ON THE USE OF DISCRETE CHOICE MODELS FOR AIRPORT CHOICE WITH APPLICATIONS TO THE SAN FRANCISCO BAY AREA AIRPORTS

Stephane Hess and John W. Polak
Centre for Transport Studies - Imperial College London, Exhibition Road - London SW7 2AZ
Tel: +44(0)20 7594-6105, +44(0)20 7594-6089 Fax: +44(0)20 7594-6102, +44(0)20 7594-6102
stephane.hess@imperial.ac.uk, j.polak@imperial.ac.uk

Abstract

In this paper, we describe an in-depth analysis of the combined choice of departure airport, airline and access-mode for passengers departing from the San Francisco Bay area. The analysis shows that several factors, most notably flight frequency and in-vehicle access-time, have a significant overall impact on the appeal of an airport, while factors such as fare and aircraft size have a significant effect only in some of the population subgroups. The analysis highlights the need to use separate models for resident and non-resident travellers, and to segment the population by journey purpose. The analysis also shows that important gains can be made through accounting for past experience at the different airports and through using a non-linear specification for the marginal returns of increases in flight frequency. In terms of model structure, the results suggest that the use of the Nested Logit model leads to significant improvements in model fit over the use of the Multinomial Logit model, although these improvements do not necessarily translate into significant advantages in prediction performance, which is already surprisingly good in the base models, showing the importance of using a detailed specification of utility.

Keywords: Airport choice behaviour; Aviation demand; Nested Logit; Airport-access; Airline choice
Topic area: A3 Airports and Aviation

1. Introduction

The analysis of air-travellers' choices of airports is an important component in the long-term transport strategies in many metropolitan areas that are served by more than one airport. A wide range of policy measures potentially affect airport choice, including expansion of airport capacity in multi-airport regions, improved access-service to an airport, changes to an airport's parking-cost structure, and the introduction of faster check-in procedures at an airport. In turn, the outcome of travellers' airport choice decisions will affect the commercial success of the single airports, the financial viability of auxiliary and complementary businesses, the congestion in the local transportation network and the local and regional environment.

Studies of airport choice have become increasingly popular over recent years. While most of the studies have used very basic models in the simple analysis of the choice of departure airport, several studies have employed advanced models in the joint analysis of the choice of airport and the choice of access mode. Recently, it has also been shown that important gains in model performance can be obtained by accommodating the fact that passenger behaviour varies not only deterministically across different groups of travellers (e.g. business/leisure), but also randomly within individual groups of travellers (c.f. Hess and Polak, 2004).

It is important to recognise that passengers do in fact not simply make a choice of airport, but make a three-way choice of airport, airline and access mode (excluding the upper-level choices of destination and main-mode). The nature of the substitution patterns amongst these choices is not clear *a priori*. While some studies have recognised this issue, the majority of published work does at best look at only two of these choice dimensions, and uses some form of simplification along the third dimension (c.f. section 2). Another problem with many existing studies is the use of over-aggregated data for the air-transport and ground-transportation level-of-service information. These deficiencies in the existing body of work are the main motivation for the present research effort.

In this paper, we look at the combined choice of airport, airline and access-mode for passengers departing from the San Francisco Bay (SF-bay) area. The aim of this study is to determine the optimal structure for the joint analysis of these choices, with a view to using this structure in a later analysis incorporating random taste heterogeneity, the prevalence of which in air-travel choice-behaviour has been highlighted by Hess and Polak (2004). Unlike many previous studies, our analysis uses highly disaggregate data, by looking at the daily frequencies for each airline on every route, and by using detailed origin-destination matrices for the ground-level access dimension. The study also investigates the use of non-linear specifications of explanatory variables such as flight frequency and access-time. Finally, the study aims to determine whether there are differences across groups of travellers in the relative importance assigned to the three separate choice-dimensions mentioned above.

In common with most previous studies, our analysis looks only at departing passengers, due to the lack of data on arriving and connecting passengers. However, by including resident as well as visiting passengers in the analysis, the models indirectly also look at the choice of arrival airport, given that for the latter group of travellers, data is collected at the return leg stage, for which the departure airport is in fact the arrival airport from the outbound flight (excluding the possibility of an open-jaw ticket). Special care however needs to be taken with the use of non-resident passengers, especially so in the case of passengers originating from multi-airport regions, where these passengers are not necessarily seen to be making a choice of airport in the study area. Similar issues arise in the case of resident passengers travelling to a multi-airport region. In this case, the choices of departure airport and destination airport are generally closely related, and it is not clear from the outset which of the two choices is more important. The discussion in section 3 shows that careful selection of destinations can however help to ensure that the passengers used do in fact make a specific choice of airport in the study area. Finally, the reason for excluding connecting passengers is that they do not generally make a choice between several connection airports that are all located in the same metropolitan area. As the scope of this analysis is limited to airport choice in a given multi-airport region, rather than choice across different (multi-airport) regions, connecting passengers are thus excluded from the analysis.

The remainder of this paper is organised as follows. In the next section, we present a brief overview of existing work in the area of airport choice. In the third section, we discuss the various datasets used, while in section four, we present the models used in the analysis. The results of the analysis are presented in section five, and are validated in section six. The paper concludes with a discussion of further avenues for research in airport choice modelling.

2. Literature review

In this section, we give a brief review of the existing body of work in the area of airport choice modelling; for other reviews on this topic, see for example Basar and Bhat (2004), Pels et al. (2003) and Hess and Polak (2004).

2.1. Existing work

One of the first major studies of airport choice was conducted by Skinner (1976), who uses a Multinomial Logit (MNL) model for airport choice in the Baltimore-Washington DC area, and identifies flight frequency and ground accessibility as the main determining factors, with travellers being more sensitive to the latter. In a more recent study using an MNL model in this area, Windle and Dresner (1995) repeat the earlier results, and also identify a significant impact of past experience; the more often a traveller has used a certain airport in the past, the more likely he/she is to choose the same airport again.

The SF-bay area has been used in several case studies of airport choice, mainly thanks to the availability of very good data. An early example is that of Harvey (1987), who uses an MNL model, and finds access time and flight frequency to be significant for both leisure and business travellers, with lower values of time for leisure travellers. More recently, Pels et al. (2001) have conducted an analysis in this area using a Nested Logit (NL) model to look at the combined choice of airport and airline. The results indicate that both business and leisure travellers have a nested choice process in which airline choice is nested within the choice of airport (notwithstanding considerations of airline brand loyalty). In a later study, Pels et al. (2003) again make use of the NL model structure, this time in the joint analysis of airport and access-mode choice, revealing high sensitivity to access time, especially for business travellers. In one of the most innovative studies of airport choice, Basar and Bhat (2004) propose the use of a two-level modelling structure in which the actual airport choice process is preceded by a choice-set generation stage, thus acknowledging the fact that some travellers only consider a subset of the available airports. The results show that this *parameterised choice set consideration* (PCMNL) model outperforms the MNL model, and suggest that flight frequency is the most important aspect in choice set composition, while access time is the dominating factor in the actual choice of airport. Finally, Hess and Polak (2004) have recently used the SF-bay area in a study that aims to show the prevalence of random taste heterogeneity in a population of air-travellers; the results show that, while a major part of the variation in tastes can be accounted for through a segmentation of the population, a remaining part of variation, namely with regards to the sensitivity to access time, is purely random.

There have also been a number of studies of airport choice in the United Kingdom. Ashford and Bencheman (1987), who use an MNL model for airport choice at five airports in England, find that access time and flight frequency are significant factors, with flight fares only having an impact for domestic passengers and for international leisure travellers. In a study that is of particular relevance to the present analysis, Ndoh et al. (1990) find that the NL model outperforms the MNL model in a study of passenger route choice in central England. Thompson and Caves (1993) use an MNL model to forecast the market share for a new airport in North England; access time, flight frequency and aircraft-size (comfort) are found to be significant, with access time being most important for travellers living close to the airport and frequency being more important for travellers living further afield. Finally, in an MNL analysis of the distribution of passengers between airports in the Midlands, Brooke et al. (1994) find flight frequency to be most the important factor.

In other studies from around the world, Ozoka and Ashford (1989) use an MNL model to forecast the effects of adding a third airport to a multi-airport region in Nigeria; the results show access-time to be very significant, making the choice of location and the provision of good ground-access facilities important determinants in the planning process. On the quality of service side, Innes and Doucet (1990) use a binary logit model for airport-choice in Canada, and show that travellers prefer jet services to turboprop flights. Furuichi and Koppelman (1994) use an NL model for departure and destination airport choice in Japan, showing significant effects by access time, access journey cost and flight-frequency. Finally, Veldhuis et al. (1999) produce the comprehensive Integrated Airport Competition Model, showing that passenger behaviour is represented most appropriately by a sequential sequential NL choice process that models the choice of main mode followed by the combined choice of airport and air-route, and finally the choice of access-mode at the chosen airport.

