AIRPORT, AIRLINE AND ACCESS MODE CHOICE IN THE SAN FRANCISCO BAY AREA

Stephane Hess

Centre for Transport Studies

Imperial College London

Exhibition Road

London SW7 2AZ

stephane.hess@imperial.ac.uk

Tel: +44(0)20 7594 6105

Fax: +44(0)20 7594 6102

John W. Polak

Centre for Transport Studies

Imperial College London

Exhibition Road

London SW7 2AZ

j.polak@imperial.ac.uk

Tel: +44(0)20 7594 6089

Fax: +44(0)20 7594 6102

Abstract

In this paper, we present an analysis of air travel choice behaviour in the San Francisco Bay area. The analysis extends existing work by considering the simultaneous choice by passengers of a departure airport, an airline, and an access mode. The analysis shows that several factors, most notably flight frequency and in-vehicle access time, have a significant overall impact on the attractiveness of an airport, airline and access mode combination, 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 the inclusion of airport-inertia variables, and through using a nonlinear specification for the marginal returns of increases in flight frequency. In terms of model structure, the results suggest that the use of the different possible two-level Nested Logit models leads to modest, yet significant gains in model fit over the corresponding Multinomial Logit models, which already exhibit very high levels of prediction performance.

Keywords: airport choice, airline choice, multi airport regions, discrete choice, nested logit

JEL codes: C25, R41

1 Introduction

The analysis of air travellers' choices of airports is an important component in 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, and the congestion in the local transportation network.

Studies of air travel choice have become increasingly popular over recent years. While most have used very basic models to analyse the choice of departure airport, several studies have employed advanced model structures allowing for correlation between different alternatives (e.g. airports). Recently, it has also been shown that significant 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 (Hess and Polak 2005).

However, a passenger's choice of airport will in general be closely related to a number of other dimensions of travel behaviour, especially the choice of airline and airport access mode and the nature of the interactions between, and substitution patterns within and across these choice dimensions is not clear *a priori*. While some studies have recognised this issue, the majority of published work at best looks 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.

The main aim and contribution of this paper is to formulate a model for the combined choice of airport, airline and access mode for passengers departing from the San Francisco Bay (SF-bay) area, and to investigate the prevalence of correlation along these three choice dimensions. Furthermore, the study aims to determine whether there are differences across groups of travellers in the substitution patterns across airport, airline and access mode alternatives.

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 visiting as well as resident passengers, the analysis indirectly also looks 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). An additional reason for excluding connecting passengers is that their choice set does not generally contain multiple airports located in the same metropolitan area, and the analysis of the choice between connecting airports spread across multiple multi airport regions is beyond the scope of the present paper.

The remainder of this article is organised as follows. In the next section, we present a brief overview of existing work in the area of air travel choice behaviour. In the third section, we discuss the various datasets used, while in Section 4, we present the models used in the analysis. The results of the analysis are presented in Section 5, and model validation is carried out in Section 6.

2 Literature review

In this section, we give a brief review of the existing body of work in the area of air travel choice behaviour modelling; for other reviews on this topic, see for example Basar & Bhat (2004), Pels et al. (2003) and Hess and Polak (2005).

One of the first 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 a MNL model, Windle & Dresner (1995) repeat the earlier results, and also reveal a significant inertia effect; 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 a 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 another study of airport choice in this area, Basar & 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 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 (2005) have recently used the SF-bay area data 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 & Bencheman (1987), who use a 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. In a study of passenger route choice in central England, Ndoh et al. (1990) find that the NL model outperforms the MNL model. Thompson & Caves (1993) use a MNL model to forecast the market share for a new airport in North England; access time, flight frequency and aircraft size 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 a MNL analysis of the distribution of passengers between airports in the Midlands, Brooke et al. (1994) find flight frequency to be the most important factor.

In other studies, Ozoka & Ashford (1989) use a 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. Innes & Doucet (1990) use a binary logit model for airport choice in Canada to show that, *ceteris paribus*, travellers prefer jet services to turboprop flights. Furuichi & Koppelman (1994) use a 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) develop the comprehensive Integrated Airport Competition Model, showing that passenger behaviour is represented most appropriately by a NL choice process that models the choice of main mode above the combined choice of airport and air route, and finally the choice of access mode at the chosen airport.

