary choices across the major carriers.

4.2.3 Air fares

The respondents were asked to provide the return fare of their flight. Six respondents only made a one-way trip, and therefore their fares have been doubled in order to make them comparable with the round trip fares. Figure 5 shows the distribution of fares of the actual flown (RP) itineraries. Most fares range between $100 and $500, while the last 10% lie between $500 and $2500. For each RP fare, ten hypothetical SP fares are generated randomly around the RP fare, such that fares above and below the RP fare are presented to the decision maker, as can be seen in the close-up in Figure 5.

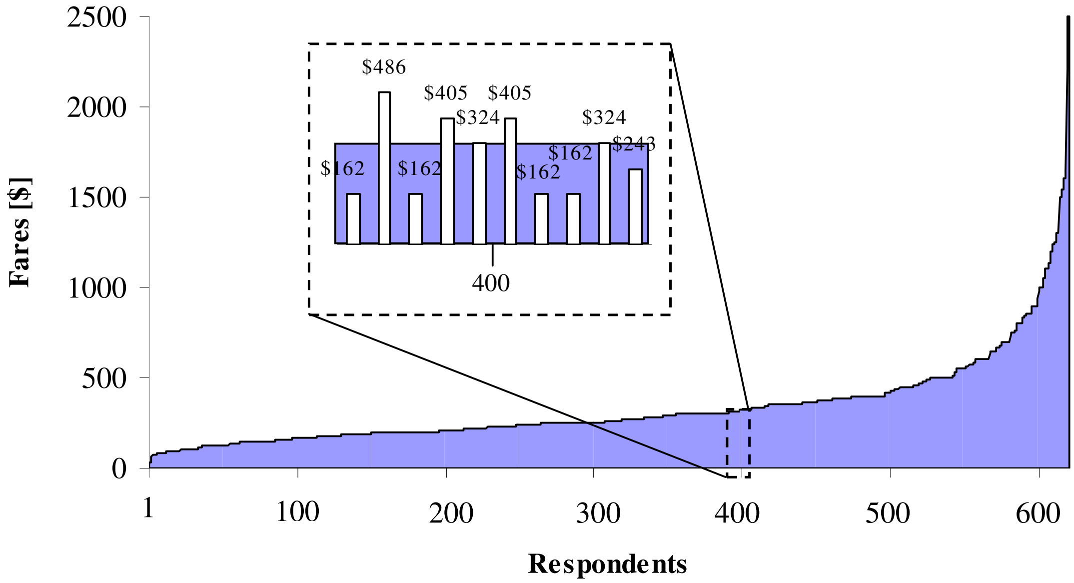

Figure 5. Distribution of return fares among the 621 sorted respondents. The close-up shows respondent no. 400, whose RP fare was $324. The generated SP fares (white columns) vary in range from $162 to $486.

If the fares are normalized by the duration of the flight to control for the different lengths of the markets served by different carriers, we can compare the different fares by carrier. Delta has the highest reported average fare of $3.44 per minute, followed by Northwest ($3.31 per minute) and American ($3.19 per minute). Continental has the lowest reported average fare ($2.02 per minute) among the six most chosen airlines, followed by Southwest ($2.27 per minute). These results indicate that fares are not the sole criterion in the choice or ranking of air carriers.

4.2.4 Departure airports

Similar to the airline rankings, the respondents were asked to rank their first, second, third, and fourth preferred home airports from a list of the closest commercial airports to their origin. Three dummy variables were created, indicating whether the passenger is flying from his/her first, second, or third preferred airport, thus making the fourth preferred airport the base. Again, as in the airline case, there are examples where the origin airport in the RP alternative is not among the four ranked airports in the actual flown itineraries.

However, this is only the case for four non-business respondents and no business respondents, so we can disregard this fact. The reason for the small number of such cases is that the airport rankings were chosen among the airports that were closest to the respondent's point of origin, so there is a very high possibility that one of these four ranked airports were used in the RP alternative. In contrast, the choice set is greater in the case of airlines, especially in large airports, so choosing an airline which is not among the four ranked airlines is more likely.

4.2.5 Access times

The respondents were asked how much time it took to get to their airport of origin. Additionally, access times to the ranked airports were included in the survey, so that access times to the airports in the SP questions (which, as mentioned before, only included ranked airports) would be available. The responses were in time intervals, which in this report have been recoded into discrete values, as shown in Table 4 and Table 5:

Table 4. Access time to airport of origin (RP alternative)

Category	Recoded to
< 15 min.	10 min.
15 – 30 min.	22.5 min.
30 – 45 min.	37.5 min.
45 – 60 min.	52.5 min.
1 – 1.5 hours	75 min.
1.5 – 2 hours	105 min.
2 – 3 hours	150 min.
> 3 hours	200 min.

Table 5: Access time to ranked airports (used in the SP alternatives)

Category	Recoded to
< 30 min.	20 min.
30 – 60 min.	45 min.
1 – 1.5 hour	75 min.
1.5 – 2 hours	105 min.
2 – 3 hours	150 min.
> 3 hours	200 min.

4.2.6 On time performance

The SP exercises stated the on time performance for the generated itineraries in percentages ranging from 50% to 90% in 10 percent increments. An on time performance of 50% would mean that the plane in the itinerary is expected to arrive on time on half of its flights. Additionally, the respondents were asked if their actual RP flight was on time, which they could answer with a yes or no. This question was recoded as percentages with yes being 100% and no being 0%. The reason for doing this is that one can interpret the percentages as probabilities of being on time. Since the respondents already know for sure whether they were on time with their current flight, these probabilities will be 100% and 0%, respectively. As seen in Appendix A, 80% of the flights in the RP alternatives were on time.

4.2.7 Aircraft type

The respondents were asked which type of aircraft they traveled with on their "primary" portion of their trip, meaning the longest flight leg. The choice alternatives were:

1) Propeller
2) Regional jet (50-100 seats)

3) Standard jet (100-200 seats)
4) Widebody jet (more than 200 seats)

Contrary to the RP alternatives, the computer-generated SP alternatives provided information about both aircraft types for connecting flights. There are no categories with three different aircraft types. This means that for double-connecting flights, the three planes only included at most two different aircraft types. Table 6 shows the distribution of the different combinations of aircraft types in the SP alternatives. The order of the aircraft is significant: the first aircraft type mentioned is the aircraft on the primary (longest) leg of the trip. Table 7 shows the similar distribution for the RP alternatives.

Table 6. Aircraft types in SP alternatives

1. Propeller	5.6%
2. Regional jet	10.3%
3. Standard jet	16.6%
4. Widebody jet	12.3%
5. Propeller and propeller	4.9%
6. Propeller and regional jet	2.0%
7. Propeller and standard jet	2.5%
8. Regional jet and propeller	6.2%
9. Regional jet and regional jet	3.9%
10. Regional jet and standard jet	3.9%
11. Standard jet and propeller	8.8%
12. Standard jet and regional jet	6.3%
13. Standard jet and standard jet	5.1%
14. Widebody jet and propeller	3.5%
15. Widebody jet and regional jet	4.4%
16. Widebody jet and standard jet	3.7%

Table 7. Aircraft types in RP alternatives

1. Propeller	2.5%
2. Regional jet	14.3%
3. Standard jet	66.4%
4. Widebody jet	16.8%

It is evident, that the SP alternatives are generated such that widebody jets are only used on primary legs, not secondary legs. This is reasonable, since widebody jets are mostly used for long-range routes. It seems that there is an underrepresentation of standard jets, and thereby and overrepresentation of the other types of aircraft in the SP alternatives compared to the RP alternatives. Thus, one could argue that the SP alternatives are less realistic. However, this misbalance has the advantage that the researcher can model passengers' responses to different aircraft types more precisely, when there is greater variation.

4.2.8 Preferred arrival time

When making the choice of itinerary, the respondent is assumed to include the preference for a certain arrival time as a factor. Therefore, the survey asked the respondent at what time he/she would prefer to arrive at the destination, and for each SP alternative, the arrival time was provided. The arrival times were recoded into *minutes after midnight* (MAM), so that they could be subtracted from each other. For this report, a new variable was created denoting the difference between preferred arrival time and arrival time offered in the itinerary choice. However, the flights that arrived around midnight would result in wrong time differences. For example, if the preferred arrival time of a certain respondent is 11 p.m. (=1380 MAM), and one of the choice itineraries arrives at 1.30 a.m. (=90 MAM), the actual time difference would be two and a half hours (=150 MAM). However, by simply subtracting the two, without controlling for the date change, we would get a time difference of -1230 MAM, indicating that the flight is 21.5 hours ahead of the preferred arrival time. To avoid these errors, two additional questions had been included in the survey:

- Which day was assumed the preferred day (compared to current itinerary's arrival time)?
- Which day is the alternative itinerary's arrival time on (compared to current itinerary's arrival time)?

Table 8 shows the proportion of arrivals for the whole dataset, 621 respondents. There are 12,420 arrivals in total (621 respondents * (10 SP itineraries + 10 RP for each SP itinerary)). Naturally, the majority of the current arrivals are on the preferred day of arrival. The ones that are not, are mostly scattered around the midnight hour, thus creating a date change. Note that no choice itineraries arrive the day before the current flight arrival, if the preferred day is the day after the current flight arrival. In other words, no flights arrive two days before the preferred day. Likewise, no flights arrive two days after the preferred day.

Table 8. Arrival times on different days		**Preferred day compared to current flight arrival**		
		Same day	Day after	Day before
Flight arrival compared to current flight	Same day	11942	176	197
	Day after	61	4	**0**
	Day before	17	**0**	23

Most business passengers in the sample (35%) prefer arriving at their destination airport in the time interval between 8 a.m. to noon, followed by the time intervals of noon to 4 p.m. (26%) and 4 p.m. to 8 p.m. (19%). 16% prefer arriving before 8 a.m. and only 3% prefer arriving after 8 p.m. These results are reasonable, considering that most business meetings take place in the morning or afternoon. For the non-business travelers, the preferred arrival times are more evenly spread during the day.

4.2.9 Frequent flyer membership

Information about the respondents' frequent flyer program (FFP) membership in their first, second, third, and least preferred airlines was asked for in the survey. For this report, the FFP information was converted into three different dummy variables: standard, silver, or gold membership, which each take a value of one, if the respondent is an FFP member of a certain level on the airlines offered in the choice set. As mentioned in section 4.2.2, the computer generates alternatives with the ranked airlines only. Thus, it will not suggest an airline, for which FFP member information is not given. For example, if it suggested the 8th preferred airline, there would not be information about the respondent's FFP member information on that particular airline. However, as in the airline case, the RP itinerary that the passenger flew, may be with an airline that is not ranked, and thus, there is no FFP information about this airline. For example, the passenger might have flown with his/her 13th preferred airline, and we cannot know whether he/she is an FFP member of this airline. In that case, the model will assume that the respondent is not an FFP member on any of the three levels. This could be a source of error.

Another issue is that most airlines that are members of alliances count miles flown on their partners as miles in their own FFP accounts. For example, a Continental FFP member earns miles on his/her account when flying Delta and Northwest, because they are all members of the Skyteam alliance. Since the respondent is only asked whether he/she is a member of the FFP of the particular airline in the choice alternative and not its alliance airlines, there can be instances, in which the respondent is not a member of the chosen airline's FFP, but is still benefiting from frequent flyer miles, by being a member of a partner airline's FFP. This issue will likely reduce the significance of the impact of FFP membership in a discrete choice model. It is difficult to control for this, because we do not have full information about all the FFP's that the respondents are members of.

4.2.10 Income

One of the demographic variables is annual household income. The respondents were to pick an interval instead of providing a continuous value. Many surveys do this, in order to make the question less sensitive, and thus increase response rate (Richardson et al., 1995). This is a disadvantage to the researcher, because essentially, it would not be appropriate to treat it as continuous data. The income distributions among passengers are shown in Figure 6.

