**An indirect flight from my local airport or a direct flight from an alternative regional airport – how does surface access influence the decision?**

Johnson, Matthews, Hess and Bielefeldt

**Abstract**

This paper reports on a study which seeks to improve our understanding of how people choose between different kinds of flight at competing regional airports, and how their choices are affected by access conditions. In particular, it investigates whether improving surface access to airports that are in relatively close proximity to one another leads people to avoid taking indirect flights from their nearest airport in favour of direct flights from an alternative airport.

## Introduction

Deregulation of US and European airline markets has allowed the development of different forms of route and network structure in air passenger transport, leading to greater choice for passengers regarding the airport they use and the type of service they fly with. The clearest examples of these differences is the contrast between the full service carrier, operating a hub and spoke network through a major hub airport (such as BA at Heathrow or Lufthansa at Frankfurt), as compared with the low-cost carrier operating a set of point to point services through a number of secondary airports. Button (2002) highlights the wide use of hub and spoke system in network based industries and the advantages they confer in terms of fares and accessible destinations. More recently, variations have emerged on these contrasting cases, involving hub and spoke networks being operated by alliances of airlines rather than by one airline and, on the other hand, point to point services serving some of the more major airports; indeed, some of these point to point services, by linking in to major hub airports, start to provide for some of the onward connections features of the hub and spoke network (though without features such as through-ticketing and connection guarantees).

What emerges, particularly in Europe, are situations where passengers are increasingly faced with choices regarding which airport they fly from and what kind of service they fly with. However, whilst the availability of greater choice amongst flight-type has broadened over recent years, it is noted by Kouwenhoven (2008), that little is known about the influence of the type or level of airline service or airport quality-related factors. In this paper we report on a study which seeks to improve our understanding of how people choose between different kinds of flight at competing airports, and how these choices are affected by access conditions. Thus, we seek to generate new evidence to develop the understanding of the interaction of the type or level of airline service and the ease of surface access upon the choice of airport. In particular, we investigate people’s preferences between direct and indirect flights from two airports that are in relatively close proximity to one another, and whether improving surface access to them leads people to avoid taking indirect flights from their nearest airport in favour of direct flights from an alternative airport. In addition, we were also interested to understand more about how people choose between surface access options.

The focus of the study is on Edinburgh Airport and Glasgow Airport, the two busiest airports in Scotland and only 67km apart from one another. With over 9m passengers and serving over 100 destinations, Edinburgh Airport proclaims itself to be Scotland’s busiest airport, whilst Glasgow, with over 7m passengers and serving over 80 destinations, is by far the second busiest. A number of destinations are served direct by both airports, such as Heathrow, Paris CDG and Amsterdam, but there are also a number of destinations for which it is only possible to fly directly from one or other of the airports. For example, there are direct flights from Edinburgh to Brussels, Frankfurt, Vilnius and Zurich but not from Glasgow. At the same time, there are direct flights from Glasgow to Plymouth, Reykjavik, Dubai and Lahore, but not from Edinburgh.

Most recent figures show that, for 2009, 70% of Edinburgh Airport’s passengers access the Airport by car or taxi, and that this figure is some 85% for Glasgow Airport. Furthermore, CAA data shows that most people use their nearest airport, with 61% of Edinburgh Airport’s domestic passengers and 58% of its international passengers coming from the Lothian region, and some 90% of Glasgow Airport’s domestic passengers and 63% of its international passengers coming from the Strathclyde region.

The primary aim of the study was to investigate how travellers may respond in a situation where, for a given trip, their *home* airport offers only connecting flights to their chosen destination, while direct flights are available from the alternative airport. As an example, would a traveller living in or around Glasgow be willing to travel to Edinburgh airport if the latter offered direct flights to the chosen destination, while only connecting flights are offered from Glasgow. A secondary aim was to discuss the impact of a new direct rail access link in this context. Our sample focused on individuals who flew from either of the two airports to locations where interchange was appropriate, eg London travellers were excluded. Using a customised SP design, systematically varying the attributes of the journey across a series of ‘games’ we analyse how the different attributes are traded off against each other. We analyse the choice of airport and access mode jointly using a cross-nested logit model to allow for flexible substitution patterns between options. The results highlight the strong aversion to connecting flights. We observe very high willingness to accept higher air fares in return for direct flights and a strong willingness to accept increased access time in return for a direct flight.

