Cross-Nested Logit modelling of the combined choice of airport, airline and access-mode

Stephane Hess* John W. Polak†

August 1, 2005

Abstract

The analysis of air-passengers' choices of departure airport in multi-airport regions is a crucial component of transportation planning in many large metropolitan areas, and has been the topic of an increasing number of studies over recent years. In this paper, we advance the state of the art of modelling in this area of research by making use of a Cross-Nested Logit (CNL) structure that allows for the joint representation of inter-alternative correlation along the three choice dimensions of airport, airline and access-mode. The analysis uses data collected in the Greater London area, which arguably has the highest levels of inter-airport competition of any multi-airport region; the authors of this paper are not aware of any previous effort to jointly analyse the choice of airport, airline and access-mode in this area. The results of the analysis reveal significant influences on passenger behaviour by access-time, access-cost, flight-frequency and flight-time. A structural comparison of the different models shows that the cross-nested structure offers significant improvements over simple Nested Logit (NL) models, which in turn outperform the Multinomial Logit (MNL) model used as the base model.

Word Count: 7,489 words, 9 tables, and 3 figures

1 Introduction

The analysis of air-travellers' choice-behaviour, particularly in multi-airpoprt regions is an important and growing area of travel demand research. The

*Centre for Transport Studies, Imperial College London, stephane.hess@imperial.ac.uk, Tel: +44(0)20 7594 6105, Fax: +44(0)20 7594 6102

†Centre for Transport Studies, Imperial College London, j.polak@imperial.ac.uk, Tel: +44(0)20 7594 6089, Fax: +44(0)20 7594 6102

modelling of such choices is not only challenging from a research perspective but is also of great practical policy importance in many large metropolitan areas, where accurate forecasts of passenger distribution are required. However, the majority of existing studies have been limited by the use of inflexible model forms and over-aggregated data. Furthermore, studies of airport choice behaviour have generally failed to acknowledge the multi-dimensional nature of the choice-set; for example, passengers not only make a choice of departure airport, but additionally choose an airline and an access-mode.

In this paper we present a joint model of these three dimension of air travel choice which uses a cross-nesting structure that allows for the flexible representation of correlation along the three choice dimensions, avoiding the constraints associated with a multi-level nesting structure. Our model uses data from the London area, which is arguably the most competitive multi-airport region in the world. Due principally to limitations in the availability of relevant data, in common with most previous work, we focus exclusively on the behaviour of departing passengers on direct flights.

The remainder of this paper is organised as follows. The next section presents a brief overview of the existing literature. This is followed in section 3 by a description of the data and choice scenarios considered. Section 4 describes the modelling approach and Section 5 presents the estimation results. The final section summarises the main results and conclusions.

2 Literature review

Although research into airport-choice behaviour stretches back to the work of Skinner [17], the majority of studies in this area date from the mid 1990s onwards and have made use of data collected in the San Francisco Bay (SF-Bay) area. Pels et al. [14] conduct a Nested Logit (NL) analysis of the combined choice of airport and airline and find that that airline choice is nested within the choice of airport, while Pels et al. [15] 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. Basar and Bhat [2] propose the use of a two-level modelling structure in which the airport choice process is preceded by a choice-set generation stage, thus acknowledging the fact that not all travellers consider all available airports. The results show flight frequency to be the most important factor in choice set composition, with access-time playing this role in the actual choice of airport. Hess and Polak [11] use the SF-bay data in a study aimed at showing the role of random taste heterogeneity in airport-choice behaviour, in turn introducing the Mixed Multinomial Logit (MMNL) model to the area of air-travel choice behaviour. 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 is purely random. Hess and Polak [9] explicitly model the combined choice of airport, airline and access-mode and Hess and Polak [10] adapt this framework to additionally allow for random taste heterogeneity within the three-dimensional choice-process.

To date, three main academic studies of airport access in the United Kingdom (UK) have been undertaken. Ashford and Bencheman [1] use an MNL model for airport choice at five airports in England (including two London airports), and find that access time and flight frequency are significant factors overall, while flight fares only have an impact for domestic passengers and for international leisure travellers. Thompson and Caves [18] use an MNL model to forecast the market share of a potential new airport in Southern England; access time, flight frequency and aircraft-size are found to be significant. Finally, in an MNL analysis of the distribution of passengers in the Midlands region, Brooke et al. [5] find flight frequency to be the most important factor, where a single London airport, Heathrow, was also included in the choice-set. The wider literature on airport choice is reviewed by Pels et al. [14, 15], Basar and Bhat [2] and Hess and Polak [9].

3 Description of data and choice scenario

3.1 Context

The Greater London area has by far the highest levels of air traffic in Europe, with, in 2002, some 117.13 million passengers using the five main airports. The area is dominated by Heathrow (LHR), the world's busiest international airport (measured in terms of the number of passengers on international routes), and the main hub in Europe. Additionally, a large number of routes are offered from Gatwick (LGW), the world's busiest single-runway airport, while Stansted (STN), Europe's fastest growing major airport, and Luton (LTN) act mainly as bases for holiday and low-cost operators. Finally, the centrally located London City (LCY) airport caters primarily to business travellers, and, due to its short runway, is restricted to short-haul flights operated by turboprop planes and small jet aircraft.

3.2 Passenger survey data

For the present analysis, data from the 1996 passenger survey[1] were obtained from the Civil Aviation Authority [6]. Although slightly dated, this is the most recent large-size full survey available for this region. The use of a more recent version of the dataset is an important avenue for further research.

The original sample obtained from the CAA contained responses from 47,831 passengers, for 31 destinations (reachable by direct flights from at least two of the five London airports), and 37 airlines. After data-cleaning, a usable sample of 33,527 passengers was obtained, which was split into four subsets, dividing the population into residents and visitors, and using a purpose split of business *vs* leisure. In each sub-group, a 95% subsample was used for model calibration, with the remaining 5% begin retained for model validation. The results of this validation process are not reported here, but are available from the first author on request; the differences in performance between the estimation and validation sample were negligible.

Of the 31 destinations used in the analysis, 5 are in Great Britain (Aberdeen, Edinburgh, Leeds, Manchester and Newcastle), 1 on the Channel Islands (Guernsey), 3 in Ireland (Cork, Dublin and Shannon), 3 in the Benelux (Amsterdam, Brussels and Rotterdam), 3 in Scandinavia and the Nordic countries (Copenhagen, Gothenburg and Helsinki), 3 in Germany (Düsseldorf, Hamburg and Munich), 3 in Austria and Switzerland (Geneva, Vienna, Zurich), 1 in France (Nice), 3 in Spain (Barcelona, Madrid, Malaga), 2 in the South East of Europe (Athens and Larnaca), 1 in the Middle East (Tel Aviv), and 3 in the United States (Boston, Detroit and Miami).

All destinations included in the sample are served by a single main airport, avoiding the problem with multi-airport destination areas described by Hess and Polak [9]. Nevertheless, some destinations remain where there is competition between air and ground-level transport, namely the 5 destinations in Great Britain, and Brussels, where there is competition with Eurostar. This competition is not taken into account in the present study, where we work on the basis of an a priori choice of main mode. The analysis of the joint choice of main mode and air-side related choices remains an important area for further research.

[1] Data collected from departing passengers at the airports.

3.3 Air-travel level-of-service data

Air-side level-of-service data were obtained from BACK aviation[2]. The dataset contains daily airline-specific information for all routes used in the analysis, including information on flight frequencies, departure times, flight-times (block times, thus taking into account airport congestion), aircraft types used and available seat capacity. Information on the fares for the different routes and airlines was compiled from two sources; the International Passenger Survey [13] and the fare supplement of the Official Airways Guide for 1996 [12]. Information on travel-class as well as ticket type (single or return) was taken into account in assembling the data. As is commonly the case in RP studies of air-travel choice behaviour, the resulting dataset is of highly aggregate nature, leading to problems in the estimation of the marginal utility of air-fares. For a broader discussion of this issue, see Hess and Polak [9].