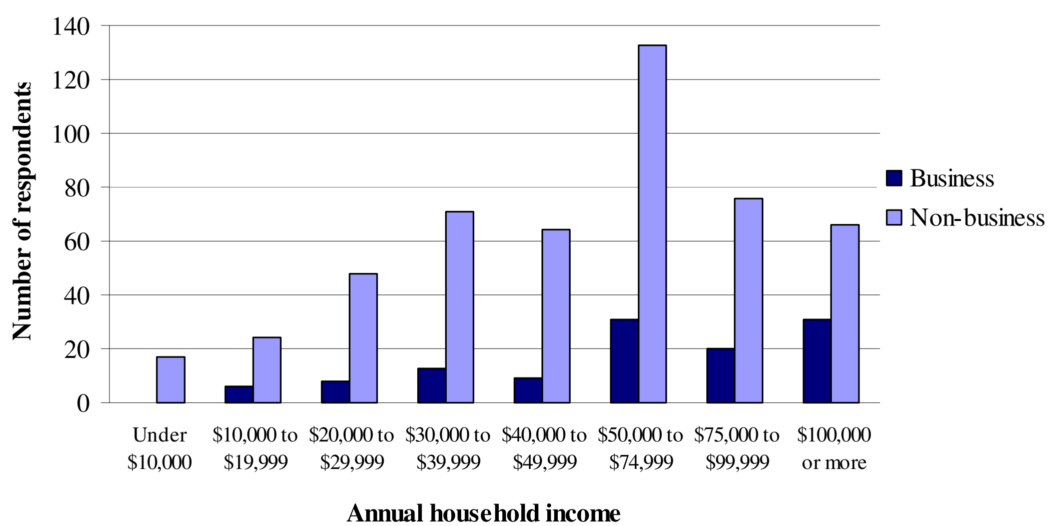

Figure 6. Distribution of annual household income among the business and non-business passengers

Figure 6 shows that there is a very high share of respondents in the highest category ($100,000 or more). This suggests that the survey designers should have added one or two more categories, to obtain a more precise depiction of the income distribution.

In order to be able to incorporate income as an interaction variable in the discrete choice models, the intervals have been recoded into discrete values, by using the average value of each interval, see Table 9. Determining a value for the highest interval is more challenging, since, as mentioned before, there is a large share of people in this income category. It is set to $120,000.

Table 9. Recoding of the income level categories

Income range	**Recoded as**
$10,000 - $19,999	$15,000
$20,000 - $29,999	$25,000
$30,000 - $39,999	$35,000
$40,000 - $49,999	$45,000
$50,000 - $74,999	$62,500
$75,000 - $99,999	$87,500
> $100,000	$120,000

4.2.11 Employment sectors

Figure 7 shows what kind of industry types, the business passengers are employed in. Employment information will not be used in the non-business model.

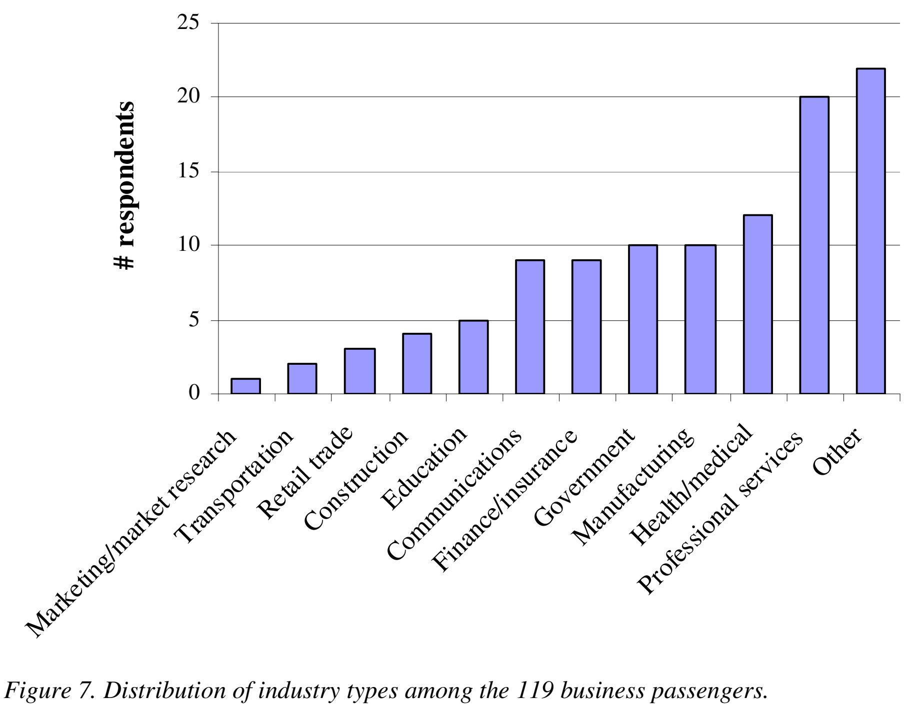

Figure 7. Distribution of industry types among the 119 business passengers.

5 MULTINOMIAL LOGIT RESULTS

This section describes the results from the MNL model, including all level-of-service variables that are considered to potentially affect itinerary choices: Fare, flight time, on time performance, access time, connections, schedule time difference, aircraft, airline, departing airport, and frequent flyer membership. In addition, an "inertia" variable was considered to incorporate an overall reluctance to shift from the revealed choice itinerary.

First, the appropriateness of distinguishing business passengers from non-business passengers, i.e. applying market segmentation in the estimation models is discussed, and next, the MNL models for each segment are presented, separately.

5.1 Market segmentation between business and non-business

Many factors suggest that market segmentation between business and non-business travelers is a good idea. First, an intuitive *a priori* assumption would be that business travelers value time more and price less than non-business travelers. Secondly, almost all the existing literature on air travel modeling proposes market segmentation, as noted in the literature review, chapter 2. In this section, it will be statistically tested and proved that the dataset should be divided into two segments: business and non-business.

The full dataset includes information about five trip purposes (business, vacation, visiting friends/relatives, attending school/college, and other). Figure 8 shows the shares between the five different trip purposes that the respondents could choose for their actual flight.

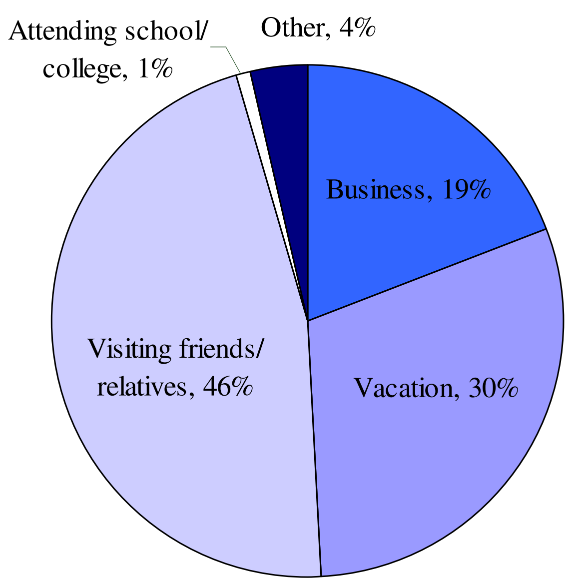

Figure 8. The percentage shares between the five different trip purposes among all 621 respondents.

5.1.1 Weight factors

The dataset does not represent a random sample of domestic US flight passengers, and there is an underrepresentation of business travelers, which only accounted for 19% of the sample, as opposed to the actual share of 43% of American flight passengers, according to the US Department of Transportation (DOT) American Travel Survey (RSG, 2001). Therefore, a weighting variable has been included to account for different ratios of trip purposes. Table 10 shows the shares in the dataset, the preferred shares, and the calculated weight factors. In the US DOT survey, the purposes *Vacation, Attend school/college* and *Other* were grouped as one.

Table 10. Shares and weight factors of different trip purposes in the full dataset

Trip purpose	Number of observations	Shares in sample (Total = 621 obs.)	Shares according to US DOT	Weight factor
Business	119	19.16 %	42.92 %	2.24
Visiting friends or relatives	287	46.22 %	26.99 %	0.584
Vacation, attend school/college, and other	215	34.62 %	30.02 %	0.867

5.1.2 Pooled model

The weight factors are used in the following standard logit model with the most important (from an *a priori* standpoint) level-of-service (LOS) variables: Fare, time, on time performance, number of connections, and access time. These five variables are expected to be the most influential in the decision process. Table 11 shows the estimated coefficients and t-statistics for the full dataset, using the weight factors above. An alternative-specific constant for the RP alternative has also been included. This constant captures the inertia effect, which will be discussed later in this chapter.

Table 11. Basic logit model for the full dataset with key LOS variables

Variables	β	*t*-stat
Constant (RP itinerary)	0.529	*11.13*
Fare ($100)	-0.013	*-30.76*
Flight time (in 100s of min.)	-0.008	*-14.77*
On time performance (%)	0.010	*10.26*
Number of connections	-0.505	*-10.24*
Access time to airport (min.)	-0.012	*-18.77*
Number of observations		6210
Number of parameters - K		6
Log-likelihood at sample shares $L(c)^{\dagger}$		-3793.47
Log-likelihood at convergence - $L(\hat{\beta})$		-2206.85
Likelihood ratio index ($\bar{\rho}_c^2$)‡		0.417

† The log-likelihood value at sample shares corresponds to the case where each individual is assigned a 0.7 probability of staying with the revealed preference alternative for each of her/his choice occasions. This is based on the 70% choice of the revealed preference alternative in the stated choice experiments.

‡ $\bar{\rho}_c^2 = 1 - \frac{[L(\hat{\beta}) - (K-1)]}{L(c)}$

From looking at the very high t-statistics, it is clear that all the variables are highly significant. The signs of the coefficients are expected: Price, flight time, access time, and number of connections pose a disutility to the traveler, while a high on time performance improves the utility of an itinerary. It is interesting that the sensitivity to access time is higher than the sensitivity to flight time. This can be compared with urban travel demand models, where out-of-vehicle travel time often has a higher value than in-vehicle travel time (Koppelman et al., 1995).

5.1.3 Segmented models

The model in Table 11 can be called the *pooled model*, because it is a cluster of all the trip purpose segments. The next step is then to estimate similar *segmented models* for each of the five segments, and test whether it is reasonable to continue with the pooled model. Table 12 shows the estimated coefficients for the same key variables as in Table 11 for each of the five segments. Note that this time, the records in each segment have not been weighted since each model only consists of the same segment.

Table 12. Logit model with market segmentation on basis of five different trip purposes

Variables	Business		Vacation		Visiting friends/relatives		Attending school or college		Other	
	β	*t*-stat	β	*t*-stat	β	*t*-stat	β	*t*-stat	β	*t*-stat
Constant (RP itinerary)	0.342	*3.38*	0.695	*7.21*	0.656	*8.73*	0.923	*1.14*	0.750	*2.46*
Fare ($)	-0.006	*-10.67*	-0.018	*-26.38*	-0.019	*-22.29*	-0.024	*-3.02*	-0.015	*-5.95*
Flight time (min.)	-0.007	*-5.88*	-0.009	*-8.33*	-0.009	*-10.53*	-0.018	*-1.72*	-0.011	*-2.77*
On time performance (%)	0.010	*4.79*	0.012	*6.25*	0.011	*7.55*	0.014	*1.32*	0.012	*2.06*
Number of connections	-0.656	*-5.96*	-0.489	*-5.16*	-0.457	*-5.96*	0.519	*0.48*	-0.671	*-2.00*
Access time to airport (min.)	-0.013	*-9.17*	-0.013	*-9.47*	-0.013	*-12.87*	-0.023	*-1.94*	-0.015	*-3.39*
Number of observations		1190		1870		2870		50		230
Number of parameters - K		6		6		6		6		6
$L(c)$ at sample shares†		-726.93		-1142.32		-1753.18		-30.54		-140.50
$L(\hat{\beta})$ at convergence		-478.24		-600.66		-914.12		-15.56		-67.67
Likelihood ratio index ($\bar{\rho}_c^2$)‡		0.335		0.470		0.476		0.327		0.483

† The log-likelihood value at sample shares corresponds to the case where each individual is assigned a 0.7 probability of staying with the revealed preference alternative for each of her/his choice occasions. This is based on the 70% choice of the revealed preference alternative in the stated choice experiments.