2.2. Summary

The review presented in this section has shown that, although there exists a large body of work on the modelling of airport choice in multi-airport regions, most of the studies use fairly basic modelling techniques, with a heavy bias towards the MNL model. However, some research has gone into the use of nesting structures, and important progress has recently been made with the use of methods that acknowledge the effects of choice-set formation, and the use of model structures that allow for the incorporation of random taste variation. This opens up the possibility of combining these three approaches. To do this, a thorough analysis of the choice of nesting structure is however required, and this is a main aim of the present paper.

3. Data

The area chosen for the present study is the SF-Bay area, which is served by three major airports, San Francisco International (SFO) being the busiest, with, in 1995 (the study year), some 15 million emplaned passengers, ahead of Oakland International (OAK), with 7.7 million passengers, and San Jose Municipal (SJC), with 4.2 million passengers. Forecasts by MTC (2000) predict significant increases in traffic; these will inevitably lead to capacity problems, and different expansion schemes are already under consideration (c.f. RAPC, 2000), making the area an ideal candidate for a study of airport choice. In this section, we give a description of the various datasets used in our analysis.

3.1. Air-passenger survey data

Data on passengers' choice behaviour were obtained from the 1995 Airline Passenger Survey conducted by the Metropolitan Transport Commission (MTC) in August and October 1995. This contained information on over 21,000 departing air-travellers. For a detailed description of the survey, see MTC (1995). Passenger interviews were conducted at the three main SF-Bay area airports, as well as at the minor Sonoma County airport (STS), which was not included in the present study. The number of passengers interviewed at the three main airports is not entirely representative of the real-world traffic at the airports; indeed, SJC is over-sampled, while OAK is under-sampled. This needs to be taken into account in the interpretation of the modelling results, but surprisingly, previous studies using this dataset have generally failed to highlight this issue (a notable exception being Basar and Bhat, 2004, who use a Weighted Exogenous Sample Maximum Likelihood approach). The weighting approach used in the present analysis is described in the section 4.

The data selection process is based on that used by Hess and Polak (2004). It was decided to use only destinations that could be reached by direct flight from all three airports, on every day of the week (at the time of the survey), leading to 14 destinations, and a sample of 9,924 respondents (after initial data cleaning). This contained some 3,246 travellers who indicated that they could not have flown out of a different airport. Possible reasons for this include unavailability (at the time of booking) for flights from other airports on the chosen flight date and time (especially likely for travellers with inflexible timing), misinformation of the traveller, or an *a priori* decision not to consider any of the other airports. Hess and Polak (2004) show that the inclusion of these travellers leads to biased results, leading to the decision to exclude these observations from the analysis. In a way, this acts as an approximation to a model that incorporates choice-set generation. From the resulting sample of 6,678 travellers, a further 1,587 passengers needed to be excluded, for four reasons. To start with, the sample contained some 111 passengers whose chosen access-mode was a hotel-courtesy-shuttle; as it was not possible to unambiguously define the availability of this mode for all passengers, it was decided to exclude these observations. At this point, five main purposes were identified in the dataset; business, holiday, visiting friends and relatives, extraordinary events, and others. As the extraordinary events group contained a wide range of trips (from funerals to weddings), it was decided that it would be wrong to simply include them in a wider leisure group, as has been done previously. Indeed, the short-term planning and/or precise timing associated with some of these events make the decision-process more similar to that of a business trip. As there are however arguably also important differences between this type of trip and a business trip, and as it was not possible to fit a sensible separate model for these 299 trips, it was decided to exclude them from the analysis. An identical decision was taken for the 360 trips that had some other purpose. A further 817 observations had to be excluded either due to the absence of information on household income (required to allow for a segmentation by income) or because of a lack of corresponding air-transport or ground-transportation level-of-service data for specific journeys. The final sample thus contained 5,091 observations, with flights to 14 destinations.

The data used are summarised in table 1, which clearly shows the oversampling of SJC. The specific choice of destinations had little effect on the distribution of observations across other dimensions, such as journey purposes and household income. Clearly, the sampling has an effect on the market shares for the different airlines; this was taken into account in the calculation of weights described in section 4. The resulting dataset was split into two parts, a dataset used in the actual analysis (4,582 observations), and a 10% sample retained for later validation of the models (509 observations).

Table 1. Destination used in the analysis

		Destination airport														
		BURBANK, CA	DENVER, CO	DALLAS, FT. WORTH, TX	LAS VEGAS, NV	LOS ANGELES, CA	ONTARIO, CA	CHICAGO, O'HARE, IL	PORTLAND, OR	PHOENIX, AZ	RENO, NV	SAN DIEGO, CA	SEATTLE, WA	SALT LAKE CITY, UT	Orange County, CA	Total
Departure Airport	SFO	55	65	36	57	199	35	89	140	128	1	258	213	42	37	1355
	SJC	167	71	91	163	367	111	58	106	133	156	248	169	61	247	2148
	OAK	211	9	25	68	381	135	1	101	51	39	139	208	43	177	1588
Total		433	145	152	288	947	281	148	347	312	196	645	590	146	461	5091

Special care is required in the case of destinations that are themselves located in multi-airport regions. Destinations from two such multi-airport regions were included in the present analysis, namely destinations in the wider Los Angeles (LA) area, as well as Chicago's O'Hare (ORD) airport. The decision to include airports from the LA area was motivated by the frequency of these destinations in the survey data. It is important to establish whether passengers are likely to make a choice of airport in the San Francisco area, besides the choice made in the LA area, especially so for passengers whose return journey started in the LA area. During the period of observation, daily flights were available between each of the three SF-Bay area airports and each of the five airports in the wider LA region. As there was relatively high frequency on all routes, passengers can be expected to make a specific choice of airport in the SF-Bay area, independently of the choice of airport in the LA area (where this choice may take precedence, especially for visiting passengers). This assumption is further supported by the fact that the differences in level-of-service attributes between the three SF-Bay area airports are similar across the five destinations in the LA area, leading to similar destination-specific choice-sets across the five airports in the Los Angeles area. The inclusion of ORD on the other hand was motivated by the comparatively very low frequency of direct flights to Chicago's alternative airport; Midway (MDW). A comparison of the results produced in two small-scale separate analyses that included, respectively excluded these destinations, revealed no major differences, suggesting that the inclusion of these airports has no ill effects on the modelling analysis. For the remaining destinations used in the analysis, there should be no explicit choice of airport in the destination area. This also makes it more likely that the non-resident passengers sampled on flights to these destinations made a conscious choice of a specific airport in the SF-Bay area on their outbound flight, where this choice is influenced only by the characteristics of the SF-Bay system and the air-travel level-of-service variables.

3.2. Air-travel level-of-service data

For the present analysis, air-travel level-of-service data were obtained from *BACK Aviation Solutions*[1]. The dataset contains daily information for each operator serving the selected routes in August and October 1995, thus making the data more detailed than that of many previous studies that have relied on the use of weekly or even monthly data. Eight airlines were used in the analysis, and these are hereafter referred to as airline $\mathbf{A_1}$ to airline $\mathbf{A_8}$ Besides the frequencies for the different operators, the dataset contains information on flight times and the type of aircraft used. Finally, information is available on the average fares paid on a given route operated by a given airline. This clearly involves a great deal of aggregation, as no differences are made between the fares for the different classes of travel, and some information is lost as no data is available on the availability of different ticket classes at the time of booking, essentially leading to an assumption of similar selling speed on all routes. Unfortunately, this assumption cannot be avoided, given the quality of the data. The availability of given fare-classes can be modelled in the presence of adequate data on the distribution of fares across classes (c.f Battersby 2003); this is beyond the scope of the current analysis, the inclusion of such an approach into a wider framework for air-travel related choice modelling is however an important avenue for further research. Finally, the dataset was complemented by information on the on-time performance of the different airlines used in the analysis, and the overall on-time performance of airlines at the three airports[2].