3.4 Ground-access level-of-service data

For the analysis of the ground-level choice dimension, data from the National Airport Access Model (NAAM) were obtained for the base year 1999 [16]. Corresponding cost information for 1996 was produced with the help of the retail price index, while assuming that relative travel-times have on average stayed constant. This dataset contains level-of-service information for travel between 455 different travel area zones and the five airports. Six different modes are considered in the analysis; private car, rental car, public transport (rail, bus, local transport), long-distance coach, taxi and minicab (MC). It is hoped that future work can make use of a higher level of disaggregation, especially for public transport.

No combinations of access modes were considered in the present analysis, and the final mode indicated in the survey was used as the chosen mode. This is a major simplification of the actual choice process, given the high incidence of access-journeys using a combination of different modes. However, in the absence of detailed route-choice information, this simplification was unavoidable. For each mode, information was included on travel time, wait time, and the number of interchanges (where appropriate). For the cost information, a fixed one-day charge of £35 was used for rental cars (in the absence of cost-bearing party information) in addition to marginal running costs (fuel only), while fare information was used for public transport, long distance coach, taxi and minicab. For car, two specifications were retained,

[2] www.backaviation.com

one using only the marginal running costs in terms of fuel, with a second also including depreciation. Parking cost information, distinguishing between short and long-term parking separately for each airport airport, was also added.

4 Model specification

4.1 Model structure

In this paper, Generalised Extreme Value (GEV) structures [c.f. 19] were used for the representation of the correlation in the unobserved part of utility. The most basic GEV nesting approach that can be used in the analysis of air-travel choice behaviour is a simple two-level NL model, where, in the context of the present research, three main possibilities arise, using nesting along a single dimension of choice, with one nest per elementary alternative represented in that dimension of choice. As an example, the appropriate structure for the NL model using nesting by airport is shown in Figure 1, with K mutually-exclusive nests, one for each airport, and where each nest has its own nesting parameter, λ_k, allowing for different substitution patterns in the different nests. The same logic applies in the case of a two-level NL treatment of the correlation along the airline or access-mode dimensions.

The NL model can be adapted to allow for correlation along more than one dimension, by using a multi-level structure. A common example in the case of air-travel is to nest the choice of airline within the choice of airport. The structure of such a model is illustrated in Figure 2, where λ_k is the nesting parameter associated with airport nest k, and π_l is the nesting parameter associated with airline nest l.

It can be seen that six possible two-level structures arise in the present context. While NL structures can, in this form, thus be used for analysing correlations along two dimensions of choice, two important shortcomings limit their potential for the analysis of choice processes of the type described in this work.

The first is that the structures can be used for the analysis of correlation along at most two dimensions of choice. Indeed, using the example shown in Figure 2, it can be seen that, by adding in an additional level of nesting by access-mode below the airline-level, each access-mode nest would contain a single alternative, such that the structural parameter for the access-mode nest would cancel out. The same principle applies in the case of the other possible four-level structures, where, in each case, the lower level of nesting becomes redundant.

While a three-level NL model can be used to analyse the correlation along two out of the three dimensions of choice, the second shortcoming of the structure means that problems arise even with this task. In fact, it can be seen that the full extent of correlation can only be taken into account along one dimension, with a limited amount along the second dimension. Indeed, by nesting the alternatives first by airport, and then by airline, the nest for airline l inside the nest for airport k will only group together the options on airline l for that airport k. As such, the model is not able to capture correlation between alternatives using airline l at airport k_1 and alternatives using airline l at airport k_2, which is clearly a restriction. This problem also applies in the other multi-level nesting approaches.

These deficiencies of multi-level nesting structures are the motivation for the efforts made in this paper to use cross-nesting structures. In the present context, a CNL model is specified by defining three groups of nests, namely K airport nests, L airline nests and M access-mode nests, and by allowing each alternative to belong to exactly one nest in each of these groups. As such, the structure addresses both of the shortcomings described above for the three-level NL model, by being able to accommodate correlation along all three dimensions, and by doing so in a simultaneous fashion. An example of such a model is shown in Figure 3, where, in addition to the previously defined λ_k and π_l, Ψ_m is used as the structural parameter for access-mode nest m, and where the allocation parameters, governing the proportion by which an alternative belongs to each of the three nests, are not shown.

Two points merit some further attention. The above discussion has looked exclusively at nesting alternatives along one or more of the three choice-dimensions. It should be noted that an infinite number of other nesting approaches are possible. Here, one promising approach for further work is to nest low-cost carriers *against* network carriers. The other point relates to the use of GEV mixture models, which, in addition to inter-alternative correlation, also allow for random variations in tastes across respondents [c.f. 7]. In this paper, such an approach was not explored, given the prohibitive cost of estimation. In this context, and given the discussion by Hess et al. [8] about the issue of confounding between simple inter-alternative correlation and random inter-agent variations in tastes, it is thus important to note that the GEV models potentially capture a mixture of both phenomena, and as such, may overstate the extent of simple inter-alternative correlation. Although the present work does not attempt to correct this *bias*, previous work has, seemingly without exception, failed to mention this issue altogether.

4.2 Choice-set

With the use of 5 departure airports, 37 airlines, and 6 access-modes, a total of $1,110$ combinations of airports, airlines and access-modes arise. However, not all airlines operate from all airports, and the total number of airport-airline pairs is actually 54 (instead of 185), which reduces the number of alternatives (airport, airline, access-mode triplets) to 324. The number of available alternatives for specific individuals in the estimation sample ranges from 6 to 58, with a mean of 31.

4.3 Re-weighting of survey data

Given that the survey data are choice-based, some form of re-weighting needs to be performed in order for the estimation to represent population-level market shares as opposed to sample-level shares (influenced by survey quotas), thus avoiding biased results. In the present analysis, multiplicative weights were used in the specification of the log-likelihood function, where, for a given respondent, the weight is proportional to the ratio between the population weight and the sample weight for the corresponding group, and where group allocation was based on a host of criteria, dominated by route and airline choice. In the case of the GEV models discussed in this paper, the correction can in fact be performed at the level of the alternative specific constants (ASC) when a full set of constants is used [c.f. 4], but given the unconventional specification of constants used in this analysis (c.f. Section 4.4), preference was given to a weighting approach, in conjunction with the use of robust estimators.

4.4 Specification of constants

An important question arises with regards to the specification of the constants in the model. In one-dimensional choice processes, a single ASC is associated with each alternative, with all but one of the constants being estimated. In the present application, such an approach led to an overloading of the model with constants. To address this deficiency, an alternative specification was attempted, using a single constant for each airport-airline pair, with separate constants for the access-mode dimension, hence associating two constants with each *combined* alternative. Separate experiments showed that this approach led to very significant gains in model performance when compared to the use of three separate sets of constants, suggesting some interaction between choice dimensions. Attempts to use airport-access constants in combination with separate constants in the airline dimension also led to gains in model fit, which were however less significant than those obtained with the airport-airline specification, which was thus retained.

4.5 Non-linearities

The final point that deserves some discussion is the way in which explanatory variables enter the utility function, in terms of the use of non-linear transforms. In the present analysis, the log-transform was used for this purpose. A preliminary analysis was conducted to determine which attributes benefited from the use of a non-linear specification. These experiments showed that important gains in model performance could be obtained by using a log-transform for flight frequency, flight time, in-vehicle access-time (IVT), and access-cost, such that this approach was adopted. Any remaining attributes were treated in a linear fashion.

5 Model estimation

This section discusses the findings of the modelling analysis. It starts with a discussion of the specification of utility in Section 5.1. This is followed by a presentation of the results for MNL (Section 5.2), NL (Section 5.3) and CNL (Section 5.4). The section concludes with a comparison of the substantive results across models in Section 5.5. All models presented in this paper were estimated with BIOGEME [3].