‡ $\bar{\rho}_c^2 = 1 - \frac{[L(\hat{\beta}) - (K-1)]}{L(c)}$

5.1.4 Test between the pooled and segmented models

In order to statistically test if this segmentation is appropriate and has a better goodness-of-fit than the original pooled model, an extended likelihood ratio test can be performed (the following procedure is from Koppelman et al., 1995). The null-hypothesis is:

$$H_0 : \beta'_{\text{pool}} = \beta'_{\text{business}} = \beta'_{\text{vacation}} = \beta'_{\text{friends/relatives}} = \beta'_{\text{school}} = \beta'_{\text{other}}$$

where β'_{pool}, β'_{business}, β'_{vacation}, $\beta'_{\text{friends/relatives}}$, β'_{school}, and β'_{other} are the vectors of coefficients for the pooled model, the business segment, vacation segment, visiting friends segment, school segment, and other segment, respectively. A wording of the null-hypothesis may be that *the sensitivities to price, time, on time performance, connections,* *and airport access time does not vary between the business passengers, vacation, visiting, school, and other non-business passengers.* The test statistic is:

$$-2 \cdot \left(L_{pool} - \sum_{s=1}^{S} L_s \right) = -2 \cdot \left[-2206.85 - (-478.24 - 600.66 - 914.12 - 15.56 - 67.67) \right] = 261.20$$

where

L_{pool} is the log likelihood for the pooled model

L_s is the log likelihood for the model for segment s

S is the number of segments (in this case $S = 5$)

The test statistic should be compared with $\chi_{n,p}$, which is the critical chi-square value with n degrees of freedom at the (1-p) confidence level. n is equal to the total number of coefficients in the segmented models minus the number of coefficients in the pooled model, i.e. in this case n = 6·5-6 = 24. At the 99.5% confidence level, $\chi^2_{24,0.005}$ = 45.56. Since 250.90>45.56, we can reject the null-hypothesis, and thereby prove that a segmentation is desirable from a statistical standpoint.

5.1.5 Merging the segmented models

Having proved that the pooled model is imprecise in its representation of the coefficients, does not necessarily mean that it is favorable to continue with the five segmented models. From looking at the coefficients of these in Table 12, there is a high possibility that some of them should be merged.

The three largest segments, in terms of number of observations, are business, vacation, and visiting friends/relatives. It is clear that the coefficients of the business segment vary significantly from the other segments. The sensitivity to price, for example, (which is the most significant of all the variables) is three times as low for business travelers as for vacationing and visiting travelers. It is also remarkable that the vacation and visiting segments have almost equal coefficients, implying that these two segments could be merged.

The segment *Attending school or college* with its 50 observations (i.e. only five individuals in the dataset) is essentially too small to be estimated, and the values of the coefficients cannot be considered credible. However, all of them are closer to the coefficients of the segments *Vacation* and *Visiting friends/relatives* than to those of the business segment, especially the important fare variable. It would thus be advisable to unite the school segment with the two previous segments.

The segment *Other* includes passengers whose trip purposes are everything that would not fit with the other four reasons. Therefore, it is a very heterogeneous group and should most likely not be treated as one distinct segment. Although not entirely obvious, the coefficients imply that it would be most appropriate to combine it with the *Vacation* and *Visiting friends/relatives* segments. Only the variable *number of connections* has a coefficient that is closer to that of the business segment.

In conclusion, it would be worthwhile to test whether the five non-business segments can be merged into one segment. When estimating a logit model for this pooled model, the non-business observations have been weighted according to new weight factors that are calculated due to their proportion of the new subset. The old weight factors cannot be used since they assumed different shares in which also business passengers were included. The new weight factors are shown in Table 13:

Table 13. Weight factors for the merged purposes in the non-business segment

Trip purpose (non-business segment)	**Number of observations**	**Shares in sample (Total = 502 obs.)**	**Shares according to US DOT**	**New weight factor**
Visiting friends or relatives	287	57.17 %	47.34 %	0.828
Vacation, attend school/college, and other	215	42.83 %	52.66 %	1.230

Table 14 shows the coefficient estimates for this new pooled non-business model along with the before-shown coefficients of the business segment.

Table 14. Logit model with market segmentation on business an non-business travelers

Variables	**Business**		**Non-business**	
	β	*t*-stat	β	*t*-stat
Constant (current itinerary)	0.342	*3.38*	0.684	*11.94*
Fare ($)	-0.006	*-10.67*	-0.018	*-38.21*
Flight time (min.)	-0.007	*-5.88*	-0.009	*-13.79*
On time performance (%)	0.010	*4.79*	0.012	*10.11*
Number of connections	-0.656	*-5.96*	-0.480	*-8.24*
Access time to airport (min.)	-0.013	*-9.17*	-0.013	*-16.65*
Number of observations	1190		5020	
Number of parameters - K	6		6	
Log-likelihood at sample shares $L(c)$†	-726.93		-3066.54	
Log-likelihood at convergence - $L(\hat{\beta})$	-478.24		-1603.16	
Likelihood ratio index ($\bar{\rho}_c^2$)‡	0.335		0.476	

† The log-likelihood value at sample shares corresponds to the case where each individual is assigned a 0.7 probability of staying with the revealed preference alternative for each of her/his choice occasions. This is based on the 70% choice of the revealed preference alternative in the stated choice experiments.

‡ $\bar{\rho}_c^2 = 1 - \frac{[L(\hat{\beta}) - (K-1)]}{L(c)}$

We can test whether the assumption that all the non-business groups should be merged, is correct. In this case, the pooled model is the non-business segment in Table 14. The stated null-hypothesis is:

$$H_0 : \beta'_{\text{non-buiness}} = \beta'_{\text{vacation}} = \beta'_{\text{friends/relatives}} = \beta'_{\text{school}} = \beta'_{\text{other}}$$

worded as: *All non-business travelers have the same average sensitivities to the basic level-of-service attributes, when choosing a flight itinerary.* The test-statistic is:

$$-2 \cdot \left(L_{pool} - \sum_{s=1}^{S} L_s \right) = -2 \cdot \left[-1603.16 - (-600.66 - 914.12 - 15.56 - 67.67) \right] = 10.30 < \chi_{18,0.005} = 37.16$$

Since the test statistic is less than the critical value, we must accept the null-hypothesis. Hence, the before-made judgments about the equality of the coefficient values for the different non-business segments are valid, and it is therefore best to proceed with only two segments: business and non-business.

The coefficients in Table 14 show that there are significant differences between the two segments. Non-business passengers seem to be three times as sensitive to fares as business passengers, while both flight time and number of connections are more important determinants for business passengers. These are reasonable and expected results. It has now been shown that business and non-business passengers should be treated in separate models, and so the next two sections will discuss these segments separately.

5.2 Business results

As *interaction variables* (demographic/trip related characteristics that are multiplied with the LOS variables), the model includes gender, income (for the proportion of business travelers who were not reimbursed by their company), employment sector, number of vehicles owned by household, duration of trip (number of nights at destination), whether or not the traveler checked bags, frequency of air travel (number of US domestic flights in the last year), and day of travel.

The final variable specifications in the models were based on a systematic procedure of eliminating statistically insignificant variables, combined with intuitive considerations and informed by the results of earlier studies. In the specifications, alternative functional forms for continuous demographic/trip variables were tested, including dummy variable effects, logarithmic functional forms, and piecewise linear effects. The final specification was selected based on statistical fit and intuitive considerations.

Table 15 presents the results of the multinomial logit models for business travelers. The first model includes service characteristics, but no demographic/trip interactions, and the second includes service characteristics and their interactions with demographic/trip variables. The rest of this section discusses the impact of each service characteristic in turn in separate sections. As mentioned in section 3.2.3, only t-statistics above 1.00 have been kept in the final model (except the inertia variable), and only if they were deemed important determinants in the choice process.

Table 15. Multinomial logit (MNL) model results for business travelers				
Variables	**Service characteristics only**		**Demographic/trip related interactions**	
	β	*t*-stat	β	*t*-stat
Fare (100 $)	-0.559	*-10.43*	-0.658	*-8.59*
Female travelers			0.225	*1.97*
Traveling in a group (2 or more)			0.170	*1.50*
Fare/income for self-paying travelers (per $100 yearly salary)			-2.682	*-3.56*
Flight time (in 100s of min.)	-0.687	*-5.39*	-1.190	*-4.43*
Employment: Transportation or retail trade			-1.469	*-2.06*
Employment: Professional services			-0.893	*-1.95*
Employment: Other* (base)			0	*N/A*
Bags checked in			0.421	*1.40*
Frequent travelers (# yearly air trips)			0.043	*2.18*
On-time performance (%)	0.866	*4.13*	1.967	*4.29*
Bags checked in			-1.132	*-2.51*
Frequent travelers (# yearly flights)			-0.098	*-3.38*
Travel on a Friday or Saturday			1.561	*2.11*
Access time to airport (in 100s of min.)	-1.072	*-6.62*	-3.685	*-6.29*
Vehicle owners (# vehicles in household)			1.265	*5.17*
Travel on a Friday or Saturday			-1.353	*-2.84*
Connecting flight	-0.711	*-5.41*	-0.908	*-3.99*
Bags checked in			0.631	*2.11*
Travel on a Friday or Saturday			-1.104	*-2.28*
Schedule time difference (in 100s of min.)				
Early arrival relative to preferred arrival time	-0.204	*-1.03*	-0.588	*-2.35*
Late arrival relative to preferred arrival time	-0.311	*-2.26*	-0.606	*-3.09*
Duration of stay (# nights at destination)			0.048	*1.68*
Aircraft type: Standard jet	0.431	*3.34*	0.416	*2.34*
Frequent travelers (# yearly flights)			2.363	*1.39*
Airline and airport preferences				
Airline: First preferred	0.492	*2.77*	0.687	*3.41*
Airline: Second or third preferred	0.236	*1.49*	0.438	*2.46*
Airport: First preferred	0.382	*2.56*	0.405	*2.45*
Frequent flyer membership				
Standard or silver member	0.317	*1.67*	0.413	*1.97*
Gold member	0.926	*1.23*	1.001	*1.33*
Inertia (staying with current RP itinerary)	0.099	*0.81*	-1.234	*-3.55*
Purchase directly from airline			1.663	*3.94*
Purchase from agent or online			1.240	*3.47*
Other means of ticket purchase** (base)			0	*N/A*
Number of observations		1180		1180
Number of parameters - K		14		32
Log-likelihood at sample shares $L(c)$†		-720.80		-720.80
Log-likelihood at convergence - $L(\hat{\beta})$		-453.49		-392.00
Likelihood ratio index ($\bar{\rho}_c^2$)‡		0.353		0.413

*Includes communications, construction, education, finance/insurance, government, health/medical, manufacturing, marketing/market research, and wholesale trade.

**Mostly corporate travel offices.

† The log-likelihood value at sample shares corresponds to the case where each individual is assigned a 0.7 probability of staying with the revealed preference alternative for each of her/his choice occasions. This is based on the 70% choice of the revealed preference alternative in the stated choice experiments.