The work was undertaken as part of the EU FP7 project INTERCONNECT. The project is concerned with how to improve interconnectivity in long distance travel, and the impacts of making such improvements.

In the rest of the paper we first review previous work to investigate airport choice and factors influencing the choice between different types of flight and access mode (section 2), before then reporting on the design and results of the SP survey conducted to analyse choices at Glasgow and Edinburgh Airports (section 3). We then close with our conclusions and recommendations.

## 2. Literature Review

**2.1 Airport Choice Models**

Work on Airport Choice dates back to Skinner’s (1976) use of multinomial logit models on air passenger survey data on the Baltimore Washington DC region, where he found higher sensitivity to ground accessibility than level of service.

The question of airport choice has been examined for many years through a number of studies, with accessibility and flight frequency consistently being highlighted as the key factors (Skinner, 1976; and Windle and Dresner, 1995). Other studies have identified aircraft type (eg Innes and Doucet, 1990) and ticket price (eg Bradley, 1998) as also being significant.

Many studies on airport choice have been based on data from the multi-airport San Francisco Bay Area. Harvey (1987) used this data to estimate separate multinomial logit models for business and non-business travel, as a function of highway access time and flight frequency, with both having non-linear effects on utility. He finds that beyond a certain threshold level additional direct services to a destination do not make airports more attractive. There is a large disutility for connecting flights. The disutility of access time decreases with total time, and shorter flights have more sensitivity than long-haul. Fare and access mode was not included.

Pels et al. (2001) use the data to analyse of the combined choice of airport and airline and find that that airline choice is correlated with the choice of airport, while Pels et al. (2003) joint analyse airport and access-mode choice, finding high sensitivity to access time, especially for business travellers.

Hess and Polak (2005) used mixed logit modelling applied to airport choice in the San Francisco Bay area. Along with the usual findings of significant influences of access time, fare and service frequency, they find significant observed heterogeneity in tastes, particularly for access time, between groups of travellers, and unobserved variations within groups of travellers. They find business travellers less sensitive to fare than leisure.

Hess and Polak (2006) developed a Cross-Nested Logit model that allows for the joint representation of correlation between airport, airline and access mode for the Greater London area (Gatwick, Luton, Stansted, London City, Heathrow). This model outperformed nested logit models. As with other studies, access time is found to be a key factor in choice of airport, along with flight time and frequency and access cost. However, due to the quality of their data and the low sensitivity of business passengers, they were not able to estimate a significant effect of air fare. Hess et al apply this approach to a stated choice dataset based on broader regional data from the East Coast of the US, to reflect that passengers travel to far outlying airports to access cheaper flights. Again they find improvements from use of the CNL model, and that the correlation between airline and access mode can only be adequately captured if additionally allowing for correlation along the airport dimension, ie nested logit models would not be appropriate.

Hess et al (20XX), use a CNL model to analyse broader, regional data from the East Coast of US, reflecting the trend that passengers aare travelling further to access cheap flights offered by low-cost airlines. Their analysis is based on SP data, which offers greater improvements over nested logit models than with RP data, as the analysis of substitution patterns is made easier by multiple observations for individual respondents.

**2.2 Access Mode Choice Models**

Gosling (2008) conducted a comprehensive review of nine airport ground access mode choice models, based on RP or SP. Whilst most models include travel time and travel cost, he concluded that there was still uncertainty over which other explanatory variables to include and the appropriate nesting structures of different modes.

* 1. **Airline Service Choice**

O’Connell and Williams (2005) highlight the growing intensity of direct competition between full service and no-frills airlines. The brand intensity of low fare airlines was such that most of those surveyed on a low cost carrier did not look at other carriers. They find that passengers on full fares will tolerate higher fares, and there are corporate deals with large, favouring the larger incumbents. Full fare passengers also prefer reliability, quality, connections, frequent flyer discounts and comfort, whilst low cost are based almost exclusively on fare. Low cost travellers are willing to connect through secondary airports.