5.1 Utility specification

A comprehensive set of variables were used in the initial modelling analysis. These included attributes relating to the air journey (frequency, fare, flight time, aircraft type, seat capacity, on-time performance of the airport and airline) and the access journey (access-cost, in-vehicle access-time, out-of-vehicle access-time, wait-time, number of interchanges, parking cost). No treatment of the distribution of departure times was used in the present analysis, for data reasons. In the absence of frequent flier information, attempts were made to account for airline allegiance by including a UK-airline dummy variable in the models for resident travellers, and a foreign-airline dummy variable in the models for visiting travellers. No further gains could be made by using separate dummy variables for all different foreign nationalities. No information on past choices was available, such that a treatment of "airport-allegiance" was not possible. Finally, attempts to model further interactions with socio-demographic attributes aside from purpose and residency status were not successful. This is characterised notably by the absence of an income effect, which can be blamed partly on the discrete nature of the income information. Taste-coefficients along the access-mode dimension were generic, and not mode-specific.

Before proceeding to the actual estimation results, some generic conclusions can be presented. As such, no effect could be identified for parking cost (possibly due to the absence of cost-bearing information), seat capacity, out-of-vehicle access-time, wait-time, on-time performance, and the number of ground-level interchanges. Furthermore, aircraft size, in the form of a dummy variable for turboprop planes, showed no effect; here however, the highly correlated flight time attribute had a significant negative effect across models. Furthermore, a significant effect of air-fare could only be identified for visiting leisure travellers, while allegiance to the national carrier played a role only for visiting business travellers. The inability to estimate a consistent fare effect across population segments can be explained mainly on the basis of the low quality of the fare data, where similar problems have arisen in other research using RP survey data [e.g 10]. Finally, the analysis showed that the use of the combined fuel and depreciation cost for car journeys is preferable to the use of fuel cost on its own.

5.2 MNL models

The estimation results for the four MNL models are shown in Table 1. The results show consistent negative effects of increases in access-cost, flight-time and in-vehicle access-time, with positive effects of increases in flight frequency (significant at the 94% level for visiting business travellers). In each case, a log-transform was used. Additionally, there is a negative effect of increases in flight-fare (linear) for visiting leisure travellers, and a positive dummy variable is associated with non-UK carriers for visiting business travellers, though this is significant only at the 90% level.

5.3 NL models

In this section, we look at the three sets of NL models fitted to the four subsamples, using nesting by airport, airline and access-mode respectively. The specification of utility used in the NL models was the same as in the MNL models. As such, this discussion centres primarily on the conclusions in terms of nesting.

5.3.1 NL model using nesting by airport

The results for the first group of NL models, which use nesting by airport, are shown in Table 2. The results show that the four NL models lead to improvements in log-likelihood (LL) over their MNL counterparts by 49.2, 120.7, 97.7 and 151.3 units respectively, at the cost of 4 additional parameters in the case of resident business travellers, and 3 additional parameters in the remaining three models. In each case, the improvement in model fit is statistically significant, with χ^2 p-values of 0 for the associated likelihood-ratio (LR) tests. Several parameters have experienced a drop in significance when compared to the MNL model, notably the fare-coefficient for visiting leisure travellers, which is now only significant at the 91% level.

In terms of nesting conclusions, a consistent pattern emerges. The nesting parameter for LHR needs to be constrained to a value of 1 across all models, indicating no heightened correlation in the unobserved utility terms between alternatives associated with LHR along the airport dimension. A similar conclusion applies for LCY, where, although for resident business travellers, the original value does not exceed 1, it is not significantly different from 1. For the remaining three airports, the values are consistently below 1, indicating heightened correlation, although, in some cases, the significance level is below the usual 95% limit.

The lowest values for the nesting parameters, and hence the highest levels of correlation, are observed for LTN and STN (with the exception of LTN for visiting leisure travellers). These two airports are different from the remaining three in terms of their route network, and in terms of being used extensively by low-cost airlines (even back in 1996). Some of these characteristics are not captured in the observed part of utility, explaining the high levels of correlation between alternatives in the unobserved part of utility.

5.3.2 NL model using nesting by airline

The results for the second group of NL models, which use nesting by airline, are shown in Table 3. The results show that the four NL models lead to improvements in LL over their MNL counterparts by 74.6, 183.9, 173.4 and 288.4 units respectively, at the cost of 19 additional parameters in the two models for resident travellers, 16 additional parameters for visiting business travellers, and 24 parameters for visiting leisure travellers. Again, all four improvements are statistically significant, with χ^2 p-values of 0 for the associated LR tests. Again, there are some changes in parameter significance, with the significance of the fare-coefficient for visiting leisure travellers decreasing to the 87% level.

Of the 37 nesting parameters, 6 had to be constrained to 1 in each of the four models, and as such, are not listed in Table 3. These relate to airlines A10, A23, A26, A33, A35, and A36. In addition to the six overall constraints, a number of other nesting parameters initially took on unacceptable values in some of the population segments. As such, out of the 37 possible parameters, 18 had to be constrained in the two models for residents, along with 21 in the model for visiting business travellers, and 13 in the model for visiting leisure travellers. A large number of the estimated structural parameters are not significantly different from a value of 1, but additional constraints led to significant drops in model performance.

5.3.3 NL model using nesting by access-mode

The results for the final group of NL models, which use nesting by access-mode, are shown in Table 4. The results show that the four NL models lead to improvements in LL over their MNL counterparts by 128.6, 45.3, 163.5 and 131.1 units respectively, at the cost of 5 additional parameters in the model for resident business travellers, and 4 additional parameters in the remaining three models. Again, all four improvements are statistically significant, with χ^2 p-values of 0 for the associated LR tests.

The common observation across models is that the structural parameter associated with the public transport nest needs to be constrained to a value of 1. Although this suggests a lack of correlation between public transport alternatives, the low level of disaggregation along the public transport dimension could play a role in this (c.f. Section 3.4), and more similarities could be expected in subgroups of public transport modes. Consistently low values for the structural parameters are obtained for car and taxi, showing high correlation within these nests. Again, not all estimated structural parameters are significantly different from a value of 1, but additional constraints led to significant drops in model performance.

5.3.4 Discussion

The presentation of the NL results has shown that each of the three nesting approaches offers significant improvements in model fit over the corresponding MNL model, across the four population groups.

At the same time, the analysis has revealed important differences across population segments in terms of the optimal two-level nesting structures. As such, using the adjusted ρ^2 measure as a means of comparison, the model using nesting by access-mode leads to the best performance for resident business travellers, while the model using nesting by airline leads to the best performance in the two leisure groups. For visiting business travellers, the performance of the two models is indistinguishable, with the higher number of parameters for the model using nesting by airline nullifying its log-likelihood advantage.

5.4 CNL models

We next turn our attention to the estimation of the CNL models, where the base specification of utility was again the same as that used for the MNL and NL models.

In the present context, a total of 48 nests were used in the CNL models (5 airports, 37 airlines, and 6 access-modes). Aside from leading to the use of 48 separate nesting parameters (to allow for differential levels of correlation in different nests), this leads, in the presence of a choice-set of 324 combined alternatives, to a total of 972 allocation parameters (324 along each dimension). As each alternative is associated with exactly one airport, one airline, and one access-mode, only one allocation parameter along each of the three dimension is not constrained a priori to zero for a given alternative. From this, it can also be seen that, given the condition that the allocation parameters for each alternative sum to 1, a total of 648 can be identified.

A preliminary analysis showed that, although the estimation of the allocation parameters leads to gains in model fit, these are not statistically significant, given the huge cost in terms of the number of parameters. Additionally, the estimation of the allocation parameters leads to major issues with parameter identification and very significant increases in computational cost. As such, the decision was taken to constrain all non-zero allocation parameters to a value of $\frac{1}{3}$, such that an alternative is associated in equal parts with an airport, an airline, and an access-mode. With the use of fixed allocation parameters, it is not immediately clear how the CNL model can reduce to one of the three NL models, although an approximation can be obtained when the structural parameters along two dimensions reduce to a value of 1. Given this complication, nested LR tests were replaced by the adjusted ρ^2 statistic.