‡ $\bar{\rho}_c^2 = 1 - \frac{[L(\hat{\beta}) - (K-1)]}{L(c)}$

5.2.1 Airfare

The fare coefficient has the expected negative, and statistically significant, value in the model with service characteristics only. The interaction model indicates statistically significant variations in the fare sensitivities based on demographic/trip interactions. In particular, women and individuals traveling in a group are less sensitive to fares than men and individuals traveling alone, respectively (however, note that the fare coefficient continues to remain negative for women and individuals traveling in a group).

Many travel demand models weight the cost coefficient with income. This is a result of an assumption that price is valued differently depending on the travelers' income (Jara-Diaz, 1998). A high-income traveler will likely not be as sensitive to fares as a low-income traveler. This is also what this model shows: for business passengers who self-pay (as opposed to their company paying) for their trip, lower household income results in a higher sensitivity to fares. This is the expected income effect on consumption of goods and services. Not surprisingly, household income had no statistically significant interaction effect for business passengers whose travel is paid by their company.

5.2.2 Flight time

Flight time also has the expected negative effect on the utility of an itinerary. The interaction model indicates higher sensitivity to flight time for business travelers employed in transportation or retail trade, and in professional services, relative to those employed in communications, construction, education, finance/insurance, government, health/medical, manufacturing, marketing/market research, and wholesale trade. This is a rather interesting result, which may be a reflection of the differential value companies place on employee time based on the overall market value of the services offered. Further, the interaction model suggests that passengers, who check in bags, and frequent business travelers, are less time-sensitive than passengers who do not check in bags and who are infrequent business travelers, respectively. Passengers who check in bags may either be intrinsically more time-patient because of personality characteristics or may be traveling for less time-sensitive activities (after all, these passengers are prepared to spend additional time at the origin and destination airports to check in and retrieve their baggage). Frequent business travelers may have become accustomed to working productively on flights, rendering flight time less onerous.

5.2.3 On-time performance

The model without interactions shows the expected positive impact of on-time performance on itinerary utility. The model with interactions indicates that this positive impact is retained for all travelers, though the effect varies across passengers. In particular, passengers who check in bags and/or travel frequently are less sensitive to on-time performance, perhaps for the same reasons that they are time-patient. For example, frequent travelers may be able to work productively not only on flights, but also at airports. Another finding is the higher sensitivity of passengers traveling on a Friday or a Saturday to on-time performance.

5.2.4 Access time to airport

A higher access time to the airport in the itinerary leads to a lower likelihood of choosing that itinerary. This is a well established result in the several studies that have focused on airport choice (see, for example, Bhat and Basar, 2003 and Pathomsiri and Haghani, 2005, who find access time as the dominant determinant of airport choice). The interaction model reflects a lower sensitivity to access time for individuals with several vehicles in their household, presumably because these passengers are more likely to drive their cars to the airport and view the private time in their vehicle in a less onerous way. On the other hand, passengers traveling on a Friday or a Saturday view access time more onerously than those traveling on other days, consistent with the "edginess" of Friday/Saturday travelers reflected in their higher sensitivity to on-time performance.

5.2.5 Connecting flight

Business travelers stay away from connecting flights, even after controlling for flight times (which includes connection times). However, a statistically significant difference between having one connection versus two or more connections was not found in the final model, as opposed to non-business, see 5.3.3. The interaction model shows that passengers with checked-in luggage are more tolerant to connections, perhaps because they would have less cabin luggage to transport between gates during connections. The "edginess" of Friday-Saturday business travelers is again apparent in their higher intolerance to itineraries with connections.

5.2.6 Schedule time difference

The schedule difference for an itinerary refers to the difference in arrival time of an itinerary at the destination airport relative to the preferred arrival time (as indicated by the respondent before the stated choice experiments). In the analysis, separate schedule difference variables were used for arrival before the preferred time (early arrival) and arrival after the preferred time (late arrival) to evaluate any asymmetric effects. Several observations can be made from the results for the schedule difference variables in Table 15. First, travelers, as expected, prefer itineraries that get them to their destination airport close to their preferred time of arrival. Second, there are no statistically significant differences in the schedule difference effect between earlier-than-preferred and later-than-preferred arrivals, although there is an indication that travelers prefer arriving too early to too late. Third, the interaction model indicates a lower sensitivity to schedule difference (applicable to both early and late arrivals relative to preferred arrival times) for trips of longer duration. This is rather intuitive, since individuals are likely to have more schedule constraints at their destination if they have short stays, while they can be more flexible if they have longer stays.

5.2.7 Aircraft type

A non-stop itinerary can be associated with four different aircraft types, as presented in the stated choice experiments. These are 1) propeller, 2) regional jet, 3) standard jet (single aisle), and 4) widebody (double aisle). For connecting flights, realistic combinations of these were presented, based on airport pair, airline, and location of connection points. In the analysis, the effect of aircraft type was tested in several ways, including assigning the aircraft type of the primary (longest) flight leg to the entire itinerary, assigning the largest aircraft type in the itinerary to the entire itinerary, assigning the smallest aircraft type to the entire itinerary, and creating separate variables for every possible combination of aircraft type. After extensive testing, the best specification was found to be the one that assigned the smallest aircraft type in the itinerary to the entire itinerary. This is reasonable from a behavioral standpoint, since passengers are likely to focus in on the "weakest" link of the entire itinerary from a safety/turbulence experience standpoint. Overall, the results indicate that business travelers prefer itineraries with standard jets relative to propeller or regional jets, and interestingly, also to widebody jets. The interaction model shows that the preference for standard jets is particularly pronounced for frequent travelers. This is reasonable, since frequent travelers are likely to be more familiar with different types of aircraft and will therefore pay more attention to this attribute when selecting an itinerary.

5.2.8 Airline and airport preferences

As indicated earlier, respondents were asked to rank their preferences of airline and airports prior to undertaking the stated choice exercises. The results in Table 15 are as expected. Individuals mostly prefer itineraries associated with the airline of their first preference, followed by itineraries associated with the airline of their second and third preferences (relative to itineraries with airlines of lower preference). Similarly, respondents prefer itineraries associated with airports of their first preference (passengers were asked to rank airports without regard to access times, so that the rankings could be a reflection of their subjective assessment of the services and quality offered by airports; however, the estimations showed that removing access time from the specification increased the magnitude and significance of the first-ranked airport, indicating that access time was considered by respondents in ranking airports). The estimation results did not indicate preferences for itineraries associated with airports ranked second, third, or fourth compared to lower-ranked airports.

5.2.9 Frequent flyer membership

Previous studies (see Proussaloglou, 1992; Proussaloglou and Koppelman, 1999; Chin, 2002 and Adler *et al.*, 2005) have indicated the strong effect of frequent flyer membership on airline choice. This finding of previous studies is reinforced in the current study. Passengers prefer itineraries with airlines with which they are frequent flyers. The loyalty effect is higher for the elite "gold" members compared to standard or "silver" members. The loyalty effects are only marginally significant, because of the correlation between the preferred airline and frequent flyer membership.

The result, however, is a bit unintuitive. One would expect there to be a difference between the standard and the silver levels. Usually, it is not difficult to become a member of an airline’s frequent flyer program, but it is a more enduring task to collect enough miles to become an elite member, which is why an airline, in which the passenger has silver status is expected to be more attractive than an airline with standard status.

5.2.10 Inertia effect

The stated choice experiments were designed such that the alternative itineraries generated by the computer were not dominated by the revealed choice alternative. However, in 70% of the stated choice experiments, respondents chose their current (revealed choice) alternative instead of the computer-generated itinerary. This is likely due to an inertia effect, where individuals are accustomed to their itinerary and stay with it. The results of the model without interactions support the inertia effect. However, the interaction model indicates that the inertia effect is only prevalent for individuals who purchase their tickets directly from the airline (such individuals are likely to be loyal to their preferred airline). For individuals who purchase their ticket from an agent or online, the inertia effect is neutralized (such individuals are likely to be those who shop around and explore different options, and thus may be more willing to accept the alternative itinerary in the SP experiments). Finally, for individuals who purchase their tickets elsewhere (predominantly corporate travel offices), there is a variety-seeking effect. That is, these individuals prefer the alternative computer generated itinerary, everything else being equal (travelers whose corporate travel offices purchase their ticket may not be involved in the process themselves, and may therefore be more disposed to question whether the travel office made the right choice when selecting an itinerary and/or may have airline preferences different from the preferred corporate airline).

5.2.11 Model fit comparison between the two MNL models

Table 15 indicates the statistical significance of several of the interaction terms in the model with demographic/trip related interactions. One can formally test the two models using a nested likelihood ratio test. The test value is -2(-453.49+392.00) = 122.98, which is substantially larger than the table chi-squared value with 18 degrees of freedom at any reasonable level of significance. Thus, one can strongly reject the null hypothesis that demographic/trip-related variables do not play a role in moderating the sensitivity to service characteristics in air itinerary choice for business travelers.

5.3 Non-business results

This section includes a description of the MNL model results of the 502 non-business passengers. Table 16 shows the results of the two trip purpose segments, with the original model specification of the business passengers and a “slimmer” specification of non-business passengers. It was difficult to obtain statistically significant and/or intuitive demographic interaction variables in the non-business segment, so only a few of these have been kept in the final model.

Table 16. MNL results – Non-business				
Variables	**Business (from Table 15)**		**Non-business**	
	β	*t*-stat	β	*t*-stat
Fare (100 $)	-0.658	*-8.59*	-1.882	*-25.22*
Female travelers	0.225	*1.97*	0.290	*3.26*
Traveling in a group (2 or more)	0.170	*1.50*		
Fare/income for self-paying travelers (per $100 yearly salary)	-2.682	*-3.56*	-4.826	*-2.09*
Flight time (in 100s of min.)	-1.190	*-4.43*	-0.871	*-12.39*
Employment: Transportation or retail trade	-1.469	*-2.06*		
Employment: Professional services	-0.893	*-1.95*		
Employment: Other* (base)	0	*N/A*		
Bags checked in	0.421	*1.40*		
Frequent travelers (# yearly air trips)	0.043	*2.18*		
On-time performance	1.967	*4.29*	0.921	*7.03*
Bags checked in	-1.132	*-2.51*		
Frequent travelers (# yearly flights)	-0.098	*-3.38*	0.932	*3.25*
Travel on a Friday or Saturday	1.561	*2.11*		
Access time to airport (in 100s of min.)	-3.685	*-6.29*	-1.264	*-6.42*
Vehicle owners (# vehicles in household)	1.265	*5.17*	0.970	*1.05*
Travel on a Friday or Saturday	-1.353	*-2.84*		
Connecting flight	-0.908	*-3.99*	-0.299	*-3.71*
Two connections			-0.351	*-2.53*
Bags checked in	0.631	*2.11*		
Travel on a Friday or Saturday	-1.104	*-2.28*		
Schedule time difference (in 100s of min.)				
Early arrival relative to preferred arrival time	-0.588	*-2.35*	-0.223	*-1.90*
Late arrival relative to preferred arrival time	-0.606	*-3.09*	-0.263	*-3.79*
Duration of stay (# nights at destination)	0.048	*1.68*		
Aircraft type				
Regional jet			0.212	*1.76*
Standard jet	0.416	*2.34*	0.380	*3.16*
Widebody			0.937	*5.94*
Frequent travelers (# yearly flights)	2.363	*1.39*		
Airline and airport preferences				
Airline: First preferred	0.687	*3.41*	0.503	*5.80*
Airline: Second or third preferred	0.438	*2.46*	0.318	*3.65*
Airport: First preferred	0.405	*2.45*	0.439	*5.24*
Frequent flyer membership				
Standard or silver member	0.413	*1.97*	0.763	*1.87*
Gold member	1.001	*1.33*		
Inertia (staying with current RP itinerary)	-1.234	*-3.55*	0.492	*6.83*
Purchase directly from airline	1.663	*3.94*		
Purchase from agent or online	1.240	*3.47*		
Other means of ticket purchase** (base)	0	*N/A*		
Number of observations		1180		5020
Number of parameters - *K*		32		20
Log-likelihood at sample shares $L(c)$ †		-720.80		-3066.54
Log-likelihood at convergence - $L(\hat{\beta})$		-392.00		-1508.62
Likelihood ratio index ($\bar{\rho}_c^2$)‡		0.413		0.502

*Includes communications, construction, education, finance/insurance, government, health/medical, manufacturing, marketing/market research, and wholesale trade.