Mason (2001) finds there that there is little distinction business travellers who use low cost and network carriers, and as such do not represent two market groups – price and value for money are prime considerations for both groups (although price more so for those on low cost)

Barrett (2004) looks at the difference in services operated between low-cost carriers and the more established airlines. De-regulated low cost airlines operate on a point to point basis so their passengers do not need to transfer at hub airports, being more willing to transfer to smaller airports outside of destination cities. Low fare airlines have brought service to underutilised secondary airports. They are clearly tough operators, and airports have to respond to the new market power.

Whitaker, Terzis, Soong and Yeh (2005) carried out a number of SP experiments to evaluate Airline passenger preferences. Passengers clearly prefer current flight schedules. Qualitative findings indicated that flights outside preferred schedules needed heavy discounting. They found , in terms of airline products and services, that many travellers were non-traders in terms of check-in queue time.

## Design of survey

The survey work made use of a sample of respondents who had recently undertaken a return journey by air (scheduled flights only) from either Glasgow or Edinburgh to one of a set of predefined destinations. This list included most of the main destinations served from either airport, but excluded a number of destinations for which connecting flights would not be seen as viable (primarily London). Respondents were asked to provide details on this specific air journey, along with details for their access journey to the airport. Respondents were also asked what their preferred schedule would have been in terms of flight departure time, which allowed us to investigate the sensitivity to schedule delay.

With the above aims in mind, a stated choice (SC) survey was designed in which respondents were faced with a number of hypothetical choice scenarios focussed around a recent trip. In particular, each respondent in the survey was presented with two sets of five such scenarios. In the first five scenarios, the respondent was given a choice between a flight at Glasgow airport and a flight at Edinburgh airport. The two flights were described on the basis of:

* the type of airline (*full service* or *no frills*);
* the flight departure time;
* whether it was a direct or connecting flight;
* the connection time (if applicable);
* the total air journey time; and
* the return air fare.

In addition, four different access options were given for each of the two airports, namely *drive & park*, *dropped off*, *taxi*, and *bus*. In the second set of five scenarios, a fifth mode, namely a *direct rail link*, was added, where this was in some cases a *high speed rail link*. The different access modes were described to respondents in terms of time and cost. For respondents who had indicated that either or both drive & park and dropped off were not available options, the choice set was adjusted accordingly.

The actual attribute levels presented varied across scenarios, with the specific variations used coming from an advanced experimental design. In particular, this design presents scenarios that require trading between attributes (i.e. gains in one attribute are traded off against losses in another attribute), where this increases the ability to retrieve meaningful estimates of the relative sensitivities to the various attributes with a finite sample size. As mentioned above, main interest was in the choice between a connecting flight at the home airport and a direct flight at the alternative airport. For obvious reasons, this situation did however not apply in each single scenario as this would have led to perfect collinearity in the data. Rather, the rate of connecting flights was simply set to a higher level at the home airport.

In each scenario, the respondent was asked to choose his or her preferred pairing of airport and access mode, while a *No travel* option was also included.

The following design specifications were used for the different attributes:

* Airline type: between two thirds and three quarters of flights are on full service airlines
* Flight departure time: pivoted around the departure time reported for the recent trip, with five levels, namely 2 hrs;-1 hr; no change;+1 hr;+2 hrs. No flights were presented with departures before 6AM or after 11PM.
* Direct or connecting flight: around 50% of flights at the home airport were direct flights, while this increased to over 2/3 at the alternative airport.
* Connection time: five different levels, namely 45 minutes, 60 minutes, 90 minutes, 120 minutes, and 180 minutes.
* Air journey time: composed of direct flight time, connection time, and an additional indirect routing effect, pivoted around values from look-up tables
* Return air fare: pivoted around the reported air fare, with five levels, namely -50%; -20%, 0%, +20%, +50%.
* Access time: pivoted around values from look-up tables, with five levels, namely -30%; -15%, 0%, +15%, +30%
* Access cost: pivoted around values from look-up tables, with five levels, namely -30%; -15%, 0%, +15%, +30%

## Data collection

The survey was implemented as an on-line self-completion questionnaire. Survey respondents were recruited in three different ways. Firstly, a survey company used a database of individuals, or panel, and contacted those in scope for the survey. A second batch of respondents were recruited at Edinburgh Airport. In the Edinburgh Airport sample, travellers at the airport were handed cards with a web-link for the survey to be completed within a three week time window. Finally, a small number of individuals were recruited through personal contacts.