The results of the CNL models are summarised in Table 5 (main results, plus structural parameters for airport and access-mode dimensions) and Table 6 (structural parameters for airline dimension). In terms of model performance, it can be seen that the four CNL models give a better fit to the data (in terms of adjusted ρ^2) than the MNL model, or any of the three NL structures. Again, however, there are some significant differences across the four population segments. Indeed, for resident business travellers, the improvement in LL offered by the CNL model over the MNL model is 35% bigger than the combined improvement offered by the three NL models. In the remaining three models, the improvement is smaller than the combined gain in LL by the three NL models, but it is in each case still significantly larger than the average improvement offered by the three NL models (twice as large in the two models for visitors). In fact, it be seen that the only group where the performance of the CNL model is not convincing is that for resident leisure travellers. Here, the performance of the model is still better than that of the NL models using nesting by airport and by access-mode, but there is little gain in performance to be obtained when compared to the model using nesting by airline. This can be seen to be a direct result of the higher number of constraints (on the structural parameters) required along the airline dimension for the CNL model in this population segment. Again, several parameters experience a drop in significance when compared to the MNL model, with the biggest drop occurring in the case of the frequency coefficient for resident business travellers. A number of the estimated structural parameters are again not statistically different from 1, while others are very close to zero. Here, it should be noted that it was not possible to produce a reliable value for the standard error of π_{A10}[3] in the model for resident business travellers, despite the use of a robust estimator. This is an indication of the complexity of the model that was estimated here.

In closing, it can be seen that the CNL model does have the potential to offer gains in performance when compared to the three two-level NL structures. Furthermore, given the problems with using multi-level NL structures[4], the CNL model has clear conceptual advantages. Finally, further gains in performance can be expected with the use of a flexible formulation (i.e. parameterisation) of the allocation parameters.

5.5 Comparison of substantive results

The final step of the analysis is concerned with a comparison of the actual substantive results across population groups and across models. To allow for a consistent comparison, only coefficients estimated across all four

[3] The estimated value of π_{A10} is arbitrarily close to zero.

[4] No multi-level NL structures were estimated in this paper, but the observations made in Section 4 apply.

population subgroups should be involved in these comparisons. In the presence of four such coefficients, namely the marginal utilities of changes in (the logarithms of) in-vehicle access-time, access-cost, flight frequency, and flight-time, three trade-offs were used. These are the willingness to accept increases in access-cost in return for decreases in access-time (i.e. VTTS), the willingness to accept increases in access-time in return for increases in flight frequency, and the willingness to accept increases in access-time in return for decreases in flight-time.

Given the use of the log-transform for each of the coefficients involved in the trade-offs, the ratio between coefficients has to be multiplied by the inverse ratio of the actual attribute values, where a mean of ratios approach was used instead of a ratio of means approach, avoiding biased results. This approach also yields respondent-specific trade-offs, allowing the calculation of a set of statistics for the distribution of the trade-offs, where it should be noted that these variations are an effect of the varying values for the concerned attributes, and do not as such give variations in tastes across respondents, but rather give an indication of the varying levels of trade-offs under different market conditions.

The results of the calculation of the trade-offs are summarised in Table 7 for the VTTS, Table 8 for the trade-off between flight-time and in-vehicle access-time and Table 9 for the trade-off between flight frequency and in-vehicle access-time.

The first observation that can be made is that there are some variations across the five models in the calculated values for the different trade-offs. In the first trade-off (VTTS), and the third trade-off (frequency *vs* IVT), the values in the MNL model and the two NL models using nesting by airport and by airline are generally roughly similar, while those produced with the NL model using nesting by access-mode and the CNL model are quite different. The second trade-off, between flight-time and in-vehicle access-time, shows a different trend, where the biggest outliers are this time generated by the model using nesting by airline, in the two models for residents and in the model for visiting leisure travellers, while the observations for visiting business travellers in the model using nesting by access-mode and the CNL model need to be put into context by noting the lower significance of the flight-time coefficient in these two models (c.f. Table 4 and Table 5). The differences in the trade-offs across models are possibly more significant than could have been expected on the basis of the small differences in model fit, highlighting the flatness of the LL function, but also stressing the differences between model structures in terms of implied choice behaviour.

Aside from comparing the calculated trade-offs across the five model structures, it is of interest to compare their values across the four population subgroups. Here, a major issue arises. Indeed, the calculation of the trade-off between access-time and access-cost produces a counter-intuitive result, suggesting that, for residents, the VTTS is higher for leisure travellers than for business travellers, depending on the model considerably so (ranging from 19% to 73%). On the other hand, for visitors, the results consistently show higher VTTS for business travellers than for leisure travellers, with differences ranging from 61% to 101%.

Clearly, this issue needs to be addressed. There are two possible reasons for an underestimated VTTS; an underestimation of the marginal utility of travel-time changes, and an overestimation of the marginal utility of access-cost changes. In the present analysis, it seems likely that that both factors play a role. This insight is partly gained from a separate analysis carried out to look at the choice-behaviour by respondents on the actual observed access-journey. For this, the market-shares of different access-modes were calculated from the survey data, using a higher level of disaggregation than was possible in the actual modelling analysis. This analysis produced three main findings:

- Business travellers are more likely to use premium PT modes for their access-journey than leisure travellers, with, for example, a lower market share for the Tube.

- On access-journeys using combinations of modes, there is, on the non-final stages, a higher market share for taxis in the case of business travellers than in the case of leisure travellers, and a much lower market share for the Tube.

- Business travellers using car as their access-mode are more likely to make use of premium parking facilities, which are more expensive, but which are located closer to the airport.

These three findings apply for residents as well as for visitors, but seem to play a bigger role in the former group. In combination, these three observations can be used to explain the counter-intuitive findings in the calculation of trade-offs. Indeed, it should be remembered that, for data reasons, the present study uses highly aggregate PT data, and does not differentiate between *standard* and *premium* modes. At the same time, again for data reasons, the chosen mode for a given traveller is defined to be the mode used in the final part of the access-journey. Finally, parking cost could not be included in the models. As such, it can be seen that, with the above three observations, the level-of-service data used in model estimation understates the access-cost for journeys by business travellers, and also overstates the access-time. This clearly has an effect on the estimated coefficients, leading to an underestimation of the access-time coefficient, and an overestimation of the access-cost coefficient. Additionally, it should be remembered that the VTTS is in the present case calculated as $\frac{\beta_{AT}}{\beta_{AC}}\frac{AC}{AT}$. At the same time as the estimates lead to a lower than warranted ratio of $\frac{\beta_{AT}}{\beta_{AC}}$, the *biased* level-of-service data leads to an underestimation of the ratio $\frac{AC}{AT}$, which further underestimates the VTTS.

To illustrate the differences, the ratio between the access-cost and the access-time variables for the actual observed journey was calculated, using the available data. This yielded mean values of 25.84 pence per minute for resident business travellers, with a corresponding value of 21.53 pence per minute for resident leisure travellers. Although this does suggest a slightly higher spending rate for business travellers, the differences are small. On the other hand, for visitors, the corresponding values are 35.56 pence per minute for business travellers, and 20.29 pence per minute for leisure travellers. It is conceivable that the spending rate for resident business travellers is indeed lower than for visiting business travellers, for example due to a lower reliance on taxis (and a higher reliance on cheaper minicabs), but this is unlikely to be on the scale indicated by the data. This argument is supported by the findings for leisure travellers, where the spending rates for residents and visitors are very similar. As such, this brief analysis does indeed suggest some bias in the data, which could explain the counter-intuitive findings.

Clearly, this issue leads to unreliable estimates of the trade-offs, making them inapplicable for use in cost-benefit analysis and forecasting. At the same time, it is not clear what effect, if any, this bias in the access-journey level-of-service data has on the results in terms of model structure. This can only be addressed with the reanalysis of the models on more disaggregate access-journey level-of-service data, which is an important avenue for further research.

Despite the above discussion of the likely bias in the estimated trade-offs, several interesting observations can nevertheless be made on the basis of Tables 7, 8 and 9.

The first observation is that the estimated VTTS, even for leisure travellers (where the issue of biased data plays a lesser role), are lower than those reported in previous studies of airport choice behaviour, where, as an example, Pels et al. [15] produce VTTS for business travellers between \$1.97 and \$2.90 per minute in the SF-bay area. While there could clearly be differences across regions[5], it seems more likely that the use of a non-linear specification is the main reason for the lower values; indeed, much higher values, together with a lower model fit, were obtained when using a linear specification. While previous research in airport-choice modelling has generally made use of a log-transform for flight-frequency, access-time and access-cost have usually been treated in a linear fashion, which could have caused the high implied VTTS.