**Mostly corporate travel offices.

† The log-likelihood value at sample shares corresponds to the case where each individual is assigned a 0.7 probability of staying with the revealed preference alternative for each of her/his choice occasions. This is based on the 70% choice of the revealed preference alternative in the stated choice experiments.

‡ $\bar{\rho}_c^2 = 1 - \frac{[L(\hat{\beta}) - (K-1)]}{L(c)}$

5.3.1 Airfare

Table 16 shows that non-business passengers are remarkably more sensitive to airfares than business passengers, which is a very realistic result. The gender interaction variable reveals, as was the case for business passengers, that women are less fare sensitive than men are. The income parameter, however, is more powerful for non-business passengers than for self-paying business passengers, because non-business passengers are expected to be more concerned with the price of the flight. The overall fare sensitivity for the different purpose segments, genders, and income levels is illustrated in Figure 9. It shows that business passengers are less sensitive to fares than non-business passengers, and that, in both segments, women are less sensitive to fares than men. Additionally, it shows that respondents with household incomes above approximately $60,000 have a common marginal utility gain for fares (within each purpose and gender segment), while incomes below that have varying marginal utility values, decreasing as income decreases.

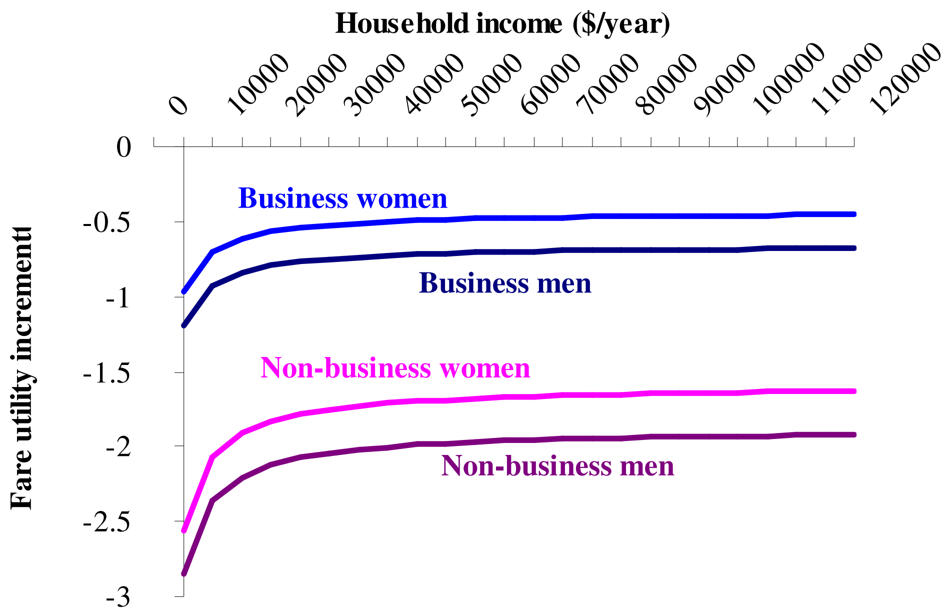

Figure 9. Fare utility increments for four different demographic groups. The utility reductions on the y-axis are for $100 fare increases.

5.3.2 Time variables

The values of the parameters of flight time, on time performance, airport access time, and schedule time difference are all indicators of how important saving time is to passengers, and how much he/she is willing to pay for it. Table 16 shows that all these parameters are numerically smaller for the non-business passengers than the business passengers. This is consistent with the expectation that business passengers value their time higher than non-business passengers. The schedule time difference parameters show that both business and non-business passengers prefer arriving too early than too late. However, as might be expected, business travelers are considerably more sensitive to arriving at a non-preferred time than non-business travelers.

5.3.3 Connecting flight

By splitting the number of connections variable into two different dummy variables, we can examine whether the sensitivity to connections is the same for all increments. In the business model, only the variable for *one or more connections* came out significant, while in the non-business model, the *two connections* variable was significant, too. Thus, the value of this coefficient denotes the difference in utility between one and two connections. What we can see from the models, is that business passengers have a high disutility (-0.908) to connecting flights in general, whether they be single- or double connecting, while non-business passengers are slightly more sensitive to shifting from one to two connections (-0.351) than shifting from zero to one connection (-0.299). In other words, business passengers will likely try to avoid connecting flights if at all possible, while non-business passengers will try to avoid double-connecting flights, but are more likely to accept a single-connecting flight. This piecewise linearity has not been shown in previous studies.

5.3.4 Aircraft type

There are differences between the two segments, concerning the specification of aircraft types. In the business model, only the standard jet variable was significant, meaning that business passengers prefer standard jets to propeller, regional, and widebody jets. However, the non-business segment has the most intuitive values: increasing utility as aircraft size increases. This may be due to the higher number of records in the non-business segment.

5.3.5 Airline and airport preferences

The airline coefficients have the expected signs, and they are almost the same for business and non-business travelers. It was also found, after extensive testing, that the specification with the second and third airline merged was the best for both segments. This phenomenon, as explained in 5.2.8, may be due to the fact that many passengers have a favorite airline, but are less decided about their second and third preferred airlines, making them interchangeable. Additionally, both segments show only the first preferred airport to be significant.

5.3.6 Frequent flyer membership

In the preliminary, unrestricted model, all frequent flyer membership levels had been included, but only some of the variables were significant. In the final business model, a combination of the standard and silver levels resulted in the best specification (albeit unintuitive, see 5.2.9). In the non-business segment, however, a more realistic result is found: it seems that only silver and gold memberships pose an improvement to an itinerary, and therefore, these two variables have been combined into an "elite member" variable. However, as mentioned in section 4.2.9, the data does not contain information about frequent flyer miles obtained from alliance airlines, and this neglect has probably reduced the significance of frequent flyer membership.

5.3.7 Inertia effect

As in the business model, there is a tendency to stay with the RP alternative for non-business travelers. This can be seen by the positive alternative-specific (inertia) variable. The result is consistent with the data: In 70% of the SP questions, the business passengers chose the actual flown RP alternative, while the same was the case for non-business passengers in 69% of the cases. However, the inertia was general for all non-business passengers, since no demographic or trip related interactions with the inertia variable were significant in the estimation tests.

6 MIXED LOGIT RESULTS

The mixed multinomial logit model is used to accommodate taste variations due to unobserved individual factors. Specifically, normal distributions were imposed for the sensitivity to the service characteristics, due to the arguments discussed in section 3.3.1. The mixed logit model was applied to both the business and the non-business segments, but the results for the non-business model were counter-intuitive, and many variables became insignificant in the estimation process. It was impossible to develop a satisfactory model specification for the non-business segment that was both descriptive and intuitive. As a cause, only the business model results are presented here.

Random coefficients were considered on all service attributes (including the inertia variable), except the fare coefficient, which was held fixed for several reasons:

- By keeping a fixed price coefficient, one avoids the fact that the Normal distribution allows the possibility of a positive price coefficient, which is highly unintuitive. In fact, preliminary estimations showed that zero was only 1.2 standard errors away from the mean, indicating that a positive price coefficient would indeed be possible for some individuals.
- Preliminary estimations showed that a random price coefficient made the gender interaction parameter negative and insignificant, contrary to the results from the MNL model. No other interaction coefficients changed signs when included in the MMNL model. As the gender variation to price sensitivity was an *a priori* assumption and very significant in the MNL model, it was desired that it had the appropriate sign in the MMNL model as well.
- Calculations of willingness-to-pay (WTP) values are much simpler when the price coefficient is fixed (Train, 2001). By only using the means and standard deviations of the β's, the WTP can be calculated directly from the estimates. If both parameters were Normally distributed, the ratio would be Gauchy distributed, which has undefined mean and variance (Nielsen and Jovicic, 2003). However, by fixing the price coefficient, one is accepting the point estimate values and thus ignoring sample variance (Hensher & Greene, 2001).

One must keep in mind that a counter-intuitive sign is also an issue for the other variables, such as flight time, access time, connections etc. These are all expected to be negative, but the mixed logit model allows them to become positive through the randomness. This is one of the disadvantages of using the Normal distribution, as discussed in 3.3.1. Note that random coefficients are not imposed on the airline and airport preference variables, and the frequent flyer membership variable, since these are already associated with the individual and are not service attributes of the itinerary.

Table 17 shows the results of the mixed multinomial logit estimation of business travelers with no interaction variables and including interactions variables. The same specification as in the MNL for business travelers (Table 15) has been kept for the

MMNL results, regardless of the values of the t-statistics, in order to compare them adequately. This means that some of the t-statistics of the variables are below 1.00. Thus, the only variables that have been eliminated because of low t-statistics are the standard deviations, because these have not previously been estimated in the MNL model.

Table 17. Mixed multinomial logit (MMNL) model results – business passengers								
Variables	**Service characteristics only**				**Demographic/trip related interactions**			
	Parameter		**Std. dev.**		**Parameter**		**Std. dev.**	
	β	*t-stat*	μ	*t-stat*	β	*t-stat*	μ	*t-stat*
Fare (100 $)	-0.887	*-12.71*			-0.872	*-8.84*		
Female travelers					0.239	*2.47*		
Traveling in a group (2 or more)					0.142	*1.25*		
Fare/income for self-paying travelers (per $100 yearly salary)					-3.482	*-3.76*		
Flight time (in 100s of min.)	-1.048	*-4.14*	0.911	*2.03*	-1.271	*-3.19*		
Employment: Transportation, retail trade or professional services					-1.402	*-2.59*		
Employment: Other* (base)					0	*N/A*		
Bags checked in					0.543	*1.19*		
Frequent travelers (# yearly air trips)					0.041	*1.08*		
On-time performance	1.611	*2.02*	3.381	*4.15*	3.970	*3.57*	2.553	*3.30*
Bags checked in					-2.526	*-2.30*		
Frequent travelers (# yearly flights)					-0.141	*-1.94*		
Travel on a Friday or Saturday					1.706	*1.19*		
Access time to airport (in 100s of min.)	-2.674	*-4.40*	2.279	*4.24*	-4.644	*-3.36*	1.132	*2.11*
Vehicle owners (# vehicles in household)					1.324	*2.19*		
Travel on a Friday or Saturday					-1.247	*-1.26*		
Connecting flight	-1.349	*-4.27*	1.219	*2.41*	-1.183	*-3.28*	1.437	*2.49*
Bags checked in								
Travel on a Friday or Saturday					-0.805	*-0.87*		
Schedule time difference (in 100s of min.)								
Early arrival relative to preferred arrival time	-0.179	*-0.51*	0.521	*0.54*	-0.449	*-1.35*		
Late arrival relative to preferred arrival time	-0.394	*-1.55*	0.977	*2.20*	-0.412	*-2.23*		
Duration of stay (# nights at destination)								
Aircraft type: Standard jet	0.360	*1.14*	0.055	*0.05*	0.414	*1.34*		
Frequent travelers (# yearly flights)								
Airline and airport preferences								
Airline: First preferred	0.928	*2.88*			0.959	*2.96*		
Airline: Second or third preferred	0.509	*1.69*			0.553	*1.78*		
Airport: First preferred	0.514	*1.91*			0.413	*1.44*		
Frequent flyer membership								
Standard or silver member	0.608	*1.61*			0.502	*1.41*		
Gold member	1.632	*1.57*			2.003	*2.04*		
Inertia (staying with current RP alternative)	0.104	*0.35*	0.035	*0.05*	-1.182	*-1.88*		
Purchase directly from airline					1.632	*1.80*		
Purchase from agent or online					1.073	*1.60*		
Other means of ticket purchase** (base)					0	*N/A*		
Number of observations	118				118			
Number of parameters - *K*	14 + 8 SD's = 22				28 + 3 SD's = 31			
Log-likelihood at sample shares $L(c)$†	-720.80				-720.80			
Log-likelihood at convergence - $L(\hat{\beta})$	-396.82				-365.16			
Likelihood ratio index ($\bar{\rho}_c^2$)‡	0.420				0.452			

*Includes communications, construction, education, finance/insurance, government, health/medical, manufacturing, marketing/market research, and wholesale trade.