This brief review has shown that although there exists a large body of work on the modelling of airport choice in multi airport regions, most studies use rather basic modelling techniques, with a heavy emphasis on the MNL model. Furthermore, to the authors' knowledge, none of the existing studies explicitly deals with the three-dimensional nature of the choice process (airport, airline and access mode), with the possible exception of the work by Veldhuis et al. (1999), which, by being applied to the Amsterdam region, cannot be seen as a multi airport study per se.

3 Data

The SF-Bay area 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 (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 (MTC 1995). This contained information on over 21,000 departing air travellers. 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 undersampled. This needs to be taken into account in the modelling approach, as described in Section 4.

It was decided to use only destinations that could be reached by direct flight from all three airports, on every day of the week, leading to 14 destinations. After extensive data cleaning, a final sample of 5,091 observations was obtained. The resulting dataset, which is summarised in Table 1, 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).

Special care is required in the case of destinations that are themselves located in multi airport regions. 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. This applies specifically in the case of non-resident passengers, where, under normal circumstances, this airport constitutes the origin of their trip. It is in this case crucial to guarantee that an explicit choice of airport was made in the SF-bay area.

Destinations from two such multi airport regions, namely the wider Los Angeles (LA) area, as well as Chicago's O'Hare (ORD) airport, were included in the present analysis. The decision to include airports from the LA area was motivated by the frequency of these destinations in the survey data. 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. 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 and then excluded these destinations, revealed no major differences, suggesting that the inclusion of these airports has no ill effects on the subsequent analysis.

3.2 Air travel level-of-service data

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 airlines $\mathbf{A_1}$ to $\mathbf{A_8}$. Besides the frequencies for the different operators, the dataset contains information on flight times and the type of aircraft used. Additionally, 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 distinction is made between the fares for the different classes of travel. Furthermore, the fact that no data is available on the availability of different ticket classes at the time of booking leads to an assumption of similar selling speed on all routes. These assumptions are a common requirement in studies of air travel choice behaviour based on revealed preference passenger

[1] Back Aviation Solutions, 6000 Lake Forrest Drive, Suite 580, Atlanta, GA 30328, www.backaviation.com

survey data, and are at least a contributing factor in the problems of retrieving a significant effect of air fares in many such studies. Finally, the dataset was complemented by information on on-time performance[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 data 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 1,099 travel area zones (TAZ) used for the SF-Bay area.

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 and the use of car pool lanes). 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. Due to complications with the treatment of rental charges, parking costs and marginal car running costs, a common car alternative was used, 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.

[2] Available from the Bureau of Transport Statistics, via www.bts.gov/programs/oai/airline_ontime_statistics

The availability of public transport was obtained from the MTC OD matrices, and, in the absence of appropriate 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

In the data used in model calibration, each respondent is observed to choose a triplet of alternatives, one in each of the three dimensions of choice. The triplet of alternatives for a given respondent forms the dependent variable for that observation in the models. 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. For each observation, data on the attributes and availability of the sub-alternatives along each dimension was appended to the survey data, taking into account the ground level origin of a traveller, the season (August vs October), the choice of destination, and the day of week and time of day (peak vs off-peak).

4 Modelling methodology

4.1 Discrete choice models

The analysis described in this paper makes use of two types of discrete choice model belonging to the family of Generalised Extreme Value (GEV) models, namely MNL and NL. The main difference between these two model structures comes in the assumptions made with regards to the error structure; here, the MNL model assumes uncorrelated errors, while the NL model allows for varying levels of correlation between the error-terms of the utility functions of the different alternatives. In the present context, this can be exploited to allow for correlation between two alternatives sharing a common airport, or a common airline, or a common access mode. This in turn leads to higher substitution patterns between these alternatives. In a NL model, alternatives that have non-zero correlation are grouped together in a nest *m*, where this nest *m* has an associated logsum (nesting) parameter λ_m, which measures the degree of independence between alternatives in the respective nest, with higher λ_m meaning more independence and hence lower correlation between the unobserved components of utility of the alternatives contained in the nest. The correlation is given by $1-\lambda_m{}^2$, such that a value of 1 for all structural parameters leads to the MNL model. For a more detailed discussion of discrete choice models, see Train (2003).

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, with separate weights calculated for separate sub-samples. In the estimation process, each term in the log-likelihood function was then multiplied by the appropriate weight for the 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 (2005) on the same data show that there exist significant differences along both dimensions, with the differences across trip purposes being generally more significant than those between residents and visitors. Following extensive diagnostic testing, separate models were used for residents and for visitors, with additional divisions into business travellers, holiday travellers, and travellers visiting friends and family (VFR), leading to a total of six distinct segments.