3.3. Ground-access level-of-service data

As was the case for the air-transport level-of-service data, the information on the chosen access mode contained in the passenger survey needs to be complemented by data on the unchosen access options at the chosen airport as well as at the different unchosen airports. For the present analysis, ground-access level-of-service information was obtained from the MTC in the form of origin-destination travel time and cost matrices for the 1099 travel area zones (TAZ) used for the SF-Bay area. Some information is clearly lost due to the aggregation into travel zones, this

1 Back Aviation Solutions, 6000 Lake Forrest Drive, Suite 580, Atlanta, GA 30328, www.backaviation.com
2 Available from the Bureau of Transport Statistics, via www.bts.gov/programs/oai/airline_ontime_statistics

is however unavoidable due to the high number of possible ground-level origins, and the effects should be minimal, given the small differences between the origins situated in a given TAZ.

The dataset contains information on travel distance, travel time and tolls for car travel, under peak and off-peak conditions, and for varying car-occupancy (which has an effect on tolls). Similarly, the dataset contains information on access time, wait time, travel time, egress time and fares for public transport journeys. Corresponding data for other modes, such as taxi, limousine and special airport bus services were calculated separately, based on current prices and the changes in the Consumer Price Index for California from August and October 1995 to September 2003. A complication arose with regards to rental car, as it is difficult to judge how the cost of renting a car is taken into account in the decision-making process (e.g. covered by employer or by traveller). Similarly, it is not clear whether the marginal costs of trips by private car (e.g. petrol) are actually taken into account by travellers (when compared to the direct cost of a road or bridge-toll). A further complication arises with regards to the cost of parking at the airport, given that no information is available as to whether a given traveller makes use of the airport parking facilities (as opposed to using *kiss-and-fly*). Attempts to include parking cost and rental cost in the models led to inconsistent results; it was thus decided to exclude these costs from the model, allowing us to merge private car and rental car into a generic car mode, where the only cost is that of any toll incurred. This led to six remaining access-modes; car, public transport (transit), scheduled airport bus services, door-to-door services, taxi and limousine. It was assumed that taxi and limousine services are available for each origin, while the availability of door-to-door and scheduled services depends on the distance to the airports. The availability of public transport was obtained from the MTC OD matrices, and, in the absence of any information on the availability of the car mode, it had to be assumed that car is always available.

3.4. Data assembly and choice-set construction

The final sample contains data on 3 departure airports, 8 airlines, and 6 access-modes, leading to 144 distinct triplets of alternatives. Given the three-dimensional choice set, any given alternative shares the attributes of 73 other alternatives along a single dimension of choice, and shares the attributes of 14 alternatives along two such dimensions. In fact, it can be seen that the 144 triplets are made up of a set of 42 distinct alternatives (8 airlines and 6 access-modes for each of the 3 airports). For each observation, data on the attributes and availability of these 42 elementary alternatives was appended to the survey data. After adding in airport-specific attributes, the combination into triplets of alternatives was performed via the specification of utilities. The attributes and availability of the access-modes depend on the ground-level origin of a traveller, while the attributes and availability of the different airline options depend on the choice of destination (where not every airline operates from each airport to all 14 destinations used). The days of week were taken into account in the definition of the attributes and availability of the different flight options, as was the season (August or October), while peak and off-peak aspects were taken into account for the access-journey attributes. The dataset now contains information on the chosen and unchosen alternatives for each of the 5,091 observations used, and the data are ready for use in the modelling analysis.

4. Modelling methodology

4.1. Discrete choice models

Discrete choice models predict the choice of a decision-maker in a given choice situation as a function of the utilities of the alternatives available to the decision-maker. The utility that a decision-maker obtains from choosing an alternative is calculated as a function of the observed attributes of this alternative and the tastes of this decision-maker together with a random component, which accounts for unobserved factors influencing choice.

Under the assumption of utility maximisation, different distributional assumptions on the random component of utility lead to different models. In the MNL model, the unobserved utility terms are assumed to be distributed *IID type I extreme value*. With this choice of distribution, the probability of decision-maker *n* choosing alternative *i* is given by:

$$P_{ni} = \frac{e^{V_{ni}}}{\sum_{j=1}^{I} e^{V_{nj}}} \quad , \qquad \ldots(1)$$

where I gives the total number of alternatives in the choice-set and V_{ni} is the observable component of utility, which is generally linear in tastes (β_n) and attributes (z_{ni}) i.e., $V_{ni}=\beta_n{}'z_{ni}$.

A key limitation of the MNL model is that the pattern of substitution between any two alternatives is the same across alternatives, making it impossible to accommodate heightened correlation between alternatives that are closer substitutes for each other. Although the MNL model has been used repeatedly in airport choice modelling, and several authors (e.g. Ashford and Bencheman, 1987 and Thompson and Caves, 1993) have justified the use of the MNL model by claiming that airports are independent entities, this is in fact far from clear. Indeed it seems highly likely that in a multi-airport region there will be varying cross-elasticities between pairs of airports in, reflecting structural similarities amongst them.

The Nested Logit (NL) model generalizes the MNL by dividing the choice set into nests of alternatives, with increased correlation, and thus higher cross-elasticities, between alternatives sharing a nest. The NL is a member of the Generalized Extreme Value (GEV) family of discrete choice models introduced by McFadden (1978). In the NL model, nests are hierarchical and mutually exclusive. The choice probabilities are represented through a product of successive logit-style choice probabilities that represent a chain from the root of the tree to the elementary alternative for which the probability is calculated. The utility of composite alternatives is determined by the utilities of the nodes situated directly below them in the tree; this way, the utility of elementary alternatives is propagated up through the tree, with the help of a logsum term. A structural (logsum) parameter is associated with each node, determining the correlation of the alternatives within the different nests. For example, the choice probability of an alternative *i* contained in nest *m* of a two-level NL model is given by:

$$P_i = P_{C_m} \cdot P_{i|C_m} = \frac{e^{\lambda_m I_m}}{\sum_{l=1}^{M} e^{\lambda_l I_l}} \cdot \frac{e^{\frac{V_i}{\lambda_m}}}{\sum_{j \in C_m} e^{\frac{V_j}{\lambda_m}}} \quad , \qquad \ldots(2)$$

with logsum term

$$I_m = \ln \sum_{j \in C_m} e^{\frac{V_j}{\lambda_m}} \quad ,$$

where is the logsum parameter associated with nest m. λ_m

The value of the different structural parametersis generally constrained to lie between 0 and 1, where values of below 0 are inconsistent with utility maximisation while values above 1 are only consistent with utility maximization for special ranges of the explanatory variables. The parameters measure the degree of independence between alternatives in the respective nest, with higher meaning more independence and hence less correlation between the unobserved components of utility of the alternatives contained in the nest. A value of 1 for all structural parameters leads to the MNL model. $\lambda_{m}\lambda_{m}\lambda_{m}\lambda_{m}$

4.2. Sampling weights

Aircraft occupancy data was used to calculate the total traffic on the different routes used in the analysis, for each of the carriers. From this, relative weights were assigned to each airport-airline pair. A similar process was used to calculate corresponding weights for the sample data used in the present analysis. The individual pairs of weights were then used to calculate multiplicative weights that could be used in the analysis, where the weight for a given airport-airline pair was given by dividing the actual population weight by the sample weight for this pair. This process was repeated for each observation used in the analysis, and it is easy to see that the resulting sum of weights over the total number of observations is equal to the number of observations used. In the estimation process, each term in the log-likelihood function was then multiplied by the appropriate weight for the associated chosen alternative.

4.3. Segmentation by purpose and residency status

An important question arises with respect to how to acknowledge the differences that exist between residents and visitors, and between travellers with different trip purposes. Results by Hess and Polak (2004) on the same data show that there exist important differences along both dimensions, with the differences across trip purposes seemingly being more important than those between residents and visitors. After a decision has been taken to segment along either or both of these dimensions, a further decision must be made whether to use separate models for the different groups of travellers or whether to use separate coefficients for each group of traveller within the same model, where the latter approach essentially assumes that any differences are restricted to a subset of the parameter space (as not all coefficients will be made group-specific).

Several tests were conducted to investigate the benefits of the different approaches, using basic MNL models with a simple specification of utility. Likelihood-ratio tests were used to assess the impact of using separate models as compared to using a common model. The tests showed significant differences between residents and visitors $(122.235 \sim \chi^2_{16})$ as well as between business trips and the combined group of holiday trips and visiting friends and relatives (VFR) trips $(241.5688 \sim \chi^2_{21})$. In the existing literature, holiday trips are generally combined with VFR trips to form a wider group of leisure trips. Several tests were conducted to assess the adequateness of such a grouping. One of these tests fits separate models for holiday trips by residents and by visitors, and separate models for VFR trips by residents and visitors, along with a common leisure model for residents and a common leisure model for visitors. In the case of resident travellers, the test reveals no significant differences between holiday trips and VFR trips $(18.629 \sim \chi^2_{20})$. However, in the case of visitors, the differences are very significant $(105.183 \sim \chi^2_{20})$. This thus not only suggests that separate models should be used, at least in the case of visitors, but also reiterates the findings that there are important differences between residents and visitors.