**Mostly corporate travel offices.

† The log-likelihood value at sample shares corresponds to the case where each individual is assigned a 0.7 probability of staying with the revealed preference alternative for each of her/his choice occasions. This is based on the 70% choice of the revealed preference alternative in the stated choice experiments.

‡ $\bar{\rho}_c^2 = 1 - \frac{[L(\hat{\beta}) - (K-1)]}{L(c)}$

In the model with service characteristics only, the standard deviations of early arrival relative to preferred arrival time, aircraft type, and the inertia variable are not significant, while the rest of the service characteristics have significant standard deviations. Overall, the results show that there is unobserved heterogeneity in the respondents' itinerary choice. The signs and magnitudes of the different parameters correspond roughly to the parameters in the MNL model, and so the interpretation of the parameters that was thoroughly discussed in chapter 5.2, can also be applied to this model.

In the MMNL model with demographic/trip-related interactions, the standard deviations characterizing random taste variations in flight time and the scheduled time difference variables also turned out to be statistically insignificant. Thus, the interaction model indicates random taste variation only in response to on-time performance, access time to the airport, and presence of a connecting flight. Further, it can be observed that the distribution due to random effects, in general, becomes narrower in the interaction model compared to the model with only service characteristics. While the coefficients of the two models cannot be directly compared due to differing overall scales, one can characterize the variation of each random coefficient relative to the mean value. For example, in the model with service characteristics only, the ratio of the standard deviation to the mean coefficient for on-time performance is 3.381/1.611 = 2.10. In the model with interactions, the equivalent ratio for individuals with the highest level of variation relative to the mean (corresponding to frequent travelers who check-in their bags), is 2.553/(3.970-0.141-2.526) = 1.96. The equivalent ratio for the vast majority of frequent travelers who do not check-in their baggage is 2.553/(3.970-0.414) = 0.67. Overall, the level of random taste variation reduces substantially after accounting for systematic taste variations. This indicates the pitfalls of ignoring systematic taste variation when accommodating random taste variation.

As a follow-up to the discussion of the choice of distribution in section 3.3.1, it could be interesting to examine whether the random parameters actually risk having counter-intuitive signs. This way, we can conclude if the Normal distribution was appropriate, or if a bounded distribution should have been chosen. By looking at the standard deviations in the demographic MMNL model, it is clear that, for on-time performance and access time, adding or subtracting one standard deviation does not move the parameter across zero. More properly, we can calculate the percentage of respondents positioned in the area with the appropriate sign, by standardizing the distribution:

$$F(z_{OTP}) = F(\tfrac{0-3.970}{2.553}) = 0.062$$

$$F(z_{Access\ time}) = F(\tfrac{0+4.644}{1.132}) = 0.99997$$

$$F(z_{Connections}) = F(\tfrac{0+1.183}{1.437}) = 0.794$$

We can see that the on-time performance variable is expected to be positive for 94% of the population, and the access time variable is negative for practically 100%. These are satisfactory results since almost all the respondents will fall into the right category. The connection parameter, however, has a 21% risk of being positive, which is a large share. However, while we may expect time parameters to be negative for *all* travelers, the connection parameters may not always be confined to one sign. Some respondents may actually favor a connecting flight to a nonstop flight, for example if they plan to do some shopping in a major transfer airport. In that regard, the 21% is not such an alarming number.

In conclusion, the random parameter results show that it seems appropriate to use the Normal distribution, taking the above considerations into account as well as the reasons discussed in 3.3.1.

6.1.1 Model fit comparisons between the two MMNL models

The model without demographic/trip-related interactions and the model with demographic/trip-related interactions can be compared using a likelihood ratio test. To be fair, the insignificant standard deviation estimates on the "early arrival relative to preferred arrival time", aircraft type, and the inertia variables from the "service characteristics only" model were dropped, for the statistical test. The likelihood ratio test value is 64, which is higher than the chi-squared table value with 13 degrees of freedom at even the 0.001 level of significance. Thus, even in the MMNL framework, one can strongly reject the null hypothesis of no demographic/trip-related interaction effects.

6.1.2 Model fit comparisons between the MMNL and MNL models

The MMNL models in Table 17 can be compared to their respective MNL counterparts in Table 15 using standard likelihood ratio tests. For both the "service characteristics only" model and the model with "demographic/trip-related interactions", the MMNL model clearly turns out to have a statistically significant superior fit. This suggests the importance of accommodating random taste variations. However, it is also illustrative to compare the performance of the MNL model with "demographic/trip-related interactions" to the MMNL model with "service characteristics only". The two models have about the same $\overline{\rho}_c^2$ value (see Table 15 and Table 17). This highlights the importance of accommodating systematic variation in as comprehensive a way as possible before proceeding to introduce random taste variations.

7 TRADE-OFF ANALYSIS

This section includes a discussion of the results of trade-off calculations for the six service characteristics of flight time, on-time performance, access time to airport, connecting flights, schedule time difference, and aircraft type. The focus is only on the MMNL models, and therefore only on business passengers. In the MMNL model with "service characteristics only", the insignificant standard deviations for "early arrival time relative to preferred arrival time", "aircraft type", and "inertia" variables are not considered. Table 18 shows the trade-off values. In the first column, the service-characteristics-only model has been used to calculate trade-off values. The second column shows the increments in willingness-to-pay (WTP) that each population group imposes. To arrive at a consolidated mean WTP and standard deviation for service characteristics in the interactions model, the WTP for each segment is weighted by the proportion of travelers in that segment in the sample and the third column shows this weighted average. For example, the increment for the fare variable for women has been multiplied by 0.38, since 38% of the sample respondents are female. Hence, the values in the third column are supposed to represent what the airlines can expect from a "typical" average business passenger.

Table 18. Trade-off values for business passengers calculated from MMNL results

Service attributes	**Service characteristics only model**		**Model with demographic/trip related interactions**		
			Separate demographic/trip related trade-offs	**Weighted averages based on proportions in sample**	
	Mean	SD		Mean	SD
Flight time (-1 hour)	$ **71**	*62*	$ 80	$ **83**	
Employment: Transportation, retail trade or professional services			$ +89		
Frequent travelers (# yearly flights)			$ -3		
Bags checked in			$ -34		
On time performance (+10 %)	$ **18**	*38*	$ 42	$ **20**	*27*
Frequent travelers (# yearly flights)			$ -1		
Bags checked in			$ -27		
Travel on a Friday or Saturday			$ +18		
Access time to airport (-1 hour)	$ **181**	*154*	$ 294	$ **136**	*72*
Vehicle owners (# vehicles in household)			$ -84		
Travel on a Friday or Saturday			$ +79		
Connecting flight (WTP for non-stop)	$ **152**	*137*	$ 125	$ **137**	*151*
Travel on a Friday or Saturday			$ +85		
Schedule time difference: early (-1 hour)	$ **20**		$ 28	$ **28**	
Schedule time difference: late (-1 hour)	$ **44**	*66*	$ 26	$ **26**	
Aircraft type: Standard jet (vs. other aircraft)	$ **41**		$ 44	$ **44**	

The results in Table 18 show that travelers are, on average, willing to pay $71 for a one-hour reduction in flight time, $18 for an improvement in on-time performance by 10%, $181 for a one hour reduction in airport access time, $152 for a non-stop itinerary compared to a connecting itinerary, $20 for reducing the schedule time difference for early arrivals by an hour, $44 for reducing the schedule time difference for late arrivals by an hour, and $41 for traveling on a standard jet compared to other aircraft types. These trade-off values provide information on the relative values of service attributes that could be used by airline carriers and airport management in evaluating service changes. It is also important to note that there is considerable random variation around these mean trade-off values in the MMNL model with service characteristics. These variations could be used to identify clusters of individuals with different preferences who could, in theory, be targeted by specific service changes.

The hourly value of flight time is quite consistent with the one found in Adler et al. (2004), which was \$69.7 with a standard deviation of \$39.2. The WTP for on time performance is lower than the findings in Adler et al. (2004): \$38.2 with an SD of \$40.8. Access time and connections WTP's are significantly higher than what was found in the 2003 data, \$62.5 and \$53.7, respectively (although the latter was for switching from single-connecting to direct flight only).

The mean trade-off values vary by demographic/trip-related characteristics of the traveler in the MMNL model with demographic/trip-related interactions. It is straightforward to compute the mean values for any group of travelers. For example, the mean willingness to pay (WTP) for a one-hour reduction in flight times for individuals with the following characteristics – male, traveling alone, company-paid ticket, not employed in transportation, retail or professional services, 2 air trips per year, no bags checked-in, and traveling Sunday through Thursday – is $[(-1.190 + 2 * 0.043)/-0.658] \cdot 60 = \$100.00$. The customized WTP values for each market segment of travelers provides information to air carriers and airport management that could be used to more effectively position and target their service improvements.

A comparison of the mean WTP values between the "service characteristics only" model and the interactions model indicates an underestimation in the WTPs for flight time, on-time performance, early arrival relative to preferred arrival time, and aircraft type in the former model, and an overestimation of the WTPs for access time, non-stop flights, and late arrival relative to preferred arrival time. Further, the overall levels of variation due to unobserved factors in the WTPs have been reduced in the interaction model for all service attributes except the "connecting flight" variable.

7.1 Application example

This section presents a basic example of how airlines can use the trade-off results to estimate the willingness-to-pay for different routes in their network, and thereby adjust the airfares accordingly.

Imagine that an airline has a route from Houston Hobby Airport to New York that connects in Chicago. The current return fare for the route is \$200, and the journey takes 5 hours including transfer time in Chicago. The airline knows that they have demand enough for a nonstop route and want to calculate how much they could charge for the new route. Because of capacity constraints at Hobby Airport, the new route will have to fly out of Bush Intercontinental Airport. See Figure 10.