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 utility functions used in the analysis. We then describe the results from the MNL models, and finally summarise the findings from the NL modelling analyses.

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 air fare and access 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 air 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, while we also explored the influence of aircraft type (jet *vs* turboprop). Both linear and various non-linear specifications of the different explanatory variables were tested, where 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. Similarly, it was not possible to identify a significant direct effect of the on-time performance of airlines or airports on the respective choice probabilities. Attempts were made to segment the population by income, where three income groups were defined, segmenting the population into low income (<$21,000 per annum), medium income (between $21,000 and $44,000 per annum) and high income (above $44,000 per annum).

A further specification issue that was explored was the inclusion of airport inertia variables in the utility functions, as discussed by Windle & 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. For each one of the three airports, a coefficient in the utility function was thus associated with the inertia variable for that airport, where, to account for cross effects, coefficients in a given airport's utility function were also associated with the inertia variables of the remaining two airports. After normalisation, this led to the use of three airport specific inertia coefficients and three cross coefficients (SJC and OAK on SFO, and SFO on SJC). The inclusion of these variables did in each case, as expected, lead to dramatic improvements in log-likelihood (LL), where the gains were even more significant when using a log-transform, such that this approach was adopted. 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. The dependence on past choices would also 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 inertia terms did not introduce major bias.

5.2 MNL models

In the following 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 air 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 indifference to fare increases on the part of business travellers, at least in 1995. Finally, increases in flight frequency lead to increases in utility, where the logarithmic transform ensures decreasing marginal returns.

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 the airport inertia variables, the estimates show positive direct effects for all three airports, with positive cross-effects of past usage of SJC and OAK on the utility of SFO, and a positive (but not significant) cross-effect of past usage of SFO on the utility of SJC.

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 was 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, while flight frequency again has a positive effect. Finally, unlike in the model for resident business travellers, the inertia cross-effect of past flights at OAK has a negative effect on the utility of SFO, while the cross-effect of past flights at SJC on the utility of SFO is now insignificant, while there is a positive cross-effect of SFO acting on the utility of SJC.

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, and a positive effect of flight frequency. All inertia coefficients are positive, though the cross-effect of past flights at SFO on the utility of SJC is not significant. Finally, for this group of travellers, a negative effect could be identified for fare (although of lower statistical significance) 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.

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. 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 inertia cross-effect estimates 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.

5.2.6 VRF trips by visitors

The final subsample used in the estimation of the 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 inertia cross-effect estimates 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 the six segments in the optimal specification of utility. The common point across all the segments is that a logarithmic specification is always preferable to a linear specification in the case of the frequency and inertia coefficients. 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, 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. 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 change in utility associated with 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 change in utility 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 $\beta_{FT}\,/\,\beta_{AT}$ 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

Several important issues arise in the specification of NL models. 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 with the current data and specification of alternatives. Thus, in this paper, we are restricted to two-level structures, where the interest now lies in a comparison of the performance of the three possible structures (i.e., 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, 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 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 (calculated with respect to unity). 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. All else being equal, passengers are not more likely to shift to another alternative at SFO than they are to shift to an alternative at another airport.

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 (i.e. higher correlation for alternatives sharing an airport) than is the case for residents.

5.3.2 Nesting by airline

The lack of information on frequent flier programme membership means that there should be some correlation in the unobserved part of utility between different alternatives that share the same airline. As such, 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.

In the models using nesting by airline, a comparatively high number of structural parameters had to be constrained to a value of 1. Nevertheless, except for the model for visitor VFR trips, the use of this structure resulted in significant increases in log-likelihood over the corresponding MNL models. 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 the models are able to pick up some effect of correlation between alternatives associated with given airlines. It can also be noted that airlines $\mathbf{A_5}$ and $\mathbf{A_8}$ on average have lower structural parameters than the other airlines. This could at least be partly related to the fact that these two carriers run a budget airline scheme; this sets them apart from other alternatives, potentially explaining the correlation, especially in the absence of an appropriate treatment of the cost structure in the models.