Given the results produced by the tests, it was decided to avoid using common models for different purposes and for residents and visitors, as it had become clear that, in many cases, not all of the differences could be explained by the use of a few separate coefficients. It was thus decided to use six separate groups of travellers, dividing the population into residents and visitors, and dividing the trips into business, holiday and VFR trips.

5. Modelling analysis

In this section, we describe the results of the modelling analysis. This is divided into three main parts. We first present a discussion of the various specifications of utility used in the analysis. We then describe the results from the MNL models, and finally summarise the findings from the analysis that uses NL models.

5.1. Utility functions

Overall, the final specifications developed for the various models are very similar, although there are some differences, notably in the inclusion of fare and cost coefficients, and in the segmentation of travellers by income. For every model, attempts were made to include coefficients showing travellers' sensitivity to various attributes of the airports, airlines and access-modes. These included factors such as flight frequency, flight time (block time, which indirectly takes into account airport congestion) and fare, as well as access-time (in-vehicle), walk-time to access-mode (e.g. to public transport station), wait-time for access-mode and access cost. We also explored the influence of aircraft type (jet *vs* turboprop) and both linear and various non-linear specifications of the various explanatory variables were tested. The best results were obtained with the use of a logarithmic transform, this however only led to an improvement in model fit when applied to flight frequency, whereas non-linear specifications of flight time, in-vehicle time, access walk time, wait time and fare led to unsatisfactory results. Also, some potentially important influences, such as carrier loyalty, could not be explored, due to lack of data (e.g. no information on frequent flyer programmes). Similarly, it was not possible to identify a direct affect of the on-time-performance of airlines or airports on the respective choice probabilities. Attempts were also made to segment the population by income, for example in order to show different values of time in different income-classes. Three income groups were defined, segmenting the population into low-income (<$21,000 p.a., accounting for around 40% of the sample population), medium-income (between $21,000 and $44,000 p.a., accounting for around 35% of the sample population) and high-income (above $44,000 p.a., accounting for the remaining 25% of the sample population).

A further specification issue that was explored was the influence of past experience on choice behaviour (c.f. Windle and Dresner, 1995). In the present analysis, we had information on the number of flights a given traveller took from each of the three SF-bay airports in the past twelve months. It might be expected that the effect of past experience is non-linear, such that most of the experience is gained in the first few flights. To this extent, a logarithmic transform was used on these factors. For the three airports, a coefficient was estimated that shows the effect of past experience at the given airport on that airport's utility. It should be recognised that there is potentially also a cross-effect of experience. Hence, coefficients were also included that show the effect of experience at one airport on another airport's utility. Clearly, some normalisation is required in this case, as the same variable cannot be included in each utility function for reasons of identification. Aside from three airport-specific experience coefficients, coefficients associated with past experience at SJC and OAK were thus included in the utility of SFO, while a coefficient of experience at SFO was included in the utility of SJC, and no cross-coefficients were included in the utility of OAK-alternatives. The inclusion of these variables did in each case, as expected, lead to a dramatic improvement in log-likelihood. As an example, in the MNL model for resident business travellers, the inclusion of these 5 coefficients led to an increase in log-likelihood by 219.583 units, which is clearly highly significant. It should of course be noted that the inclusion of these coefficients could lead to problems with endogeneity, as the values of the past choice indicators may be closely correlated with the other explanatory variables and with unobservables. This would make this approach inapplicable in the case where the model was used for forecasting. However, this is not the main purpose of the present analysis; furthermore, in each one of the models used, the values of the remaining coefficients remained largely unaffected, suggesting that the inclusion of these experience terms did not introduce major bias.

5.2. MNL models

In the following six paragraphs, we describe the findings of the analysis fitting MNL models to the six separate estimation datasets. The results of the various models are summarised in table 2 for residents and table 3 for visitors.

5.2.1. Business trips by residents

The estimation dataset contains information on 1,098 business trips by residents. The estimation process revealed significant effects of walk-access time, access-cost, in-vehicle access-time, flight-time and frequency. Also, a negative impact on utility is associated with turboprop planes. The initial estimation revealed an effect of fare, however, this effect was of the wrong sign (positive) for medium and high-income traveller, while the effect for low-income travellers was negative, but not significant. As these results are counterintuitive, it was decided to drop these coefficients from the model. The fact that no significant negative effect of fare could be identified can be partly explained by the poor quality of the (highly aggregate) fare data, but could also signal sheer indifference to fare increases by business travellers.

It was possible to segment the sensitivity to walk time and access cost by income, although, given very low differences between the estimates in the low and medium income group, only two coefficients were retained, one for people earning less than $44,000 per annum, and one for the remaining travellers. The results show lower sensitivity to cost for people with higher income, along with higher sensitivity to increases in walk time. In terms of past experience, the estimates show positive direct effects of past experience for all three airports, with positive cross-effects of experience at SJC and OAK on the utility of SFO, and a positive (but not significant) cross-effect of experience at SFO on the utility of SJC. Finally, increases in flight frequency lead to increases in utility, where the logarithmic transform ensures decreasing marginal returns.

5.2.2. Business trips by visitors

The estimation dataset contains information on 1,057 business trips by visitors. Just as for resident business travellers, the initial modelling estimates showed a positive (but insignificant) effect of fare for high and medium-income business travellers, while the effect for low-income travellers was negative, but not significant. Again, fare was thus excluded from the models. In-vehicle access -time and access-cost are again significant, and negative, with increasing sensitivity to in-vehicle access-time with higher income (only two groups could be used) and lower sensitivity to cost with higher income (two groups only). Whereas it was not possible to estimate a significant effect of wait-time for resident business travellers, a significant negative effect could be identified for their non-resident counterparts. However, the estimate for flight-time no longer significant (but still negative), and it was not possible to include an effect of equipment type, as flights using turboprop planes were never chosen. Also, with this model, no effect could be associated with access walk-time. Frequency again has a positive effect, but, unlike for access-time and access-cost, it was not possible to identify significantly different estimates for the different income groups. Finally, unlike in the model for resident business travellers, experience at OAK has a negative effect on the utility of SFO. Also, the cross-effect of experience at SJC on the utility of SFO is insignificant.

5.2.3. Holiday trips by residents

The model estimated on the 831 observations for residents' holiday trips suggests a lower utility for flights using turboprop aircraft, negative impacts by access-cost and in-vehicle time, a positive effect of flight frequency, positive direct effects of past experience as well as positive cross-effects of past experience at SJC and OAK on the utility of SFO, and of past experience at SFO on the utility of SJC (not significant). Finally, for this group of travellers, a negative effect could be identified for fare, although the t-value is just below the 95% limit, while no effect could be associated with flight-time and access walk-time. No significant gains could be made through segmenting the population by income for any of the coefficients.

Table 2. MNL results for residents (selected coefficients)

		Business		Holiday		VFR	
		estimate	t-test	estimate	t-test	estimate	t-test
	Access cost			-0.02077	-2.21	-0.02232	-2.29
	Access cost, >$44,000 p.a.	-0.02436	-2.86				
	Access cost, <$44,000 p.a.	-0.03575	-4.17				
	Access in-vehicle time	-0.05215	-12.13	-0.05935	-12.94	-0.04903	-9.43
	Walk-time, >$44,000 p.a.	-0.15311	-2.97				
	Walk-time, <$44,000 p.a.	-0.11391	-2.47				
	Fare			-0.01307	-1.90	-0.02673	-3.03
	Flight time	-0.04707	-2.37				
	Flight frequency	1.31825	10.77	1.32348	9.35	1.44471	7.87
	Turboprop	-2.52961	-3.20	-4.22941	-2.70		
Effect of past experience	OAK on OAK	1.99926	9.44	2.10237	5.09	2.29190	5.24
Effect of past experience	SFO on SFO	1.18295	9.62	1.18871	7.89	2.04880	8.83
Effect of past experience	SJC on SJC	1.96406	8.49	2.59094	5.04	3.16900	5.87
Effect of past experience	OAK on SFO	0.66199	3.37	0.83278	1.98	0.44133	1.02
Effect of past experience	SJC on SFO	0.78451	3.68	1.43022	2.71	0.55743	1.10
Effect of past experience	SFO on SJC	0.17309	1.07	0.16181	0.79	0.02924	0.09
	Observations	1098		831		641	
	Log-likelihood	-1551.6171		-1384.8137		-1050.8398	
	ρ^2	0.5934		0.5198		0.5157	

5.2.4. Holiday trips by visitors

For the 534 visitors on holiday trips, no significant effect of fare could be identified, and the effect of access-cost, although of the correct sign, is not significant at the 95% level. In-vehicle time has a significant negative effect, as has flight time, while increases in frequency lead to increases in utility. Past experience at SJC has a positive impact on the utility of SFO, while the other two cross-effect estimates are not significant. Finally, the aircraft-type coefficient had to be excluded from the model (never chosen), while no effect could be identified for wait-time, and segmentations by income did not lead to any gains in model fit.