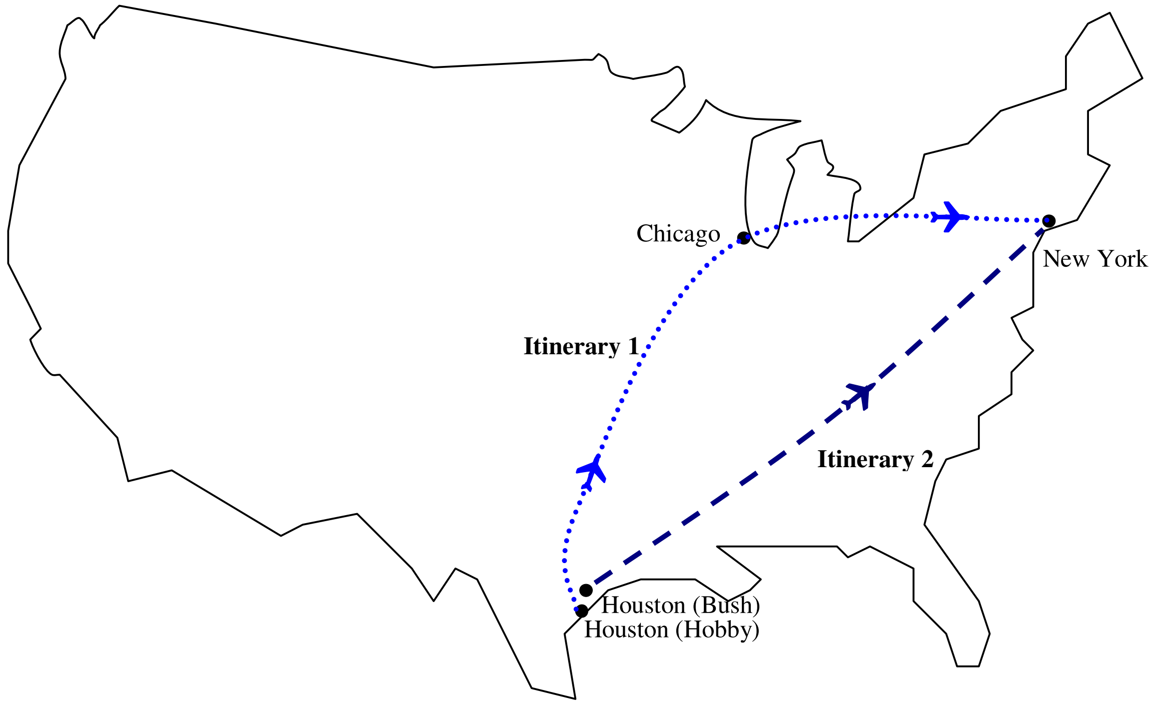

Figure 10. Map of the existing itinerary 1 and the proposed new itinerary 2 for a hypothetic airline.

Table 19 shows the different attribute values for the two itineraries. The new nonstop route (Itinerary 2) will cut the travel time down to 3.5 hours. On time performance for the old itinerary was 90% for each flight leg, which gives a total on time performance of 81%. The new route is estimated to have an on time performance of 90%. Hobby Airport is closer to downtown Houston than Bush Airport, so the airline estimates that the average travel time to Hobby is 30 minutes and 1 hour for Bush. Itinerary 1 has one connection, while Itinerary 2 has none. Finally, the flight leg Houston-Chicago is served by a standard jet, while Chicago-New York is served by a regional jet. Itinerary 2 will be flown by a standard jet.

Table 19. Service attributes for the two hypothetical itineraries

Service attributes	**Itinerary 1** Houston (Hobby) – Chicago – New York	**Itinerary 2** Houston (Bush) – New York
Flight time	5 hours	3.5 hours
On time performance	90% · 90% = 81%	90%
Access time to departure airport	30 minutes	1 hour
Connections	1	0
Aircraft type	Standard jet + regional jet	Standard jet
Fare	$200	**??? – we want to calculate this**

If we assume that the sample used in this report is a good representation of the current airline's market on the route Houston-New York, we can use the weighted WTP values from Table 18 to calculate what the average willingness-to-pay for business travelers would be for the new route. This would be the fare, in which there is equilibrium between the two itineraries, i.e. passengers would have equal utility gains by choosing either one:

$$\begin{aligned} WTP &= \$200 + 1.5\text{ hr} \cdot \$83\text{ per hr} + 9\%\text{ otp} \cdot \$2\text{ per }\%\text{ otp - }0.5\text{hr} \cdot \$136\text{ per hr} \\ &\quad + \$137\text{ for connecting flight} + \$44\text{ for standard vs. regional jet} = \mathbf{\$456} \end{aligned}$$

Similarly, the standard deviation of this fare can be calculated:

$$SD = 9\%\text{ otp} \cdot \$2.7\text{ per }\%\text{ otp} + 0.5\text{ hr} \cdot \$72\text{ per hr} + \$151\text{ per connecting flight} = \pm 211$$

This means that it would be feasible to set the fare for Itinerary 2 between $245 and $667, with a mean of $456. Of course, this is a very simplified example, and in reality, factors such as competing airlines, attracted customers, advertising issues, and the demand on Itinerary 1 would very much influence the designated fare for the new route. However, it does provide an idea of how the airline can forecast the demand for future routes and charge their passengers accordingly.

8 CONCLUSION

This report contributes to the air travel behavior analysis area by estimating an itinerary choice model that considers taste variations in air service characteristics due to both observed traveler/trip characteristics and unobserved factors. This is an important shift in the field from examining isolated air travel choice dimensions to analyzing the multidimensional set of choices implicit in selecting an itinerary.

The sample used in this report is drawn from a 2001 online survey of air travelers. The online survey collected information on the respondents' most recent flight within the US. The survey also generated ten alternative itineraries and presented these itineraries to respondents along with the actual chosen itinerary, in a series of 10 binary stated choice exercises. Each itinerary was characterized by several air service characteristics, including fare, flight time, on-time performance, access time to airport, presence or absence of connecting flights in the itinerary, and time difference between the traveler's preferred arrival time at the destinations and the itinerary's arrival time. The empirical analysis considered all of these service characteristics, along with an inertia variable to stay with the actual chosen itinerary and interactions of individual/trip attributes with the service characteristics. The estimations used a simple multinomial logit structure for the business and non-business segments and a mixed multinomial logit structure for the business segment, the latter structure being used to accommodate random taste variations (across travelers) to service characteristics. In both these structures, specifications without and with individual/trip interactions with service characteristics were considered.

Two general observations may be drawn from the empirical results. First, and most importantly, the results highlight the importance of considering demographic/trip interactions with service characteristics in a comprehensive manner. There are important and statistically significant sensitivity variations to service characteristics across traveler and trip segments, which are masked when the interactions are ignored. Consequently, ignoring the interactions leads to

- An inability to understand the differential trade-offs between, and willingness to pay (WTP) for, service characteristics across different sub-populations of the air travel market.
- Inconsistent estimates of the trade-offs and WTP values for the air travel market as a whole.
- Inappropriately high manifestations of random taste variation.
- Significantly poorer fit measures. In fact, with respect to model fit, the results indicated that a simple multinomial logit model with comprehensive consideration of demographic/trip interactions provides as good a fit as a mixed multinomial logit model with no consideration of demographic/trip interactions.

In addition to these empirical considerations, from a fundamental theoretical and model-building standpoint, it is critical to first accommodate systematic variations comprehensively before proceeding to consider random taste variation. The introduction of unobserved taste variation should not be in lieu of observed taste variation, but only to recognize the inevitable presence of unobserved random variations even after the most comprehensive control of observed factors.

Second, there are statistically significant random taste variations (across individuals and trips) to air service characteristics, even after incorporating demographic/trip interactions. Ignoring these random taste variations leads to inconsistent trade-off and WTP values, and significantly poorer model fit statistics. Overall, the results highlight the importance of considering both systematic and random taste variations to (1) effectively position and target air service improvements and (2) accurately predict air travel demand.

There are also several specific and important results from the analysis regarding the service determinants of itinerary choice of business travelers. Women, individuals traveling in a group, and high income earners are less sensitive to fares then men, individuals traveling alone, and low income earners, respectively (the result regarding income is applicable only to paying travelers, not company-paid travelers). Frequent travelers and travelers who check in bags are more time-patient, less likely to be influenced by on-time performance, and more tolerant to connections than occasional travelers and travelers who do not check bags. Travelers on Fridays and Saturdays highly value on-time performance, access time, and non-stop flights relative to business travelers on other days of the week. Travelers staying at their destinations for short periods of time are more sensitive (than travelers staying for long periods) to arriving at their destination airport close to their desired arrival time. Travelers very highly value reductions in access times to airports and non-stop flights; specifically, business travelers are willing to pay, on average, about $136 for a one-hour saving in access time to the airport and about $137 more for a non-stop flight itinerary relative to a connecting flight itinerary.

Some of these tendencies were also found for non-business travelers. However, the demographic variation was not as distinct. Women and high-income earners are less fare sensitive than men and low-income earners, frequent travelers are concerned about on-time performance, and vehicle owners are less concerned about airport access time. Unlike the business travelers, it was found that the shift from single- to double-connecting flights pose a greater disutility than the shift from nonstop to single-connecting flights.

We can summarize the most important findings of this report with the following observations:

- The focus on itinerary choice as a whole is a major advance, and results in a more realistic representation of the choice process that air passengers are going through.
- The report presents, for the first time, a combination of observed and unobserved demographic and trip related variations. Many of the unobserved variations disappear or become narrower, when controlling for observed variations.
- The trade-off values for business passengers are consistent with the findings in Adler *et al.* (2005). Only the parameters on-time performance, access time, and connections were associated with randomness, after eliminating insignificant standard deviations on the remaining parameters.

9 REFERENCES

Adler, T., C. S. Falzarano, and G. Spitz). Modeling Trade-offs in Air Itinerary Choices. Presented at 84th Annual Meeting of TRB, Washington, D.C., 2005.

Algers, S., P. Bergström, M. Dahlberg, and J. L. Dillén. Mixed Logit Estimation of the Value of Travel Time. Working Paper, Department of Economics, Uppsala University, Uppsala, 1998.

Algers, S. and M. Beser. Modeling choice of flight and booking class – a study using Stated Preference and Revealed Preference data. *International Journal of Service Technology and Management*, Vol. 2, Nos.1-2, 2001.

Ashford, N. and M. Benchemam. Passengers' Choice of Airport: An Application of the Multinomial Logit Model, *Transportation Research Record*, No. 1147, 1987, pp. 1-5.

Basar, G. and C. R. Bhat. A parameterized Consideration Set Model for Airport Choice: An Application to the San Francisco Bay Area. Research Report, Center for Transportation Research, Austin, Texas, 2003.

Bhat, C.R. Incorporating Observed and Unobserved Heterogeneity in Urban Work Mode Choice Modeling. *Transportation Science*, Vol. 34, No. 2, 2000, pp. 228-238.

Bhat, C.R. Quasi-Random Maximum Simulated Likelihood Estimation of the Mixed Multinomial Logit Model. *Transportation Research Part B*, Vol. 35, 2001, pp. 677-693.

Bhat, C.R. Simulation Estimation of Mixed Discrete Choice Models Using Randomized and Scrambled Halton Sequences. *Transportation Research Part B*, Vol. 37, No. 9, 2003, pp. 837-855.

Bhat, C.R., and R. Sardesai. On Examining the Impact of Stop-Making and Travel Time Reliability on Commute Mode Choice: An Application to Predict Commuter Rail Transit Mode for Austin, TX. Technical paper, Department of Civil Engineering, The University of Texas at Austin, August 2004.

Brownstone, D., C. Kazimi, A. Ghosh, T. F. Golob, and D. van Amelsfort. Drivers' Willingness-to-Pay to Reduce Travel Time: Evidence from the San Diego I-15 Congestion Pricing Project. *Transportation Research A*, Vol. 37, 2003, pp. 373-387.

Carson, R. T., J. J. Louviere, D. A. Anderson, P. Arabie, D. S. Bunch, D. A. Hensher, R. M. Johnson, W. F. Kuhfeld, D. Steinberg, J. Swait, H. Timmermans, and J. B. Wiley. Experimental Analysis of Choice. *Marketing Letters*, Vol. 5, No. 4, 1994, pp. 351–368.

Chin, A. T. H. Impact of Frequent Flyer Programs on the Demand for Air Travel. *Journal of Air Transportation,* Vol. 7, 2002, pp. 53-86.