5.3.3 Nesting by access mode

The results of this analysis are summarised in Table 6. In many regards, nesting by access mode proved to be the most promising approach. 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' strong allegiance to car as an access mode. A comparably 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 three of the structural parameters reported in Table 6 are not statistically different from 1. Setting these parameters to 1 however either led to a significant drop in log-likelihood or did not lead to significantly changed values of the other structural parameters and coefficients. 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 show that important gains can be made by using a structure that nests alternatives by access mode, suggesting that a number of attributes that could not be included in the utility functions lead to heightened correlation between alternatives sharing the same mode.

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 correlation between the unobserved components of utility. The first explanation interprets the similarity in the performances of the two models as an endorsement of the MNL models. This would mean that the (observed) utility specification used captures almost all of the correlation 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 specific nesting structures used are little better than the MNL model in capturing the true structure of the underlying correlations in the unobserved component of utility. The same conclusion would extend to the multiple-level NL structures initially explored. It is not clear from the empirical results alone which of these potential explanations is most appropriate. Perhaps the most promising direction for future research is to explore the applicability of more flexible structures such as a cross-nested form. If these also prove to offer little empirical advantage over the MNL then clearly this would reinforce confidence in the MNL structure (and conversely if a cross-nested structure is empirically superior). However, the findings are in each case clearly specific to the data and utility specification used in the present analysis.

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. 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 model fit 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 model fit for business and holiday trips. Finally, nesting by airline never leads to the biggest improvements in model fit.

6 Model validation

Model validation consisted of using the estimated models in conjunction with the validation sub-sample of 519 observations (not used in model estimation) in order to test the ability of the models to correctly recover the observed choices and market shares for the various airports, airlines and access modes.

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 prediction of the 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 validation sample, using the root-mean-squared error (RMSE) between the observed and predicted shares (in percentage points) for the different composite alternatives.

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 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 in predicting 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 process is potentially influenced by factors that were not directly measurable and could thus not be 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 access 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 in performance between the individual structures 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 access 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. Finally, even though the NL models do thus not lead to very important gains in model fit or prediction performance, they should be preferred, given their more intuitively correct behaviour in terms of the substitution patterns between alternatives. This comes despite a slight increase in the cost of estimation for these structures, which is however nowhere nearly as severe as when comparing closed form models to mixture structures such as Mixed Logit.

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.

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 suggests that the models have not been overfitted on the estimation data. In a direct comparison with the previous analysis conducted by Hess and Polak (2005), 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. avoiding the use of measures of overall service at an airport), and by explicitly modelling the choice of airline and access mode.

Conclusions

In this paper, we have presented a detailed analysis of the joint 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. Here, it should be noted that, except for passengers on very flexible tickets, frequency is not taken into account directly by the respondents, but captures a host of effects, including visibility, capacity, and schedule delay (under the assumption of a relatively even spread of departure times).

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. Finally, the inclusion of airport inertia variables led to very significant improvements in model fit across all population segments.

In terms of model structure, the analysis has shown that statistically 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. The modest extent of the gains in performance is at least partly due to the inability to fit a model allowing for correlation along multiple dimensions through using a more complicated nesting structure. As it is however clearly desirable to simultaneously account for the correlations in unobserved utility components along these three dimensions, the use of a cross-nesting structure is an important avenue for further research. Here, the upper level would contain a nest for each of the 17 composite alternatives, and each elementary alternative would belong to exactly one nest 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. Independently of this, the paper has clearly shown the benefit of explicitly modelling the three separate choice dimensions of airport, airline and access mode.

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 for support and guidance in the estimation stages of this project. Finally, we have benefited from the comments of three anonymous referees.

References

Ashford N, Bencheman M (1987) Passengers' choice of airport: an application of the multinomial logit model. *Transportation Research Record* 1147: 1-5

Basar G, Bhat C (2004) A parameterized consideration set model for airport choice: an application to the San Francisco Bay area. *Transportation Research* B38: 889-904

Brooke AS, Caves RE, Pitfield DE (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, Polak JW (2005) Mixed Logit modelling of airport choice in multi-airport regions. *Journal of Air Transport Management* 11(2): 59-68

Innes JD, Doucet DH (1990) Effects of access distance and level of service on airport choice. *Journal of Transportation Engineering* 116(4): 507-516

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

MTC (1995) *Metropolitan Transportation Commission Airline Passenger Survey: Final Report*. J.D. Franz Research, Oakland, CA.