5.2.5. VFR trips by residents

The estimates for the model fitted to the sample of 641 residents on VFR trips show significant negative effects of access-cost, in-vehicle time and flight fare, along with positive effects of flight frequency. The cross-effect estimates of past experience are not significant, equipment-size could not be included and no effects could be identified for walk-time, wait-time and flight time, while segmentations by income led to a loss of information in the model.

Table 3. MNL results for visitors (selected coefficients)

		Business		Holiday		VFR	
		estimate	t-test	estimate	t-test	estimate	t-test
	Access cost			-0.01446	-1.66		
	Access cost, >$44,000 p.a.	-0.02186	-2.55				
	Access cost, <$44,000 p.a.	-0.02862	-3.94				
	Access in-vehicle time			-0.07694	-13.22	-0.06980	-11.06
	In-vehicle time, >$22,000 p.a.	-0.08201	-14.43				
	In-vehicle time, <$22,000 p.a.	-0.04964	-7.18				
	Wait time	-0.25067	-3.28				
	Fare, <$21,000 p.a.					-0.05009	-3.55
	Fare, [$21,000,$44,000] p.a.					-0.02671	-1.95
	Flight time	-0.02925	-1.39	-0.09080	-3.42	-0.15216	-5.12
	Flight frequency	1.30655	11.34	1.07830	7.51	0.72443	4.41
Effect of past experience	OAK on OAK	1.18806	6.57	1.25285	2.90	1.38995	2.96
	SFO on SFO	1.93240	9.39	0.75140	3.97	1.09905	3.35
	SJC on SJC	1.39732	6.10	2.05636	4.42	2.25691	4.17
	OAK on SFO	-0.71724	-3.36	-0.47413	-0.99	0.18871	0.35
	SJC on SFO	0.00748	0.03	0.83176	1.86	-0.12191	-0.17
	SFO on SJC	0.50315	2.38	-0.10838	-0.34	0.18095	0.42
	Observations	1057		534		421	
	Log-likelihood	-1517.6772		-1018.2452		-621.805	
	ρ^2	0.4477		0.387		0.5236	

5.2.6. VRF trips by visitors

The final subsample used in the estimation of MNL models contains information on 421 VFR trips by visitors. The results show negative impacts of fare in the medium and low-income classes (with higher sensitivity in the low-income class), while the effect for high-earners was insignificant and was dropped from the model. In-vehicle time and flight-time have a negative effect, with a positive effect for frequency increases. Again, the cross-effect estimates of past experience are insignificant, while no effect could be associated with access walk-time, wait-time, and access-cost, and the turboprop coefficient had to be excluded.

5.2.7. Comparison

The discussions in sections 5.2.1 to 5.2.6 have revealed that there are important differences across models (and thus data sub-samples) in the optimal specification of utility. The common point across models is that a logarithmic specification of frequency and past experience is always preferable to a linear specification. Significant effects of flight fare could only be identified for resident holiday and VFR travellers, as well as for visiting VFR travellers, where there are also differences across income groups in fare-sensitivity. In terms of model fit (ρ^2), the models for residents perform better than those for visitors for business and holiday trips, while the opposite is the case for VFR trips. Finally, it is of interest to compare the substantive results across models. Given the potential differences in scale, such comparisons should only be made in the form of ratios in parameters. As fare is only used in three of the models, it was decided to give preference to the trade-off between flight frequency and in-vehicle time (over the trade-off between fare and frequency). The coefficient estimated for in-vehicle time β_{AT} gives the marginal change in utility resulting from an increase in in-vehicle time by one minute. The corresponding estimate for flight frequency gives the return of an increase in the logarithm of frequency by one unit, such that, with a base frequency of f flights, and coefficient estimate β_{FT}, the return of an increase by one flight is equal to $\beta_{FT}(ln(f+1)-ln(f))$. The trade-off between increases in flight frequency and increases in access in-vehicle time is thus given by $\beta_{FT}(ln(f+1)-ln(f)) / \beta_{AT}$. The results show a higher willingness to accept increases in access-time for residents (values of β_{FT}/β_{AT} equal to 25.28, 22.3 and 29.47 minutes per additional flight for business, holiday and VFR trips respectively) than for visitors (values of 15.93 and 26.32 minutes respectively for high and low-income business travellers, and 14.01 and 10.38 minutes respectively for holiday and VFR trips). The differences are especially significant in the case of VFR trips, where the relative value of frequency increases is at its highest for residents, while it is at its lowest for visitors.

5.3. NL models

After the estimation of the MNL models described in section 5.2, the next aim of the analysis was to find appropriate specifications for NL models. Several important issues arise in this case. The analysis looks at the combined choice of airport, airline and access-mode. While heightened correlation is generally expected between the different flight options at a given airport, it must equally well be assumed that there is heightened correlation between the different flights operated by a given carrier, and also between two alternatives sharing the same access mode. As such, there is potentially a need to nest by airport, airline, and access-mode. However, a four-level NL model (root, plus three additional levels of nesting) would not be appropriate as the lower level of nesting would be obsolete, given that each nest would contain just a single elementary alternative (e.g. after the choice of airport and airline, there is only one remaining alternative for each access-mode). This thus means that at best, a three-level structure can be used, discarding one of the three possible nesting levels. This leads to six possible tree structures, when one notes that a tree structure with airport above airline is not equivalent to a tree structure with airline above airport. The use of each of these six three-level structures was attempted, however, none of them led to satisfactory results. This suggests that a multi-level structure is not applicable for the modelling of airport choice, at least with the current data. A two-level structure would thus have to be used, nesting either by airport, or airline, or access-mode. In this section, we describe the results obtained with each of these approaches. Due to space constraints, only a very limited part of the results is reproduced here; the optimal utility function specifications of the various models were however essentially identical to those of the corresponding MNL models (as were the substantive results in terms of ratios of coefficients), although the use of a nesting structure occasionally led to a drop in significance of individual coefficients.

5.3.1 Nesting by airport

The first set of models nest the elementary alternatives by airport, leading to 48 alternatives per nest (8 airlines and 6 access-modes). The results are summarised in table 4, with t-statistics for the structural parameters given in brackets. For comparison, the table again gives the final log-likelihood of the corresponding MNL models. The results show that, for every single model, the structural parameter of the nest containing the SFO alternatives had to be constrained to a value of 1, as it would otherwise have exceeded this value, becoming inconsistent with utility maximisation. This suggests that there is no heightened correlation between the different alternatives available from SFO. Passengers are not more likely to shift to another alternative at SFO than they are to shift to an alternative at another airport. This can at least partly be explained by the overall poor on-time performance of SFO.

Except for the case of visitors on VFR trips, where the structural parameter for OAK had to be constrained to 1, the estimates for the structural parameters of the other two airports are always below 1. There are differences across models in the values of the structural parameters, and also in the relative values of the structural parameters for the SJC and OAK nests (although λ_{SJC} is generally lower than λ_{OAK}), suggesting important differences between the different groups of travellers. In terms of model fit, the use of the NL models leads to a significant increase in log-likelihood, except in the case of visitors on VFR trips, where the log-likelihood is virtually identical to that of the MNL model, as is the NL model itself, given that the SFO and OAK structural parameters are equal to 1, while the structural parameter for SJC is very close to 1. Except for VFR trips, the improvements in model fit are more important for visitors than for residents, and the lower structural parameters for visitors on business and holiday (only for SJC) trips suggest a lower substitution effect between airports than is the case for residents.

5.3.2 Nesting by airline

The lack of information on frequent-flier programme membership means that the models are largely unable to pick up the full effect of the correlation between different alternatives that refer to the same airline. Nevertheless, it is of interest to attempt to use a nesting structure that uses a single nest for each airline, leading to 8 nests, with 18 alternatives each. The results of this analysis are summarised in table 5.