The second observation relates to the relative sensitivity to access-time and flight-time. Here, the findings for visitors can again be judged to be more reliable, and would indicate that, while business travellers are relatively equally sensitive to access-time and flight-time, leisure travellers are far more sensitive to flight-time. Here, the correlation between flight-time and aircraft-type plays an important role, and the lower objection to using turboprop flights by business travellers than by leisure travellers can help to explain the results. Additionally, the average flight-time is longer for leisure travellers than for business travellers, further reducing the appeal of turboprop flights. Here, there is little opportunity for comparing the results to those obtained in other studies, where this trade-off is often not available.

Finally, for the willingness to accept increases in access-time in return for increases in flight frequency, the values are higher for leisure travellers than for business travellers, which is a reflection of lower VTTS for leisure travellers. This finding applies to residents as well as visitors (c.f. Table 9), suggesting that the main source of bias in the VTTS estimates could be access-cost, rather than access-time. The low value for the trade-off for resident business travellers in the CNL model needs to be put into context by noting the high associated standard error (c.f. Table 5). The actual implied values equate to between 10% and 30% of the average observed access-time, and as such, are possibly on the low end of the scale.

6 Summary and Conclusions

This paper has described an analysis of the combined choice of airport, airline and access-mode for passengers departing from the London area, using three different types of GEV structures; MNL, two-level NL, and CNL.

In common with most previous studies, the analysis has shown that access-time is a prime determining factor in travellers' choices of departure airport, while flight frequency, access-cost and flight-time also play a role.

[5] Here, no comparable values for other studies involving the London airports were available.

At this point, it should be noted that the frequency coefficient can be seen as a proxy for visibility and scheduling convenience, while the flight-time coefficient can also be seen as a proxy for smaller aircraft, and for on-time performance, given that the block-time incorporates taxi-time, and hence takes into account congestion. As in many previous studies, it was not possible to estimate a consistent significant effect of air-fare, nor of airline-allegiance, a fact that is down to the general low quality of the level-of-service data for the associated attributes.

In terms of model performance, all attempted nesting approaches lead to significant gains in fit. The use of two-level NL models, which allow for the treatment of correlation along a single dimension of choice, lead to improvements in performance over the MNL model, and show differences across population groups in terms of the optimal nesting structure. The theoretical advantages of the CNL model are reflected in the estimation results, showing gains in model performance and insights into choice behaviour, suggesting that this model form can indeed serve as a valuable tool in the analysis of air-travel choice behaviour.

The study did however produce a counter-intuitive result in the models for residents, showing lower VTTS for business travellers than for leisure travellers. While it was possible to identify the most likely source for this bias, the absence of more detailed access-journey level-of-service data prevented the estimation of more refined models. As such, it remains to be seen what effect, if any, the data problems had on the conclusions in terms of model structure. The relatively consistent results across the four population segments in terms of the advantages of advanced model structures however somewhat increase the confidence in these findings.

A number of avenues for further research can be identified, not least of which the use of more advanced model structures, allowing jointly for cross-nesting, continuous deterministic and random taste heterogeneity. Further refinement of the auxiliary datasets can also be expected to lead to gains in model performance. Finally, aside from accounting for correlation between alternatives sharing a given airport, airline or access-mode (or a combination thereof), it is also of interest to test for correlation between alternatives at different, yet comparable airports (e.g. STN & LTN), or different airlines and access-modes.

Acknowledgements

The analysis presented in this paper was supported by the Civil Aviation Authority (CAA) and the UK Department for Transport (DfT). However, the analysis was carried out independently, such that the conclusions reached in the paper do not necessarily represent the views of the CAA or the DfT.

References

[1] Ashford, N. and Bencheman, M. [1987], 'Passengers' choice of airport: an application of the Multinomial Logit model', *Transportation Research Record* **1147**, 1–5.

[2] 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 Part B: Methodological* **38**(10), 889–904.

[3] Bierlaire, M. [2003], *BIOGEME: a free package for the estimation of discrete choice models*, Proceedings of the 3^{rd} Swiss Transport Research Conference, Monte Verita, Ascona, Switzerland.

[4] Bierlaire, M., Bolduc, D. and McFadden, D. [2003], *Characteristics of Generalized Extreme Value Distributions*, Technical report, Econometrics Laboratory, University of California, Berkeley, CA.

[5] Brooke, A. S., Caves, R. 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.

[6] CAA [1996], *UK airports Passenger survey, CAP 677 - Passengers at Birmingham, Gatwick, Heathrow, London City, Luton, Manchester and Stansted in 1996*, The Civil Aviation Authority, London.

[7] Hess, S., Bierlaire, M. and Polak, J. W. [2005*a*], Capturing taste heterogeneity and correlation structure with Mixed GEV models, *in* R. Scarpa and A. Alberini, eds, 'Applications of Simulation Methods in Environmental and Resource Economics', Springer Publisher, Dordrecht, The Netherlands, pp. 55–76.

[8] Hess, S., Bierlaire, M. and Polak, J. W. [2005*b*], *Confounding between substitution patterns and random taste heterogeneity*, CTS Working Paper, Centre for Transport Studies, Imperial College London.

[9] Hess, S. and Polak, J. W. [2004], *On the use of Discrete Choice Models for Airport Competition with Applications to the San Francisco Bay area Airports*, paper presented at the 10^{th} triennial *WCTR* conference, Istanbul, Turkey.

[10] Hess, S. and Polak, J. W. [2005*a*], *Accounting for random taste heterogeneity in airport-choice modelling*, Transportation Research Record, forthcoming.

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

[12] OAG [1996], *World Airways Guide Fares Supplement 1996*, Reed Elsevier, Dunstable, United Kingdom.

[13] ONS [1996], *International Passenger Survey*, The Office for National Statistics, London.

[14] 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.

[15] 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 Part A: Policy and Practice* **37**(1), 71–83.

[16] Scott Wilson Kirkpatrick [1999], *National airport accessibility model: model status report*, prepared for the Department for Transport by Scott Wilson Kirkpatrick & Co Ltd, Basingstoke, United Kingdom.

[17] Skinner, R. E. J. [1976], 'Airport choice: An empirical study', *Transportation Engineering Journal* **102**(4), 871–883.

[18] 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.

[19] Train, K. [2003], *Discrete Choice Methods with Simulation*, Cambridge University Press, Cambridge, MA.