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

Ndoh NN, Pitfield DE, Caves RR (1990) Air transportation passenger route choice: a nested multinomial logit analysis. In: Fisher MM, Nijkamp P, Papageorgiou YY (eds) *Spatial choices and processes*. North Holland, Amsterdam, pp 349-365

Ozoka AI, 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, 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, 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 RE (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, 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, Kroes, E (1999) The integrated airport competition model, 1998. *Journal of Air Transportation World Wide* 4(2): 100-120

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

TABLE 1: Destinations used in the analysis (number of respondents)

		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[1]	55	65	36	57	199	35	89	140	128	1	258	213	42	37	1,355
	SJC[2]	167	71	91	163	367	111	58	106	133	156	248	169	61	247	2,148
	OAK[3]	211	9	25	68	381	135	1	101	51	39	139	208	43	177	1,588
Total		433	145	152	288	947	281	148	347	312	196	645	590	146	461	5,091

1. SFO = San Francisco International

2. SJC = San Jose Municipal

3. OAK = Oakland International

TABLE 2: MNL results for residents (selected coefficients)

	Business		Holiday		VFR	
	estimate	**t-test**	**estimate**	**t-test**	**estimate**	**t-test**
Access cost ($)			-0.0208	-2.21	-0.0223	-2.29
Access cost ($), inc. >$44,000 p.a.	-0.0244	-2.86				
Access cost ($), inc. <$44,000 p.a.	-0.0358	-4.17				
Access in-vehicle time (min)	-0.0522	-12.13	-0.0594	-12.94	-0.0490	-9.43
Walk time (min), inc. >$44,000 p.a.	-0.1531	-2.97				
Walk time (min), inc. <$44,000 p.a.	-0.1139	-2.47				
Fare ($)			-0.0131	-1.90	-0.0267	-3.03
Flight time (min)	-0.0471	-2.37				
Flight frequency (log of frequency)	1.3183	10.77	1.3235	9.35	1.4447	7.87
Turboprop (dummy)	-2.5296	-3.20	-4.2294	-2.70		
Inertia variables (log of flights in last 12 months)						
OAK on OAK	1.9993	9.44	2.1024	5.09	2.2919	5.24
SFO on SFO	1.1829	9.62	1.1887	7.89	2.0488	8.83
SJC on SJC	1.9641	8.49	2.5909	5.04	3.1690	5.87
OAK on SFO	0.6619	3.37	0.8328	1.98	0.4413	1.02
SJC on SFO	0.7845	3.68	1.4302	2.71	0.5574	1.10
SFO on SJC	0.1731	1.07	0.1618	0.79	0.0292	0.09
Observations	1,098		831		641	
Log-likelihood	-1551.62		-1384.81		-1050.84	
ρ^2	0.5934		0.5198		0.5157	

TABLE 3: MNL results for visitors (selected coefficients)

		Business		Holiday		VFR	
		estimate	**t-test**	**estimate**	**t-test**	**estimate**	**t-test**
	Access cost ($)			-0.0145	-1.66		
	Access cost ($), inc. >$44,000 p.a.	-0.0219	-2.55				
	Access cost ($), inc. <$44,000 p.a.	-0.0286	-3.94				
	Access in-vehicle time (min)			-0.0769	-13.22	-0.0698	-11.06
	In-vehicle time (min), inc. >$22,000 p.a.	-0.0820	-14.43				
	In-vehicle time (min), inc. <$22,000 p.a.	-0.0496	-7.18				
	Wait time (min)	-0.2507	-3.28				
	Fare ($), inc. <$21,000 p.a.					-0.0501	-3.55
	Fare ($), inc. [$21,000,$44,000] p.a.					-0.0267	-1.95
	Flight time (min)	-0.0293	-1.39	-0.0908	-3.42	-0.1522	-5.12
	Flight frequency (log of frequency)	1.3066	11.34	1.0783	7.51	0.7244	4.41
Inertia variables (log of flights in last 12 months)	**OAK on OAK**	1.1881	6.57	1.2529	2.90	1.3899	2.96
	SFO on SFO	1.9324	9.39	0.7514	3.97	1.0991	3.35
	SJC on SJC	1.3973	6.10	2.0564	4.42	2.2569	4.17
	OAK on SFO	-0.7172	-3.36	-0.4741	-0.99	0.1887	0.35
	SJC on SFO	0.0075	0.03	0.8318	1.86	-0.1219	-0.17
	SFO on SJC	0.5032	2.38	-0.1084	-0.34	0.1809	0.42
	Observations	1,057		534		421	
	Log-likelihood	-1517.68		-1018.24		-621.81	
	ρ^2	0.4477		0.387		0.5236	