The effects of the lack of information relating to airline-allegiance become visible in the high number of structural parameters that had to be constrained to a value of 1. Nevertheless, except for the model for visitor VFR trips, the use the NL model resulted in a significant increase in log-likelihood over the MNL model. Also, the great variability in the values of the structural parameters for given airlines across the different models suggests significant differences in the cross-elasticities in the different models. The exact analysis of these cross-elasticities is beyond the scope of the present paper (given the very high number of elementary alternatives); however, the results in table 5 could suggest that, despite the lack of adequate data on allegiances to airlines, the models are able to pick up some effect of correlation between alternatives associated with given airlines. Also, a brief analysis revealed some correlation between the nesting parameters associated with a given airline and the on-time performance of that airline. Indeed, airline $\mathbf{A_7}$, for which the structural parameter had to be constrained to a value of 1 in each model, was plagued by very poor on-time performance in 1995, as were airlines $\mathbf{A_1}$ and $\mathbf{A_3}$, for which the structural parameters are also generally closer to 1; passenger allegiance to such airlines is clearly affected in a negative way. Finally, airlines $\mathbf{A_5}$ and $\mathbf{A_8}$ on average have lower structural parameters than the other airlines. This could at least be partly be related to the fact that these two carriers run a budget-airline scheme; travellers on such airlines are generally far more likely to shift to another service operated by this same or another low-cost carrier than to shift to another airline. It could also signal that these travellers make a decision to travel by air solely because of the low cost offered by these carriers, and are thus very unlikely to consider flying on another airline.

Table 4. NL results for nesting by airport

	Business		Holiday		VFR	
	Resident	Visitor	Resident	Visitor	Resident	Visitor
MNL LL	-1551.6171	-1517.6772	-1384.8137	-1018.2452	-1050.8398	-621.805
NL LL	-1545.1353	-1487.7122	-1372.1879	-999.5086	-1039.6656	-621.6214
NL ρ^2	0.5951	0.4586	0.5242	0.3983	0.5208	0.5237
λ_{SFO}	1.00	1.00	1.00	1.00	1.00	1.00
λ_{SJC}	0.7829 (14.5)	0.5259 (11.8)	0.7627 (13.1)	0.4399 (6.9)	0.6708 (11.2)	0.9333 (8.8)
λ_{OAK}	0.8925 (13.6)	0.7178 (9.4)	0.7258 (12.2)	0.7373 (6.3)	0.7828 (10.8)	1.00

5.3.3 Nesting by access-mode

In many regards, the use of nesting by access-mode is the most promising approach. Indeed, in this case, there are no important problems with missing data such as in the case of the nesting by airline. Furthermore, given the heavy bias towards car in the access-journeys to the different SF-bay area airports (over 80% in the present sample), it can be expected that the nesting parameter for this nest especially should be very low, to reflect the higher substitution effects between alternatives sharing the car mode, and the low cross-elasticities between the car mode and alternative access-modes. The results of this analysis are summarised in table 6.

Except for the model for business trips by visitors (for whom the car and rental car market shares are generally lower than for other groups), the structural parameter for car is always very low, illustrating travellers' *allegiance* to the car mode; they generally seem to be more willing to accept a change of airline or airport than to accept a change of access-mode. A comparable constant low structural parameter is observed for the taxi nest, while the structural parameter for the scheduled nest especially varies widely across models. Unlike in the models using nesting by airport and airline, the present nesting approach leads to universal significant increases in log-likelihood, including the model for VFR trips by visitors. Also, in total, only three of the structural parameters had to be constrained to a value of 1. Nevertheless, it should be noted that some of the structural parameters reported in table 6 are insignificant. The exclusion of these parameters however either led to a significant drop in log-likelihood (in the case of the model for holiday trips by residents and visitors) or did not lead to significantly changed values of the other structural parameters and coefficients (in the case of VFR trips). Finally, it should be noted that, for holiday trips by visitors, the structural parameters of the car, door-to-door and taxi nests were constrained to have the same value, given that the initial estimates were almost indistinguishable. This led to a drop in the log-likelihood by a mere 0.028 points. Overall, the results from this section suggest that important gains can be made by using a structure that nests alternatives by access mode. This reinforces the belief that access-related factors often play a crucial role in the choice of airport.

Table 5. NL results for nesting by airline

	Business		Holiday		VFR	
	Resident	Visitor	Resident	Visitor	Resident	Visitor
MNL LL	-1551.6171	-1517.6772	-1384.8137	-1018.2452	-1050.8398	-621.805
NL LL	-1536.6592	-1507.6170	-1371.2103	-1003.9270	-1034.0716	-6020.2399
NL ρ^2	0.5974	0.4514	0.5245	0.3956	0.5234	0.5248
λ_{A1}	0.9499 (4.7)	0.9617 (3.5)	0.9237 (3.9)	0.6989 (3.1)	1.00	1.00
λ_{A2}	0.6108 (7.2)	0.9822 (8.7)	0.7841 (3.8)	0.6249 (7.7)	0.8663 (9.5)	0.8606 (7.2)
λ_{A3}	1.00	0.8895 (2.9)	1.00	0.7697 (3.9)	0.8617 (2.7)	0.8549 (3.6)
λ_{A4}	1.00	0.6538 (4.2)	1.00	0.7237 (2.8)	1.00	0.6762 (2.6)
λ_{A5}	0.7433 (9.7)	0.6317 (3.8)	0.7379 (7.5)	0.3917 (3.2)	0.6344 (6.8)	1.00
λ_{A6}	1.00	1.00	0.9967 (10.2)	0.6761 (5.1)	1.00	0.7935 (8.2)
λ_{A7}	1.00	1.00	1.00	1.00	1.00	1.00
λ_{A8}	0.8389 (4.7)	0.7921 (4.3)	0.7240 (8.6)	0.5298 (7.9)	0.6664 (2.7)	0.8399 (3.7)

Table 6. . NL results for nesting by access-mode

	Business		Holiday		VFR	
	Resident	Visitor	Resident	Visitor	Resident	Visitor
MNL LL	-1551.6171	-1517.6772	-1384.8137	-1018.2452	-1050.8398	-621.805
NL LL	-1520.4187	-1508.7863	-1351.1774	-1004.2646	-1007.2021	-603.072
NL ρ^2	0.6016	0.4510	0.5315	0.3954	0.5358	0.5379
λ_{car}	0.1793 (3.4)	0.4531 (6.1)	0.1252 (3.0)	0.1632 (2.3)	0.1325 (3.3)	0.0871 (2.1)
$\lambda_{scheduled}$	0.1919 (2.5)	0.6378 (2.1)	0.1763 (1.9)	0.1455 (1.3)	0.0455 (1.9)	0.7961 (1.2)
$\lambda_{transit}$	0.3118 (2.4)	0.2473 (1.5)	0.3023 (2.2)	0.3299 (1.3)	1.00	0.0180 (0.9)
$\lambda_{door\text{-}2\text{-}door}$	0.2929 (2.6)	0.4988 (1.6)	0.1796 (2.7)	0.1632 (2.3)	0.1792 (2.0)	0.1192 (1.7)
λ_{taxi}	0.1283 (2.9)	0.3805 (4.4)	0.0901 (2.9)	0.1632 (2.3)	0.1731 (2.2)	0.0543 (1.6)
$\lambda_{limousine}$	1.00	0.3636 (2.6)	0.2211 (1.6)	0.2475 (1.3)	0.3094 (2.3)	1.00

5.3.4 Summary of NL results

The analysis has shown that some gains in model fit can be obtained by using a nesting structure, although these gains are often not as significant as expected. This could be due to two very distinct reasons. Nested Logit models differ from the MNL model in that they accommodate a correlation between the unobserved components of utility. The first explanation interprets the closeness between the performances of the two models as a good performance by the MNL models. This would mean that the (observed) utility specification used captures almost all of the variation in utility across alternatives, reducing the scope of the NL model to capture any correlation patterns in the remaining unobserved part of utility. An alternative explanation is based on the reasoning that the nesting structure used is little better that the MNL in capturing the true structure of the underlying correlations in the unobserved component of utility. It is not clear from the outset which of these reasons is more likely; it seems that, while the specification of utility used in the MNL models was clearly quite good (relatively good ρ^2 values), further gains could be made by using a more flexible nesting structure that allows for the correlation along the three choice dimensions. As mentioned previously, attempts made during the present analysis to accommodate correlation along two dimensions resulted in non-converging models. It seems that a promising alternative would be to use cross-nested models; this approach is discussed in more detail in section 7.