In order to be in scope for the survey, individuals had to have undertaken scheduled return flight in the last six months to one of a list of popular destinations. This list of destinations was censored to exclude any destinations where only direct flights were judged feasible, eg London was excluded on these grounds.

To support the customisation of the survey to the individual’s access origin, time and cost look-up tables were constructed detailing times and costs of representative taxi, bus and car options from 163 postcode districts (eg EH6, for Edinburgh areas of Leith and Newhaven) close to the airport. Areas further away were broken down into 11 larger postal areas (eg IV for Inverness).

Table 1 shows the responses from each survey recruitment approach.

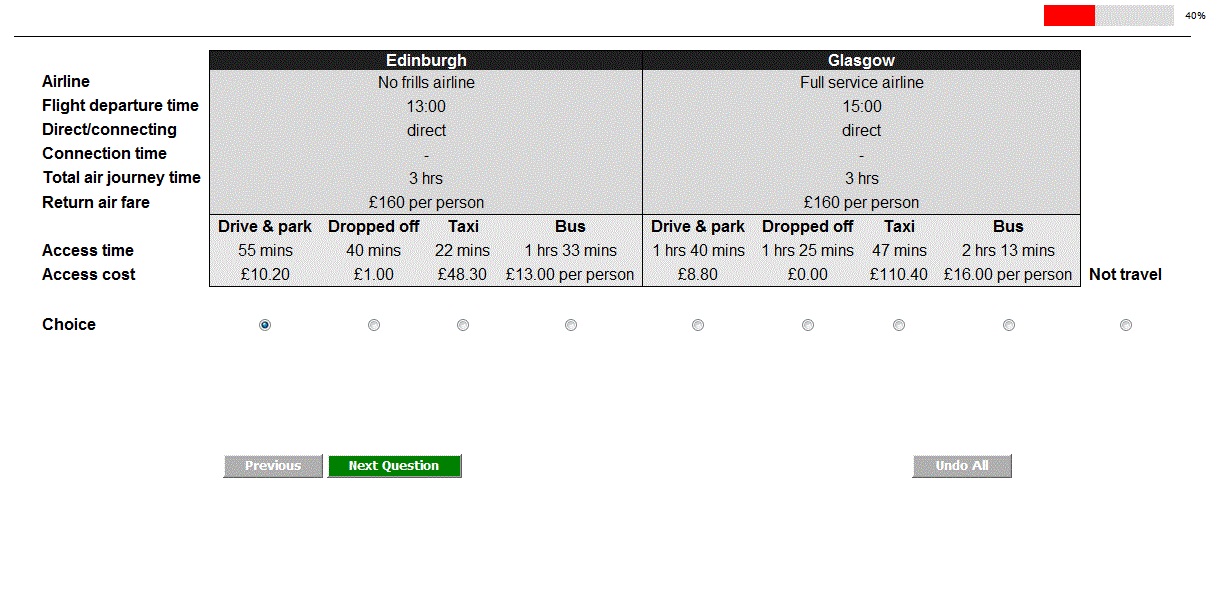
**Table 1: Completed Surveys**

|  |  |
| --- | --- |
| **Respondent type** | **Completed responses** |
| On-line panel | 250 |
| Edinburgh Airport | 80 |
| Personal contacts | 12 |
| **Total** | **342** |

After data cleaning, a final sample of 303 respondents was used for the analysis, leading to 3,030 separate choice scenarios.

Examples of choice scenarios are shown below, the first one taken from the first five choices and the second from the second five choices, including the rail access option. These are based on a respondent whose reference trip was a flight from Edinburgh to Rome, costing £200.

**Figure 1: Example choice scenario (excluding rail access option)**



**Figure 2: Example choice scenario (including rail access)**

# C:\Users\the_triffid\Pictures\ic2.gifModel specification and estimation

### Introduction

The SC data collected in this project was analysed with the help of discrete choice models belonging to the family of random utility models. These models are estimated on data containing information on real world choices or data from hypothetical choice scenarios (stated choice). They explain individual choices on the basis of the concept of utility maximisation; decision makers evaluate the various alternatives that are available to them and choose the one that provides them with the greatest utility (or smallest disutility). The utility of an alternative is a function of attributes of the alternative and characteristics of the decision maker. This includes both observable variables, such as socio-demographic characteristics, and unobservable components that need to be estimated by the analyst. The main emphasis is on the sensitivities of the respondent to changes in the attributes, commonly referred to as tastes or marginal utilities. Data coming from SC surveys are by nature very well suited for the computation of relative sensitivities (i.e. willingness to pay measures), but are not suitable for producing forecasts of likely demand.