TABLE 4: NL results for nesting by airport (t-statistics calculated with respect to 1)

	Business		Holiday		VFR	
	Resident	Visitor	Resident	Visitor	Resident	Visitor
MNL LL	-1551.62	-1517.68	-1384.81	-1018.25	-1050.84	-621.81
NL LL	-1545.14	-1487.71	-1372.19	-999.51	-1039.67	-621.62
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 (4.02)	0.5259 (10.64)	0.7627 (4.08)	0.4399 (8.79)	0.6708 (5.5)	0.9333 (0.63)
λ_{OAK}	0.8925 (1.64)	0.7178 (3.7)	0.7258 (4.61)	0.7373 (2.24)	0.7828 (3)	1.00

TABLE 5: NL results for nesting by airline (t-statistics calculated with respect to 1)

	Business		Holiday		VFR	
	Resident	Visitor	Resident	Visitor	Resident	Visitor
MNL LL	-1551.62	-1517.68	-1384.81	-1018.25	-1050.84	-621.81
NL LL	-1536.66	-1507.62	-1371.21	-1003.93	-1034.07	-620.24
NL ρ^2	0.5974	0.4514	0.5245	0.3956	0.5234	0.5248
λ_{A1}	0.9499 (0.25)	0.9617 (0.14)	0.9237 (0.32)	0.6989 (1.34)	1.00	1.00
λ_{A2}	0.6108 (4.59)	0.9822 (0.16)	0.7841 (1.05)	0.6249 (4.62)	0.8663 (1.47)	0.8606 (1.17)
λ_{A3}	1.00	0.8895 (0.36)	1.00	0.7697 (1.17)	0.8617 (0.43)	0.8549 (0.61)
λ_{A4}	1.00	0.6538 (2.22)	1.00	0.7237 (1.07)	1.00	0.6762 (1.25)
λ_{A5}	0.7433 (3.35)	0.6317 (2.22)	0.7379 (2.66)	0.3917 (4.97)	0.6344 (3.92)	1.00
λ_{A6}	1.00	1.00	0.9967 (0.03)	0.6761 (2.44)	1.00	0.7935 (2.13)
λ_{A7}	1.00	1.00	1.00	1.00	1.00	1.00
λ_{A8}	0.8389 (0.9)	0.7921 (1.13)	0.7240 (3.28)	0.5298 (7.01)	0.6664 (1.35)	0.8399 (0.71)

TABLE 6: NL results for nesting by access mode (t-statistics calculated with respect to 1)

	Business		Holiday		VFR	
	Resident	Visitor	Resident	Visitor	Resident	Visitor
MNL LL	-1551.62	-1517.68	-1384.81	-1018.25	-1050.84	-621.81
NL LL	-1520.42	-1508.79	-1351.18	-1004.26	-1007.20	-603.07
NL ρ^2	0.6016	0.4510	0.5315	0.3954	0.5358	0.5379
λ_{car}	0.1793 (15.6)	0.4531 (7.4)	0.1252 (20.9)	0.1632 (11.8)	0.1325 (21.6)	0.0871 (22.0)
$\lambda_{scheduled}$	0.1919 (10.5)	0.6378 (1.2)	0.1763 (8.9)	0.1455 (7.6)	0.0455 (39.9)	0.7961 (0.3)
$\lambda_{transit}$	0.3118 (5.3)	0.2473 (4.6)	0.3023 (5.1)	0.3299 (2.6)	1.00	0.0180 (49.1)
$\lambda_{door\text{-}2\text{-}door}$	0.2929 (6.3)	0.4988 (1.6)	0.1796 (12.3)	0.1632 (11.8)	0.1792 (9.2)	0.1192 (12.6)
λ_{taxi}	0.1283 (19.7)	0.3805 (7.2)	0.0901 (29.3)	0.1632 (11.8)	0.1731 (10.5)	0.0543 (27.9)
$\lambda_{limousine}$	1.00	0.3636 (4.6)	0.2211 (5.6)	0.2475 (3.9)	0.3094 (5.1)	1.00

TABLE 7: Model validation using control sample

Segment	Model structure	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%