Although the gains in model fit were not as important as expected, several conclusions can be drawn from the analysis discussed above. First, there seem to be important differences across population groups in the values of the structural parameters. This suggests that gains in performance could be made by using a modelling structure in which the structural parameters themselves are functions of socio-economic characteristics of travellers, as discussed by Bhat (1997) in the context of intercity travel. Secondly, the results indicate differences in performance between the three nesting structures across the six datasets used. As such, the models nesting by access-mode lead to the biggest gains in LL for the three datasets with resident travellers, while for visitors this is only the case for VFR trips, with nesting by airport leading to the biggest gains in LL for business and holiday trips. This again suggests differences in behaviour between residents and visitors. The fact that nesting by airport produces the best results in two out of the three models for visitors indicates that these travellers are less likely to choose an alternative airport, which could suggest that these travellers often simply choose the airport that is closest to their intended ground-level destination (keeping in mind that the chosen airport is actually their arrival airport from the outbound leg). This is consistent with the high values of access-time reported for these travellers by Hess and Polak (2004). Finally, nesting by airline never leads to the biggest improvements in LL; this is at least partly explainable by the lack of good data on allegiance-related factors.

6. Model validation

To validate the various models described in the paper, they were applied to the validation sample of 519 observations (divided into 6 groups for the separate models) specifically retained for this use, using the coefficient values produced during the estimation process. This enabled us to compare the models' performance in terms of correctly predicting the observed choices and in terms of recovering the market shares for the various airports, airlines and access modes, using data that is *unknown* to the models.

The validation approach produces, for every observation, a choice probability for each of the 144 elementary alternatives, where this choice probability is adjusted using the weights employed during estimation. From this, the average probability of correct prediction for the actual choice in the validation sample can be calculated. Aside from this probability of the choice of the actual triplet of airport, airline and access-mode, it is also of interest to look at the probability of correct choice for just the airport, just the airline, and just the access-mode. These probabilities can be obtained through summing the probabilities of the single elementary alternatives falling into the given group. Given the high number of elementary alternatives used in the models, the choice probability estimated for the actual chosen alternative will not necessarily be very high (although the relative probability should be); the use of these aggregated choice probabilities is thus a more accurate measure of model performance. Additionally, the choice probabilities for the individual elementary alternatives were used to calculate the weighted predicted market shares for individual airports, airlines and access-modes, which could then be compared to the actual shares of these alternatives in the control sample, using the root-mean-squared error (RMSE) between the observed and predicted shares (in percentage points) for the different composite alternatives. The RMSE was preferred to the other commonly used mean absolute percentage error (MAPE), as it is less susceptible to bias introduced by small prediction errors for alternatives with a low market share.

The results of this analysis are summarised in table 7. The first observation that can be made from this table is the surprisingly high probability of correct prediction of the actual chosen elementary alternative. Indeed, even in the poorest-fitting model (holiday trips by visitors), the probability of correct prediction is close to 30%, which is very high when one takes into account the extent of the choice-set. In terms of the correct prediction of airport choice, the probabilities range from 68.51% to as high as 85.39%. This compares very well to results in other studies, and the rates obtained in some of the models in fact exceed those obtained in many previous studies.

The performance in terms of the choice of access-mode is also very good, although generally slightly poorer than the performance in the case of airport choice, which can at least be partly explained by data problems in terms of the availability of the car mode, and lack of information on parking behaviour. The performance of the models to predict the correct choice of airline is poorer than that for the choice of airport and access-mode; however the values still always exceed 50%, despite the extensive choice set of eight airlines, and the lack of information on airline allegiance. Again, superior performance could be expected if better data were available, notably with regards to fare structures and frequent flyer programmes. The comparatively poor performance of the models for holiday trips (especially by visiting travellers, see also section 5.2.4) can possibly partly be explained by the fact that at least some of the travellers on such holiday trips have purchased a package holiday (or special flight deal); for such deals, the choice of departure airport is potentially influenced by factors that were not directly measurable and could thus not included in the models.

In terms of a comparison between the NL and MNL models, the results show that in general, the NL models perform slightly better than the corresponding MNL models. Even more so than was the case for the differences in model fit described in section 5, these differences are however far less significant than expected. This can again be seen as a reflection of the good performance of the MNL models, or the inability of the NL models to recover meaningful underlying correlation patterns in the unobserved utility components. Given the high correct prediction probability, the former reasoning however seems more likely. Overall, the best performance seems to be given by the models using nesting by mode, while nesting by airport leads to good results especially for visitors on business and holiday trips (reflected in the good model fits reported in section 5). However, the differences between the performances of the individual models are very low, and it is not directly clear what measure of error should be associated with these probabilities, such that no certain conclusions can be drawn. Nevertheless, it is interesting to note that, while the models using nesting by mode regularly outperform the other models in the correct prediction of the choice of airport and airline, this form of nesting never leads to the best results in terms of the correct prediction of mode-choice. Indeed, the best performance is in this case always obtained by the model using nesting by airport. Given the good performance of the nesting by mode for predicting the correct choice of airport, these results indicate high correlation between these two choice dimensions. Overall, the results suggest that the nesting by access-mode should be the preferred option, reinforcing earlier beliefs that the access-journey plays a crucial role in the choice of airport. Finally, even though the NL models do thus not lead to a very important gain in model fit or prediction performance, they should be preferred, given their more intuitively correct behaviour in terms of the substitution effects between alternatives.

In terms of the models' ability to recover the sample shares of the different composite alternatives, the performance is again very good, with the poorest performance being a *RMSE* of a mere 5.65 percentage points. With regards to a comparison between the performance of the MNL and NL models, the results on average show very similar performance, with the only major outlier being the poor performance in terms of airport shares by the NL model using nesting by mode in the model for VFR trips by residents.

Table 7. Model validation using control sample

		Average probability of correct prediction				Recovery of weighted sample shares (RMSE in percentage points)		
		Elementary alternative	Airport	Access mode	Airline	Airport	Access mode	Airline
Resident business	MNL	47.13%	84.04%	84.04%	60.68%	4.22%	2.26%	4.18%
	NL nesting by airport	48.02%	83.69%	85.22%	61.06%	4.34%	1.80%	4.12%
	NL nesting by airline	47.90%	84.18%	84.92%	60.30%	4.02%	1.94%	4.21%
	NL nesting by mode	48.41%	85.39%	83.76%	61.33%	3.16%	2.41%	3.87%
Visitor business	MNL	34.33%	70.69%	70.18%	55.39%	3.02%	2.32%	2.30%
	NL nesting by airport	36.19%	70.69%	72.39%	55.90%	3.10%	2.39%	2.46%
	NL nesting by airline	35.00%	71.21%	71.08%	55.27%	3.09%	2.26%	2.27%
	NL nesting by mode	34.65%	71.11%	70.25%	55.49%	2.83%	2.37%	2.19%
Resident holiday	MNL	30.56%	69.58%	67.72%	54.93%	1.90%	2.88%	3.64%
	NL nesting by airport	31.39%	69.16%	68.91%	55.03%	1.84%	3.48%	3.55%
	NL nesting by airline	31.82%	70.24%	68.64%	54.79%	1.99%	3.16%	3.64%
	NL nesting by mode	31.38%	70.98%	67.29%	55.46%	2.22%	2.66%	3.60%
Visitor holiday	MNL	27.21%	69.53%	63.22%	53.31%	3.51%	2.89%	5.65%
	NL nesting by airport	28.97%	68.51%	66.41%	54.34%	3.19%	2.97%	5.19%
	NL nesting by airline	27.78%	68.61%	64.24%	51.60%	3.62%	2.95%	5.60%
	NL nesting by mode	27.78%	72.41%	62.11%	53.49%	3.51%	3.05%	5.65%
Resident VFR	MNL	36.58%	80.83%	66.47%	60.26%	0.83%	2.27%	1.50%
	NL nesting by airport	36.74%	80.07%	67.50%	60.08%	0.99%	2.44%	1.61%
	NL nesting by airline	36.50%	80.36%	67.26%	59.41%	0.51%	2.37%	1.58%
	NL nesting by mode	39.60%	84.97%	66.16%	61.36%	2.46%	2.38%	1.25%
Visitor VFR	MNL	36.83%	73.20%	77.08%	60.97%	3.07%	5.39%	4.30%
	NL nesting by airport	36.81%	73.13%	77.25%	60.73%	3.09%	5.45%	4.33%
	NL nesting by airline	36.93%	73.26%	76.96%	60.52%	3.19%	5.38%	4.29%
	NL nesting by mode	37.83%	74.46%	76.98%	61.04%	3.08%	5.19%	4.11%

In summary, the results show very good prediction performance for the different models, where the performance is comparable, and occasionally even better than the performance obtained during a comparable application run on the actual data used during estimation (detailed results available on request). This is remarkable, given that this validation data was *unknown* to the models, and suggests that the model could be successfully used as a prediction tool (although the issue of endogeneity caused by the past experience variables would need to be addressed). In a direct comparison with the previous analysis conducted by Hess and Polak (2004), the models presented in the present paper on average lead to a better correct prediction rate (with a corresponding rate of around 72% in the previous study), showing that important gains can be made by using disaggregate level-of-service information for air-travel (i.e. avoid the use of measures of overall service at an airport), and by explicitly modelling the choice of airline and access-mode. When taking into account the fact that this was not the case in the study conducted by Hess and Polak (2004), the performance obtained in that study is in fact very good, signalling the advantages of using an approach that allows for a random distribution in the tastes of decision-makers. The combined conclusions from these two analyses suggest that important gains could be made by using a mixing framework on a purely disaggregate dataset and by explicitly modelling the individual choices of airport, airline and access mode. This is the topic of ongoing research (c.f. section 7).