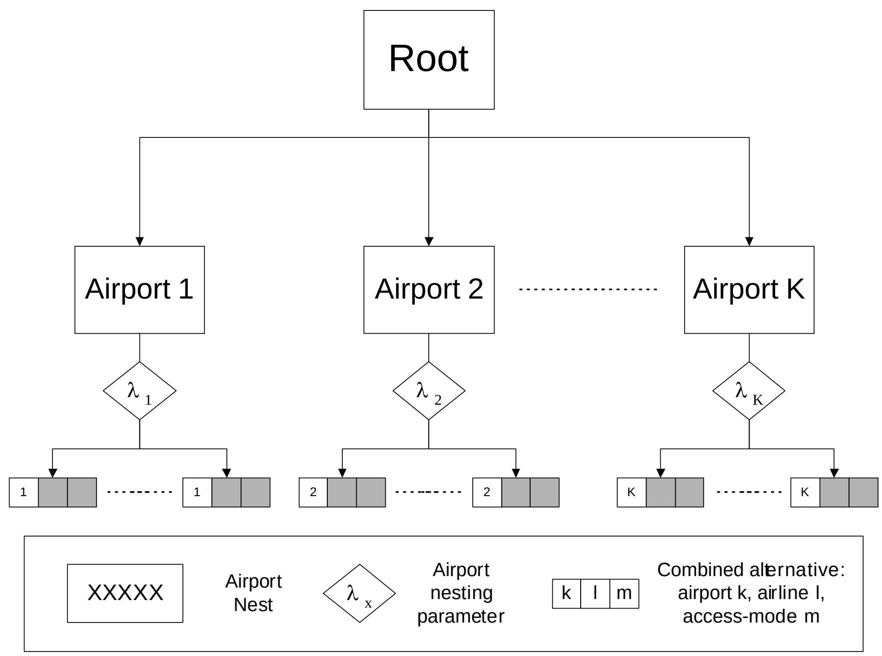

Figure 1: Structure of two-level NL model, using nesting along airport-dimension

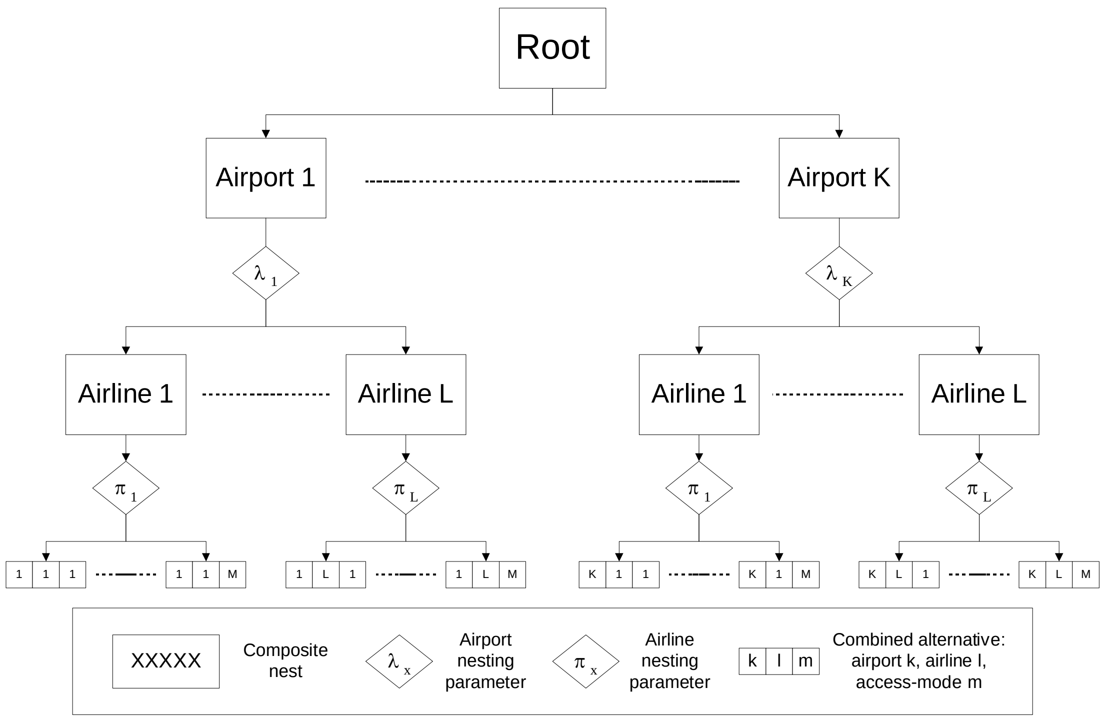

Figure 2: Structure of three-level NL model, using nesting along airport-dimension and airline-dimension

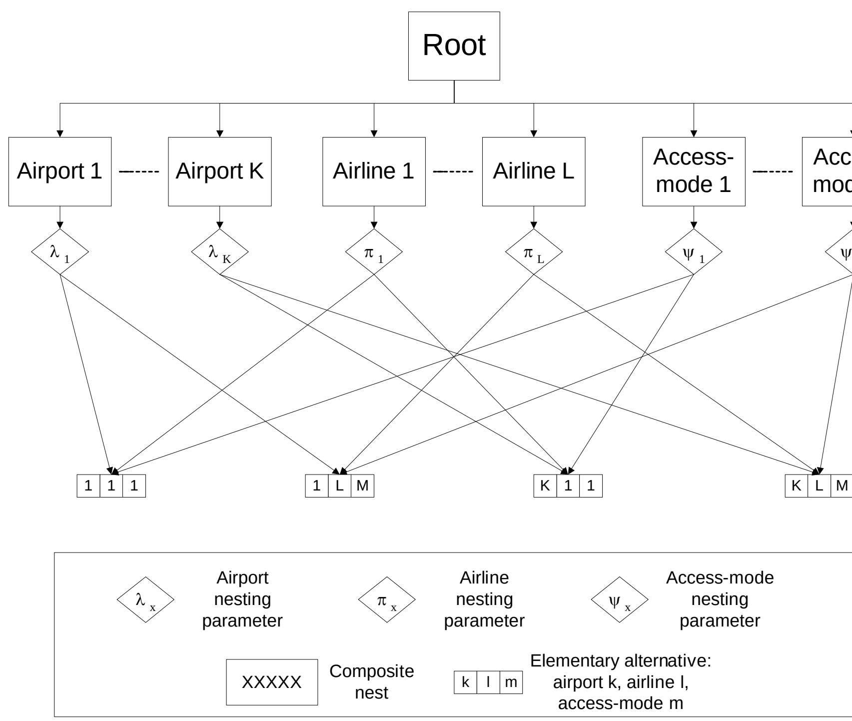

Figure 3: Structure of CNL model for the joint analysis of correlation along the airport, airline and access-mode dimensions

	Resident				Visitor			
	Business		Leisure		Business		Leisure	
Observations	6,706		8,269		7,207		9,667	
Parameters	55		57		57		58	
Final LL	-14945.3		-17627.1		-15278.1		-20553.8	
Adjusted ρ^2	0.3416		0.3529		0.3549		0.3418	
	Est.	t-stat.	Est.	t-stat.	Est.	t-stat.	Est.	t-stat.
$\beta_{\text{flight}-\text{fare}}$	-	-	-	-	-	-	-0.0026	-1.93
$\beta_{\text{LN}(\text{access}-\text{cost})}$	-1.2831	-8.05	-0.9083	-7.20	-0.9004	-7.74	-0.7097	-5.04
$\beta_{\text{LN}(\text{flight}-\text{time})}$	-2.2963	-3.17	-2.7678	-3.18	-2.1878	-2.01	-4.3711	-4.48
$\beta_{\text{LN}(\text{frequency})}$	0.5641	2.42	0.9776	4.73	0.5070	1.85	0.7024	3.22
$\beta_{\text{LN}(\text{IVT}^\dagger)}$	-1.4440	-6.21	-1.6898	-10.75	-1.6319	-10.67	-1.4025	-7.48
$\delta_{\text{national carrier}}$	-	-	-	-	0.4653	1.69	-	-

† IVT = in-vehicle access-time

Table 1: MNL estimation results

	Resident				Visitor			
	Business		Leisure		Business		Leisure	
Observations	6,706		8,269		7,207		9,667	
Parameters	59		60		60		61	
Final LL	-14896.1		-17506.4		-15180.4		-20402.5	
Adjusted ρ^2	0.3436		0.3572		0.3589		0.3465	
	Est.	t-stat.	Est.	t-stat.	Est.	t-stat.	Est.	t-stat.
$\beta_{\text{flight}-\text{fare}}$	-	-	-	-	-	-	-0.0021	-1.73
$\beta_{\text{LN(access}-\text{cost)}}$	-1.1807	-7.83	-0.8556	-8.01	-0.8370	-8.48	-0.7292	-5.94
$\beta_{\text{LN(flight}-\text{time)}}$	-2.1002	-2.91	-2.0361	-2.48	-1.9097	-1.86	-3.6898	-3.90
$\beta_{\text{LN(frequency)}}$	0.5446	2.40	0.9460	4.95	0.4883	1.86	0.7335	3.59
$\beta_{\text{LN(IVT}^{\dagger})}$	-1.4610	-6.75	-1.5081	-10.54	-1.5896	-11.15	-1.2530	-7.42
$\delta_{\text{national carrier}}$	-	-	-	-	0.4223	1.57	-	-
	Est.	t-stat.	Est.	t-stat.	Est.	t-stat.	Est.	t-stat.
λ_{LCY}	0.8730	0.43	1.00	-	1.00	-	1.00	-
λ_{LGW}	0.8266	1.90	0.7205	4.16	0.8114	1.68	0.6939	4.28
λ_{LHR}	1.00	-	1.00	-	1.00	-	1.00	-
λ_{LTN}	0.5470	2.14	0.7029	1.17	0.7312	2.10	0.8377	1.36
λ_{STN}	0.7568	1.27	0.6519	2.82	0.4415	3.64	0.6773	1.55

† IVT = in-vehicle access-time

T-statistics calculated with respect to 0 for taste coefficients, and with respect to 1 for structural parameters.