Coldren, G. M. and F. S. Koppelman. Modeling the Underlying Competitive Dynamic among Air-travel Itineraries with Ordered Generalized Extreme Value Models. Presented at 84th Annual Meeting of TRB, Washington, D.C., 2005.

Furuichi, M. and F. S. Koppelman. An analysis of air traveler's departure airport and destination choice behaviour. *Transportation Research*, Vol. 28A, 1994, pp. 187-195.

Gunn, H. F., M. A. Bradley, and D. A. Hensher. High Speed Rail Market Projection: Survey Design and Analysis. *Transportation*, Vol. 19, 1992, pp. 117-139.

Harvey, G. Airport Choice in a Multiple Airport Region. *Transportation Research*, Vol. 21A, No. 6, 1987, pp. 439-449.

Hensher D. *Estimation and application of discrete choice models using combined stated and revealed preference data: Comments and advice on the mode choice study of train services*. CIS Associados, 1999.

Hensher, D. and J. Louviere. Combining sources of preference data for modeling complex decision processes. *Marketing Letters*, Vol. 10, No. 3, 1999, pp. 205-218.

Hensher, D. A. and W. H. Greene. *The Mixed Logit Model: The State of Practice and Warnings for the Unwary*, 2001.

Hensher, D. Identifying the influence of stated choice design dimensionality on willingness to pay for travel time. *Journal of Transport Economics & Policy*, Vol.38, No. 3, 2004, pp. 425-446.

Hess, S., M. Bierlaire, and J. W. Polak. Estimation of Value-of-Time using Mixed Logit models. Presented at 84th Annual Meeting of TRB, Washington, D.C., 2005.

Hess, S. and J. W. Polak. Mixed Logit Modeling of Airport Choice in Multi-Airport Regions. Presented at 84th Annual Meeting of TRB, Washington, D.C., 2005.

Hess, S. and J. W. Polak. On the use of Discrete Choice Models for Airport Competition with Applications to the San Francisco Bay area Airports. Paper presented at the 10th triennial World Conference on Transport Research, Istanbul, 2004.

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

Innes, J. D. and D. H. Doucet. Effects of access distance and level of service on airport choice. *Journal of Transportation Engineering*, Vol. 116, 1990, pp. 507-516.

Jara-Díaz, S.R. Time and Income in Travel Demand - Towards a microeconomic activity framework. In *Theoretical Foundations of Travel Choice Modelling*. Eds. Gärling, T., T. Laitila, and K. Westin. Pergamon, 1998, pp. 51-74.

Koppelman, F., C. Bhat, V. Sethi, and B. Williams. *A Self-Instructing Manual on Discrete Choice Modeling*. Unpublished report prepared for Federal Highway Administration, 1995.

Louviere, J. J., D. Street, R. Carson, A. Ainslie, J. R. DeShazo, T. Cameron, D. Hensher, R. Kohn. and T. Marley. Dissecting the Random component of Utility. *Marketing Letters*, Vol. 13, No. 3, 2002, pp. 177-193.

Nielsen, O. A. *Tidsværdier i trafikmodeller*. Working paper, Center for Traffic and Transport, Technical University of Denmark, Lyngby, 2005.

Nielsen, O. A. and G. Jovicic. The AKTA road pricing experiment in Copenhagen. Presented at the 10th International Conference on Travel Behaviour Research, Lucerne, 2003.

Ndoh, N. N., D. E. Pitfield, and R. R. Caves. Air transportation passenger route choice: a Nested Multinomial Logit analysis. In Fisher, M. M., P. Nijkamp, and Y. Y. Papageorgiou (Eds.), *Spatial Choices and Processes*. Elsevier Science Publishers, Amsterdam, 1990.

OAG Worldwide Limited. *Official Airline Guide*. Bedfordshire, LU5 4HB, United Kingdom, 2001.

Ozoka, A. I. and N. Ashford. Application of disaggregate modelling in aviation systems planning in Nigeria: a case study. *Transportation Research Record*, No. 1214, 1989, pp. 10-20.

Pathomsiri, S. and A. Haghani. Taste Variation in Airport Choice Models. Presented at 84th Annual Meeting of TRB, Washington, D.C., 2005.

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

Pels, E., P. Nijkamp, and P. Rietveld. Airport and airline choice in a multi-airport region: an empirical analysis for the San Francisco bay area. *Regional Studies*, Vol. 35, 2001, pp. 1-9.

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

Proussaloglou, K. E. *Air Carrier, Fare Class and Flight Demand: An Individual Choice Approach*. PhD Dissertation. Northwestern University, Evanston, Illinois, 1992.

Richardson, A. J., E. S. Ampt, and A. H. Meyburg (1995). *Survey Methods for Transport Planning*. Eucalyptus Press. Melbourne, Australia.

Resource Systems Group. *Air Travel 2001. What do they tell us about the future of US air travel?* Industry report, 2001.

Sivakumar, A., C. R. Bhat, and G. Ökten. Simulation Estimation of Mixed Discrete Choice Models Using Randomized Quasi-Monte Carlo Sequences: A Comparison of Alternative Sequences, Scrambling Methods, and Uniform-to-Normal Variate Transformation Techniques. Presented at 84th Annual Meeting of TRB, Washington, D.C., 2005.

Skinner, R. E. Jr. Airport Choice: An empirical study. *Transportation Engineering Journal,* Vol. 102, 1976, pp. 871-883.

Swait, J., and J. Louviere. The Role of the Scale Parameter in the Estimation and Comparison of Multinomial Logit Models. *Journal of Marketing Research*, Vol. 30, 1993, pp. 305-314.

Sørensen, M. J. *Discrete Choice Models. Estimation of Passenger Traffic.* Ph.D. Thesis, Center for Traffic and Transport, Technical University of Denmark, Lyngby, 2003.

Thompson, A. and R. Caves. The projected market share for a new small airport in the south of England. *Regional Studies*, Vol. 27, 1993, pp. 137-147.

Train, K. *A Comparison of Hierarchical Bayes and Maximum Simulated Likelihood for Mixed Logit*. University of California, Berkeley, 2001.

Train, K. *Discrete Choice Methods with Simulation.* Cambridge University Press, Cambridge, 2003.

Whitaker, B., G. Terzis, E. Soong, and W. Yeh. Stated Preference as a Tool to Evaluate Airline Passenger Preferences and Priorities. Presented at 84th Annual Meeting of TRB, Washington, D.C., 2005.

Windle, R. and M. Dresner. Airport choice in multi-airport regions. *Journal of Transportation Engineering*, Vol. 121, 1995, pp. 332-337.

Yoo, K. E. and N. Ashford. Carrier Choices of Air Passengers in Pacific Rim: Using Comparative Analysis and Complementary Interpretation of Revealed Preference and Stated Preference Data. *Transportation Research Record*, No. 1562, 1996, pp. 1-7.

Appendix A: Survey instrument

This table lists selected questions and results from the online survey that is used in this report. The percentages are for the business passengers only.

Questions about the trip		
When did you make the trip you have in mind?	☐ In the past month ☐ 1-6 months ago ☐ 7-12 months ago	34% 40% 26%
How did you acquire your ticket?	☐ Travel agent ○ Local ○ Local office of national agent ○ National agent without local office ☐ Directly from the airline ○ By telephone ○ In person ☐ Online – using the airline website ☐ Online – from a travel site other than the airline ☐ Other	46% *45%* *27%* *27%* 15% *72%* *28%* 19% 12% 8%
How many associates, friends, or family members traveled together in your party?	☐ 1 (traveled alone) ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 or more	65% 19% 4% 3% 5% 4%
How many nights were you away on your trip?	☐ 0 (left and returned on the same day) ☐ 1 night ☐ 2 nights ☐ 3 nights ☐ 4 nights ☐ 5 nights ☐ 6 nights ☐ 7 nights ☐ 8-14 nights ☐ 15-21 nights ☐ 22 nights or more	8% 11% 23% 19% 10% 8% 8% 5% 3% 2% 3%
Who paid for the ticket?	☐ I paid, personally ☐ My company paid or reimbursed me	24% 76%
How much did the ticket cost?	(continuous variable)	
Round trip or One-way	☐ Round trip ☐ One-way	95% 5%
On what day of the week did you depart?	☐ Monday ☐ Tuesday ☐ Wednesday ☐ Thursday ☐ Friday ☐ Saturday ☐ Sunday	18% 18% 21% 14% 9% 6% 14%
What travel mode did you use to get to the airport?	☐ Car: Drove myself and parked ☐ Car: Was dropped off ☐ Taxi ☐ Shuttle ☐ Bus ☐ Train	44% 39% 9% 6% 0% 2%
Did you check any baggage for your flight?	☐ Yes ☐ No	53% 47%

Questions about the flight flown (RP alternative)		
Was your flight on time?	☐ Yes	80%
	☐ No	20%
If late, how late was the flight?	(continuous variable)	
Did you have a direct or connecting flight?	☐ Direct flight	72%
	☐ One connection	27%
	☐ Two connections	1%
What airline did you use?	☐ Airtran	2%
	☐ America West	2%
	☐ American	9%
	☐ American Eagle	1%
	☐ Alaskan	2%
	☐ Continental	9%
	☐ Continental Express	1%
	☐ Delta	18%
	☐ Delta Express	2%
	☐ Midway Airlines	1%
	☐ Northwest	10%
	☐ Southwest	15%
	☐ TWA	2%
	☐ United	12%
	☐ United Express	1%
	☐ United Shuttle	1%
	☐ US Airways	8%
	☐ US Airways Express	1%
	☐ US Airways Shuttle	1%
	☐ Other	2%
How full was your flight?	☐ Full	48%
	☐ More than three-quarters, but not completely full	44%
	☐ Half to three-quarters full	7%
	☐ Less than half full	1%
What class of service did you use?	☐ Economy or coach	77%
	☐ Business	19%
	☐ First class	4%
Did you receive a free upgrade to this class?	☐ Yes	21%
	☐ No	79%

Demographic questions		
What is your gender?	☐ Male	62%
	☐ Female	38%
What is your employment status?	☐ Full-time worker	78%
	☐ Part-time worker	3%
	☐ Self-employed	9%
	☐ Student	3%
	☐ Retired	4%
	☐ Homemaker	1%
	☐ Unemployed	1%
What is your employment position?	☐ Clerical/secreterial	3%
	☐ Executive/managerial	32%
	☐ Professional/technical	44%
	☐ Sales/buyer	9%
	☐ Teacher/professor	4%
	☐ Retail/service	5%
	☐ Other non-office worker	3%
What type of industry do you work in?	☐ Communications	8%
	☐ Construction	4%
	☐ Education	5%
	☐ Finance/insurance	8%
	☐ Government	9%
	☐ Health/medical	11%
	☐ Manufacturing	9%
	☐ Marketing/market research	1%
	☐ Retail trade	3%
	☐ Transportation	2%
	☐ Professional services	19%
	☐ Other	21%
What is your annual household income?	☐ $10,000 - $19,999	5%
	☐ $20,000 - $29,999	7%
	☐ $30,000 - $39,999	11%
	☐ $40,000 - $49,999	8%
	☐ $50,000 - $74,999	26%
	☐ $75,000 - $99,000	17%
	☐ $100,000 or more	26%
How many US air trips did you make in the last year (paid for personally or by company)?	☐ 1	15%
	☐ 2	15%
	☐ 3	14%
	☐ 4	7%
	☐ 5	9%
	☐ 6	8%
	☐ 7	6%
	☐ 8-10	9%
	☐ 11-15	11%
	☐ 15 or more	6%
How often do you make domestic trips for this purpose?	☐ Once per week or more	1%
	☐ 1-3 times per month	17%
	☐ 7-11 times per year	17%
	☐ 1-6 times per year	55%
	☐ Less than once per year	10%