7. Further research

The analysis described in this paper has shown the importance of explicitly modelling the separate, but related choices of departure airport, airline and access-mode. The analysis has also highlighted the prevalence of correlation between the unobserved components of utility along each of these three choice dimensions. However, while the three types of NL model fitted showed the effect of individually allowing for correlation along either of these dimensions, it was not possible to fit a model that allowed for correlation along multiple dimensions through using a more complicated nesting structure. This at least partly explains the lack of significant improvements in prediction performance when using the NL models instead of the MNL model. Furthermore, with this layout of the data, where each triplet of composite alternatives corresponds to exactly one elementary alternative, it is not possible fit a multi-level NL model with one level per choice dimension. It is however clearly desirable to simultaneously account for the correlations in unobserved utility components along these three dimensions, and there is a way around using the NL model while still allowing for correlation along the three choice dimensions. Indeed, it is possible to use a Cross-Nested Logit (CNL) model, where the upper level contains a nest for each of the 17 composite alternatives, and where each elementary alternative belongs to exactly three nests, one in each group (one airport, one airline and one access-mode). By using separate structural parameters, such a model would be able to show the relative level of correlation between the unobserved utility components along each of the three dimensions. For more information on this type of model, see for example Koppelman and Sethi (2000). Also, while the results of the present analysis have shown the importance of explicitly modelling the individual choice components, and of using separate models for different population groups, the results by Hess and Polak (2004) have shown the benefit of allowing for a random distribution in tastes in given subpopulations, in addition to the use of deterministic variations in tastes through a segmentation of the population. It seems likely that important gains in model performance could be made by combining the two approaches described above, thus allowing for varying correlation between the unobserved utility components, as well as a random distribution of tastes. This is possible by noting that, just as the MNL model is the most basic type of GEV model, the Mixed Multinomial Logit (MMNL) model used by Hess and Polak (2004) is simply the most basic member of a wider class of mixed GEV (MGEV) models (c.f. Train, 2003). This opens up the possibility of the use of a Mixed CNL model for the joint analysis of airport, airline and access mode; this is the topic of an ongoing study, and is an important avenue for further research. Finally, two other aspects that should be considered are choice-set generation (as proposed by Basar and Bhat, 2004) and the availability of different fare classes (c.f. Battersby, 2003). These additional modelling components could relatively readily be incorporated into a wider framework of an MGEV model, and could lead to important further gains in model accuracy.

8. Conclusions

In this paper, we have described an in-depth analysis of the combined choices of departure airport, airline and access-mode for passengers departing from the San Francisco Bay area. The analysis has shown that several factors, most notably flight frequency and in-vehicle access-time have a significant overall impact on the appeal of a given airport, while factors such as fare and aircraft size have a visible impact only for some of the population subgroups.. Our study has highlighted the need to use separate models for resident and non-resident travellers, and has also shown the benefit of using individual models for different journey purposes. From a utility specification perspective, the research has shown that important gains in model fit can be obtained through the use of a non-linear specification of flight frequency, and for some journey purposes, through a segmentation of the population into different income classes. Accommodating the effects of past experience on the current choice also leads to very significant improvements in model fit.

In terms of modelling structure, the analysis has shown that significant gains in model fit can be obtained through the use of a Nested Logit model, although these improvements are less significant than expected and do not in general translate into important advantages in terms of model prediction performance, which is already very good for the MNL models used as a base. This could suggest that an appropriate specification of utility is more important than the use of an adequate nesting structure. Nevertheless, the paper has highlighted the importance of explicitly modelling the three separate choice dimensions of airport choice, airline and access mode, and while the advantages of the nesting approach are less significant than expected, these models should still be preferred, as they give more intuitively correct representation of the choice process.

Acknowledgements

The authors would like to acknowledge the cooperation of the SF-Bay area Metropolitan Transport Commission, especially Chuck Purvis. Similarly, the authors would like to acknowledge the support of *Back Aviation Solutions*, especially John Weber. We would also like to thank Andrew Daly and Michel Bierlaire for support and guidance in the estimation stages of this project.

References

Ashford, N., and Bencheman, M. 1987. Passengers' choice of airport: an application of the Multinomial Logit model, Transportation Research Record, vol. 1147, pp. 1-5.

Basar, G., and Bhat, C. R. 2004. “A Parameterized Consideration Set model for airport choice: an application to the San Francisco Bay area”, Transportation Research, B38, pp. 889-904.

Battersby, B. 2003. Quasi-maximum likelihood estimation of censored multi-class air-travel demand in regional Australia, paper presented at the 2003 Australian Conference of Economists, October 2003, Canberra.

Bhat, C. 1997. Covariance heterogeneity in nested logit models: Econometric structure and application to intercity travel, Transportation Research, B31, 11–21.

Brooke, A.S., Caves, R. E., and Pitfield, D.E. 1994. Methodology for predicting European Short-Haul Air Transport Demand from Regional Airports: An Application to East Midlands International Airport, Journal of Air Transport Management, 1 (1) 37-46.

Harvey, G. 1987. Airport choice in a multiple airport region, Transportation Research, A21 (6) 439-449.

Hess, S., and Polak, J.W., 2004. Mixed Logit modelling of airport choice in multi-airport regions, Journal of Air Transport Management, forthcoming.

Innes, J.D., and Doucet, D.H., 1990. Effects of access distance and level of service on airport choice, Journal of Transportation Engineering, 116 (4) 507-516.

Furuichi, M., and Koppelman, F. S. 1994. An analysis of air traveler's departure airport and destination choice behaviour, Transportation Research, A28(3) 187-195.

Koppelman, F.S., and Sethi, V., 2000. Closed Form Discrete Choice Models, in Handbook of Transport Modeling, D. Hensher and K. J. Button, eds., Pergamon Press, Oxford.

McFadden, D. 1978a. Modelling the choice of residential location, in A. Karlquist, ed., Spatial Interaction Theory and Planning Models, North Holland, Amsterdam, chapter 25, 75–96.

MTC, 1995. Metropolitan Transportation Commission Airline Passenger Survey: Final Report, prepared by J.D. Franz Research, Oakland, CA.

MTC, 2000. Aviation Demands Forecast: Executive Summary, Metropolitan Transportation Commission, Oakland, California.

Ndoh, N.N., Pitfield, D.E., and Caves, R.R. 1990. Air transportation passenger route choice: a Nested Multinomial Logit analysis, in Spatial Choices and Processes, M. M. Fisher, P. Nijkamp, and Y. Y. Papageorgiou, eds., Elsevier Science Publishers B.V., North Holland, pp. 349-365.

Ozoka, A.I., and Ashford, N., 1989. Application of disaggregate modelling in aviation systems planning in Nigeria: a case study, Transportation Research Record, 1214, 10-20.

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

Pels, E., Nijkamp, P., and Rietveld, P. 2003. Access to and competition between airports: a case study for the San Francisco Bay area, Transportation Research, A37 71-83.

RAPC, 2000. Regional Airport System Plan Update 2000, Regional Airport Planning Committee, Oakland, CA.

Skinner, R.E. Jr., 1976. Airport choice: An empirical study, Transportation Engineering Journal, 102 (4) 871-883.

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

Thompson, A., and Caves, R., 1993. The projected market share for a new small airport in the south of England, Regional Studies, 27 (2) 137-147.

Veldhuis, J., Essers, I., Bakker, D., Cohn, N., and Kroes, E., 1999. The Integrated Airport Competition Model, 1998, Journal of Air Transportation World Wide, 4 (2).

Windle, R., and Dresner, M., 1995. Airport choice in multi-airport regions, Journal of Transportation Engineering, 121 (4) 332-337.