Table 2: NL estimation results using nesting by airport

	Resident		Visitor	
	Business	Leisure	Business	Leisure
Observations	6,706	8,269	7,207	9,667
Parameters	74	76	73	82
Final LL	-14870.7	-17443.2	-15104.7	-20265.4
Adjusted ρ^2	0.3441	0.3590	0.3615	0.3502

	Est.	t-stat.	Est.	t-stat.	Est.	t-stat.	Est.	t-stat.
$\beta_{\text{flight-fare}}$	-	-	-	-	-	-	-0.0021	-1.53
$\beta_{\text{LN(access-cost)}}$	-1.1331	-7.04	-0.7850	-6.92	-0.8011	-8.12	-0.6316	-4.86
$\beta_{\text{LN(flight-time)}}$	-2.3415	-3.44	-2.8067	-3.45	-2.1560	-2.14	-4.4597	-4.93
$\beta_{\text{LN(frequency)}}$	0.5716	2.50	0.8593	4.17	0.4707	1.86	0.6014	3.00
$\beta_{\text{LN(IVT}^{\dagger}\text{)}}$	-1.3946	-6.27	-1.4594	-10.04	-1.5659	-11.38	-1.1370	-6.50
$\delta_{\text{national carrier}}$	-	-	-	-	0.4328	1.52	-	-

	Est.	t-stat.	Est.	t-stat.	Est.	t-stat.	Est.	t-stat.
π_{A1}	0.6113	1.36	0.9511	0.15	0.7126	1.32	0.9222	0.32
π_{A2}	1.0000	-	1.0000	-	1.0000	-	0.8736	0.21
π_{A3}	1.0000	-	0.2609	1.92	0.5595	1.45	1.0000	-
π_{A4}	0.8940	1.29	0.8333	2.17	1.0000	-	0.8259	1.35
π_{A5}	1.0000	-	1.0000	-	0.7268	1.90	0.6572	2.15
π_{A6}	0.7795	0.44	0.2099	1.52	1.0000	-	1.0000	-
π_{A7}	1.0000	-	0.9025	0.36	1.0000	-	0.8585	0.51
π_{A8}	0.6339	0.84	0.6199	1.04	0.3981	0.63	0.4927	1.97
π_{A9}	1.0000	-	0.4089	3.70	1.0000	-	1.0000	-
π_{A11}	0.7186	1.46	0.6923	2.38	0.5391	3.03	0.5260	2.34
π_{A12}	0.6185	2.02	0.9020	0.43	1.0000	-	0.2869	2.80
π_{A13}	1.0000	-	0.6950	1.84	1.0000	-	0.8392	0.63
π_{A14}	0.7758	0.39	1.0000	-	0.4072	2.01	0.7271	0.79
π_{A15}	0.5633	1.48	0.3234	3.57	0.4669	1.04	0.2337	3.09
π_{A16}	1.0000	-	1.0000	-	0.8942	1.33	0.8575	0.93
π_{A17}	0.8992	0.37	0.7721	0.60	0.8096	1.21	1.0000	-
π_{A18}	0.6091	0.93	1.0000	-	1.0000	-	1.0000	-
π_{A19}	0.7654	0.41	0.4931	1.53	0.7494	0.67	0.5711	1.60
π_{A20}	0.4341	2.50	1.0000	-	0.8362	0.58	0.8649	0.46
π_{A21}	0.4869	1.77	0.5568	1.61	0.1316	2.48	0.4243	2.34
π_{A22}	1.0000	-	1.0000	-	1.0000	-	0.7133	0.91
π_{A24}	1.0000	-	1.0000	-	0.8487	0.58	1.0000	-
π_{A25}	1.0000	-	1.0000	-	1.0000	-	0.8156	0.62
π_{A27}	0.7238	1.15	1.0000	-	1.0000	-	0.8578	0.66
π_{A28}	0.8700	0.76	1.0000	-	1.0000	-	1.0000	-
π_{A29}	1.0000	-	1.0000	-	1.0000	-	0.6307	1.85
π_{A30}	0.3878	2.67	0.5162	1.48	1.0000	-	0.7903	1.22
π_{A31}	0.6622	1.42	0.5528	2.40	0.4298	3.40	0.5109	3.05
π_{A32}	0.6874	0.97	0.8226	0.67	0.2730	3.14	0.6720	0.78
π_{A34}	1.0000	-	0.6940	2.28	0.6711	1.50	0.5483	2.30
π_{A37}	0.4285	0.85	0.7507	1.45	1.0000	-	0.8065	0.44

† IVT = in-vehicle access-time

T-statistics calculated with respect to 0 for taste coefficients, and with respect to 1 for structural parameters.

Table 3: NL estimation results using nesting by airline

	Resident				Visitor			
	Business		Leisure		Business		Leisure	
Observations	6,706		8,269		7,207		9,667	
Parameters	60		61		61		62	
Final LL	-14816.7		-17581.8		-15114.6		-20422.7	
Adjusted ρ^2	0.3470		0.3544		0.3616		0.3458	
	Est.	t-stat.	Est.	t-stat.	Est.	t-stat.	Est.	t-stat.
$\beta_{\text{flight-fare}}$	-	-	-	-	-	-	-0.0021	-2.14
$\beta_{\text{LN(access-cost)}}$	-1.0197	-7.23	-0.7841	-5.83	-0.8258	-7.61	-0.6838	-5.27
$\beta_{\text{LN(flight-time)}}$	-1.4941	-2.95	-2.2619	-2.87	-1.3507	-1.72	-3.4752	-4.15
$\beta_{\text{LN(frequency)}}$	0.3196	1.96	0.8227	4.28	0.3746	1.93	0.5419	2.94
$\beta_{\text{LN(IVT}^\dagger)}$	-0.9553	-3.27	-1.5280	-8.95	-1.3575	-9.36	-1.1774	-6.99
$\delta_{\text{national carrier}}$	-	-	-	-	0.3869	1.89	-	-
	Est.	t-stat.	Est.	t-stat.	Est.	t-stat.	Est.	t-stat.
Ψ_{Car}	0.6062	2.56	0.7815	2.62	0.7553	2.11	0.7145	3.10
Ψ_{Hire}	0.3700	2.40	0.9244	0.21	0.4598	3.66	1.00	-
Ψ_{LDC}	0.7635	0.69	1.00	-	1.00	-	0.5954	2.63
Ψ_{MC}	0.5778	2.15	0.9548	0.41	0.5081	4.59	0.7206	1.92
Ψ_{PT}	1.00	-	1.00	-	1.00	-	1.00	-
Ψ_{Taxi}	0.6356	2.53	0.7986	1.77	0.7156	2.44	0.6467	2.32

† IVT = in-vehicle access-time

T-statistics calculated with respect to 0 for taste coefficients, and with respect to 1 for structural parameters.