### Utility specification

The final specification used in our models estimated the following parameters.

* b\_ac: marginal utility of changes in access cost, with the attribute being measured in £
* b\_at: marginal utility of changes in access time, with the attribute being measured in minutes
* b\_ctime\_120: parameter estimated for connecting flights, with a connection time of 120 minutes
* b\_ctime\_180: parameter estimated for connecting flights, with a connection time of 180 minutes
* b\_ctime\_45: parameter estimated for connecting flights, with a connection time of 45 minutes
* b\_ctime\_60: parameter estimated for connecting flights, with a connection time of 60 minutes
* b\_ctime\_90: parameter estimated for connecting flights, with a connection time of 90 minutes
* b\_fare: marginal utility of changes in air fare, with the attribute being measured in £
* b\_sde: marginal utility of changes in early schedule delay (i.e. arrival earlier than preferred arrival time), with the attribute being measured in minutes
* b\_sdl: marginal utility of changes in late schedule delay (i.e. arrival later than preferred arrival time), with the attribute being measured in minutes
* d\_coach: constant for coach
* d\_glas: constant for Glasgow
* d\_home: constant for *home* airport
* d\_hsr: constant for high speed rail
* d\_kf: constant for dropped off
* d\_nofrills: constant for no frills airlines
* d\_notr: constant for no travel alternative
* d\_rail: constant for conventional rail
* d\_taxi: constant for taxi

It is important to recognise that the degree of randomness in the choices (from an analyst’s perspective) may differ between the two sets of five choices (given the additional rail alternative in game 2). With this in mind, we also estimated an additional scale parameter (scale2) for the second game. This scale parameter (which is normalised to 1 for game 1) is inversely proportional to the variance of the error term. This means that an estimated value greater than 1 for scale2 means less response error and more deterministic choices (from the analyst’s perspective) in the second set of five choices.

The sample used for estimation was too small for segmentation by purpose, but we allowed for continuous interactions with journey time and air fare (as in Bierlaire et al XXX). In particular (where used), for attribute *x*, we estimated β∙(zn/z)λ∙x. Here, we interact the sensitivity to attribute x with the socio-demographic variable z (e.g. income), which has a mean value of z in the sample population, and a value of zn for respondent n. The additional parameter λ now gives the elasticity of the sensitivity β to changes in z.

We estimated the following interaction terms using this specification:

* lambda\_ac\_inc: income elasticity for access cost sensitivity
* lambda\_at\_ft: total journey time elasticity for access time sensitivity
* lambda\_conn\_ft: total journey time elasticity for sensitivity to connections
* lambda\_fare\_inc: income elasticity for fare sensitivity
* lambda\_sde\_ft: total journey time elasticity for sensitivity to early schedule delay
* lambda\_sdl\_ft: total journey time elasticity for sensitivity to late schedule delay

### Model structure

For each respondent in each choice task, there are two possible airports, each with up to five access modes, depending on personal availabilities, as well as whether the choice scenario is for game 1 or game 2. There is also a no-travel alternative. This thus leads to a choice between up to 9 different alternatives in game 1, and up to 11 different alternatives in game 2.