Table 4: NL estimation results using nesting by access-mode

	Resident				Visitor			
	Business		Leisure		Business		Leisure	
Observations	6,706		8,269		7,207		9,667	
Parameters	91		74		89		87	
Final LL	-14603.3		-17437.6		-14988.2		-20142.9	
Adjusted ρ^2	0.3551		0.3592		0.3658		0.3540	
	Est.	t-stat.	Est.	t-stat.	Est.	t-stat.	Est.	t-stat.
$\beta_{\text{flight-fare}}$	-	-	-	-	-	-	-0.0020	-1.39
$\beta_{\text{LN(access-cost)}}$	-0.9911	-8.02	-0.7901	-5.87	-0.7975	-8.62	-0.6708	-6.92
$\beta_{\text{LN(flight-time)}}$	-1.4201	-3.16	-1.8270	-2.50	-1.3789	-1.71	-3.5148	-3.21
$\beta_{\text{LN(frequency)}}$	0.2453	1.10	0.9027	4.60	0.5306	1.61	0.5523	2.56
$\beta_{\text{LN(IVT}^\dagger)}$	-1.0718	-9.19	-1.5515	-12.04	-1.4368	-11.14	-1.0552	-6.67
$\delta_{\text{national carrier}}$	-	-	-	-	0.4081	1.35	-	-
	Est.	t-stat.	Est.	t-stat.	Est.	t-stat.	Est.	t-stat.
λ_{LCY}	0.5412	1.23	1.0000	-	0.8346	0.78	1.0000	-
λ_{LGW}	0.6177	2.63	0.1050	0.67	0.1000	0.60	0.0783	0.25
λ_{LHR}	1.0000	-	1.0000	-	1.0000	-	1.0000	-
λ_{LTN}	0.2627	2.95	0.0961	4.53	0.2644	2.47	0.6873	1.74
λ_{STN}	0.2608	2.18	0.5311	2.04	0.1806	2.28	0.1603	0.30
	Est.	t-stat.	Est.	t-stat.	Est.	t-stat.	Est.	t-stat.
Ψ_{Car}	0.0402	-	0.7250	0.98	0.5778	1.20	0.4581	0.40
Ψ_{Hire}	0.3087	1.03	0.2124	-	0.1048	0.60	0.6127	1.20
Ψ_{LDC}	1.0000	-	1.0000	-	0.7936	0.25	0.2734	2.15
Ψ_{MC}	0.1493	2.58	0.7165	1.14	0.2924	2.09	0.4573	1.38
Ψ_{PT}	1.0000	-	1.0000	-	1.0000	-	1.0000	-
Ψ_{Taxi}	0.4877	1.13	0.6670	1.32	0.6540	0.93	0.1109	2.82

† IVT = in-vehicle access-time

T-statistics calculated with respect to 0 for taste coefficients, and with respect to 1 for structural parameters.

Table 5: CNL estimation results, part 1

	Resident				Visitor			
	Business		Leisure		Business		Leisure	
	Est.	t-stat.	Est.	t-stat.	Est.	t-stat.	Est.	t-stat.
π_{A1}	0.1888	1.05	1.0000	-	0.4976	0.84	1.0000	-
π_{A2}	0.1726	1.88	1.0000	-	0.4936	2.39	0.1367	1.71
π_{A3}	1.0000	-	1.0000	-	0.1628	1.68	1.0000	-
π_{A4}	0.4727	2.01	1.0000	-	1.0000	-	0.7661	0.81
π_{A5}	0.1233	1.52	1.0000	-	0.4404	1.47	0.3054	-
π_{A6}	0.1016	0.14	1.0000	-	0.2101	1.10	0.1271	7.20
π_{A7}	1.0000	-	1.0000	-	1.0000	-	0.8871	0.25
π_{A8}	0.4773	0.40	0.6816	0.47	0.3354	0.33	0.5610	0.89
π_{A9}	1.0000	-	1.0000	-	1.0000	-	1.0000	-
π_{A10}	0.0752	†	1.0000	-	1.0000	-	1.0000	-
π_{A11}	0.5512	1.00	0.6829	0.96	0.1196	0.61	0.0613	0.30
π_{A12}	0.2936	2.17	1.0000	-	0.7029	0.30	0.2192	2.52
π_{A13}	0.9049	0.26	0.4500	1.41	0.8986	0.20	0.8286	0.36
π_{A14}	0.0998	2.96	1.0000	-	0.2796	4.43	0.2348	3.41
π_{A15}	0.5752	0.82	0.1554	2.06	0.1083	3.94	0.1079	1.24
π_{A16}	0.4087	1.77	1.0000	-	0.6396	1.70	0.1000	2.03
π_{A17}	0.2772	1.75	0.0584	-	0.4494	1.32	1.0000	-
π_{A18}	0.3386	0.88	1.0000	-	0.9075	0.17	1.0000	-
π_{A19}	0.1051	-	0.0502	1.47	0.6685	0.24	0.2115	1.71
π_{A20}	0.1012	5.53	1.0000	-	1.0000	-	1.0000	-
π_{A21}	0.4790	1.28	0.0949	1.18	0.1072	3.04	1.0000	-
π_{A22}	1.0000	-	1.0000	-	1.0000	-	0.6505	0.37
π_{A23}	1.0000	-	1.0000	-	1.0000	-	1.0000	-
π_{A24}	0.7482	0.70	1.0000	-	0.6214	0.50	1.0000	-
π_{A25}	1.0000	-	1.0000	-	0.8376	0.22	0.7344	0.52
π_{A26}	1.0000	-	1.0000	-	1.0000	-	1.0000	-
π_{A27}	0.2320	2.21	1.0000	-	1.0000	-	0.8216	0.37
π_{A28}	0.4722	1.78	1.0000	-	1.0000	-	1.0000	-
π_{A29}	0.8907	0.20	1.0000	-	0.4366	-	0.2525	0.99
π_{A30}	0.2394	1.29	0.0510	1.39	1.0000	-	0.9219	0.28
π_{A31}	0.1197	5.56	0.1835	0.75	0.2278	2.07	0.5838	1.58
π_{A32}	0.5304	2.09	1.0000	-	0.2502	1.25	0.0475	2.85
π_{A33}	0.8717	0.26	1.0000	-	1.0000	-	1.0000	-
π_{A34}	1.0000	-	1.0000	-	0.5096	0.81	0.5889	0.91
π_{A35}	0.0993	5.83	1.0000	-	1.0000	-	1.0000	-
π_{A36}	1.0000	-	1.0000	-	1.0000	-	1.0000	-
π_{A37}	0.1233	0.83	0.8262	0.73	0.6993	0.68	1.0000	-

† Estimate of standard error not reliable

T-statistics calculated with respect to 1.

Table 6: CNL estimation results, part 2

		Resident		Visitor	
		Business	Leisure	Business	Leisure
MNL	mean	17.45	24.03	38.66	24.07
	std.dev.	24.83	37.60	45.69	35.93
NL by airport	mean	19.18	22.77	40.52	20.93
	std.dev.	27.30	35.62	47.88	31.25
NL by airline	mean	19.08	24.02	41.70	21.93
	std.dev.	27.15	37.57	49.28	32.74
NL by access-mode	mean	14.52	25.18	35.07	20.97
	std.dev.	20.67	39.38	41.44	31.31
CNL	mean	16.76	25.37	38.44	19.16
	std.dev.	23.86	39.69	45.42	28.60

Table 7: Trade-off between in-vehicle access-time and access-cost (£/hour)

		Resident		Visitor	
		Business	Leisure	Business	Leisure
MNL	mean	1.07	1.14	1.09	2.96
	std.dev.	0.70	0.90	0.98	2.79
NL by airport	mean	0.97	0.94	0.98	2.80
	std.dev.	0.63	0.74	0.88	2.63
NL by airline	mean	1.13	1.34	1.12	3.73
	std.dev.	0.74	1.06	1.01	3.51
NL by access-mode	mean	1.06	1.03	0.81	2.81
	std.dev.	0.69	0.81	0.73	2.64
CNL	mean	0.89	0.82	0.78	3.17
	std.dev.	0.58	0.65	0.70	2.98

Table 8: Trade-off between flight-time and in-vehicle access-time

		Resident		Visitor	
		Business	Leisure	Business	Leisure
MNL	mean	6.76	19.67	5.83	16.90
	std.dev.	10.02	29.36	8.14	22.64
NL by airport	mean	6.45	21.32	5.76	19.75
	std.dev.	9.56	31.84	8.05	26.46
NL by airline	mean	7.10	20.02	5.64	17.84
	std.dev.	10.51	29.88	7.88	23.91
NL by access-mode	mean	5.79	18.30	5.17	15.53
	std.dev.	8.58	27.33	7.23	20.81
CNL	mean	3.96	19.78	6.93	17.66
	std.dev.	5.87	29.53	9.68	23.66

Table 9: Trade-off between frequency and in-vehicle access-time (min/flight)