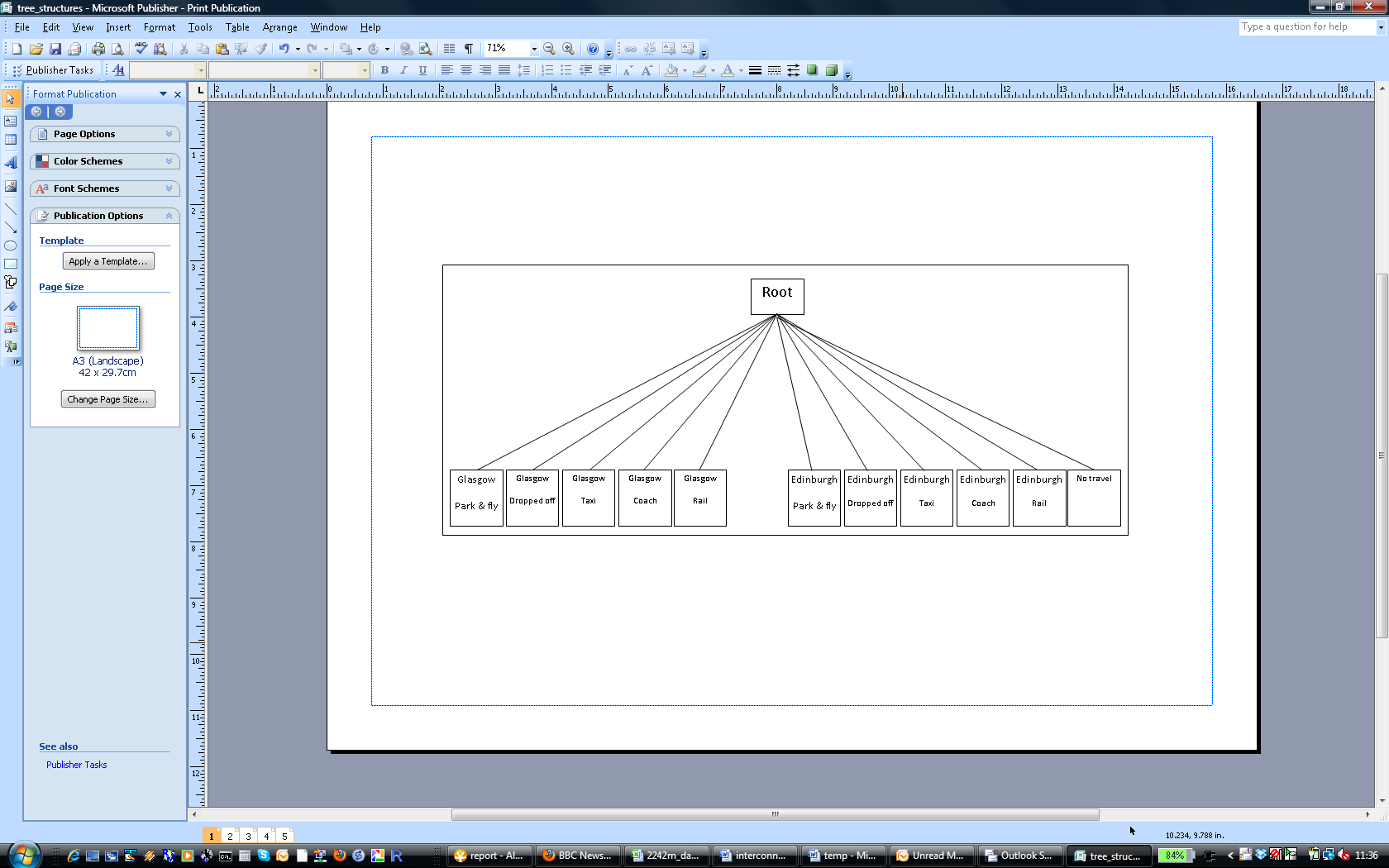
The model structure chosen for the present analysis is a Cross-Nested Logit (CNL) model. The justification for this is now explained in stages.

The main point is that the choice process in the survey is two dimensional, with a respondent being asked to make a choice both along the airport dimension and along the access mode dimension. Here, we would expect heightened substitution between two alternatives sharing a given airport, but also between two options sharing an access mode. In other words, if a respondent’s current airport-access mode combination becomes unavailable, he/she will be more likely to switch to another option that keeps one of the dimensions of choice constant, rather than changing both airport and access mode.

We thus aim to allow for correlation along both dimensions of choice. In other words, we want to allow for heightened correlation between two options sharing the same airport, as well as heightened correlation between two options sharing the same access mode, but no correlation between options at different airports using different access mode.

For this purpose, each alternative is in these models specified as a pair of alternatives, made up of one airport and one access mode, with the exception of the *No Travel* option. When estimating simple Multinomial Logit (MNL) models (cf. McFadden, 1974), we do not allow for any correlation between the random part of the utility[[1]](#footnote-1). The simple MNL model would thus fail to account for the correlation between alternatives at the same airport, or alternatives using the same access mode at the two different airports. The resulting structure is illustrated in Figure 3.

**Figure 3: MNL structure**

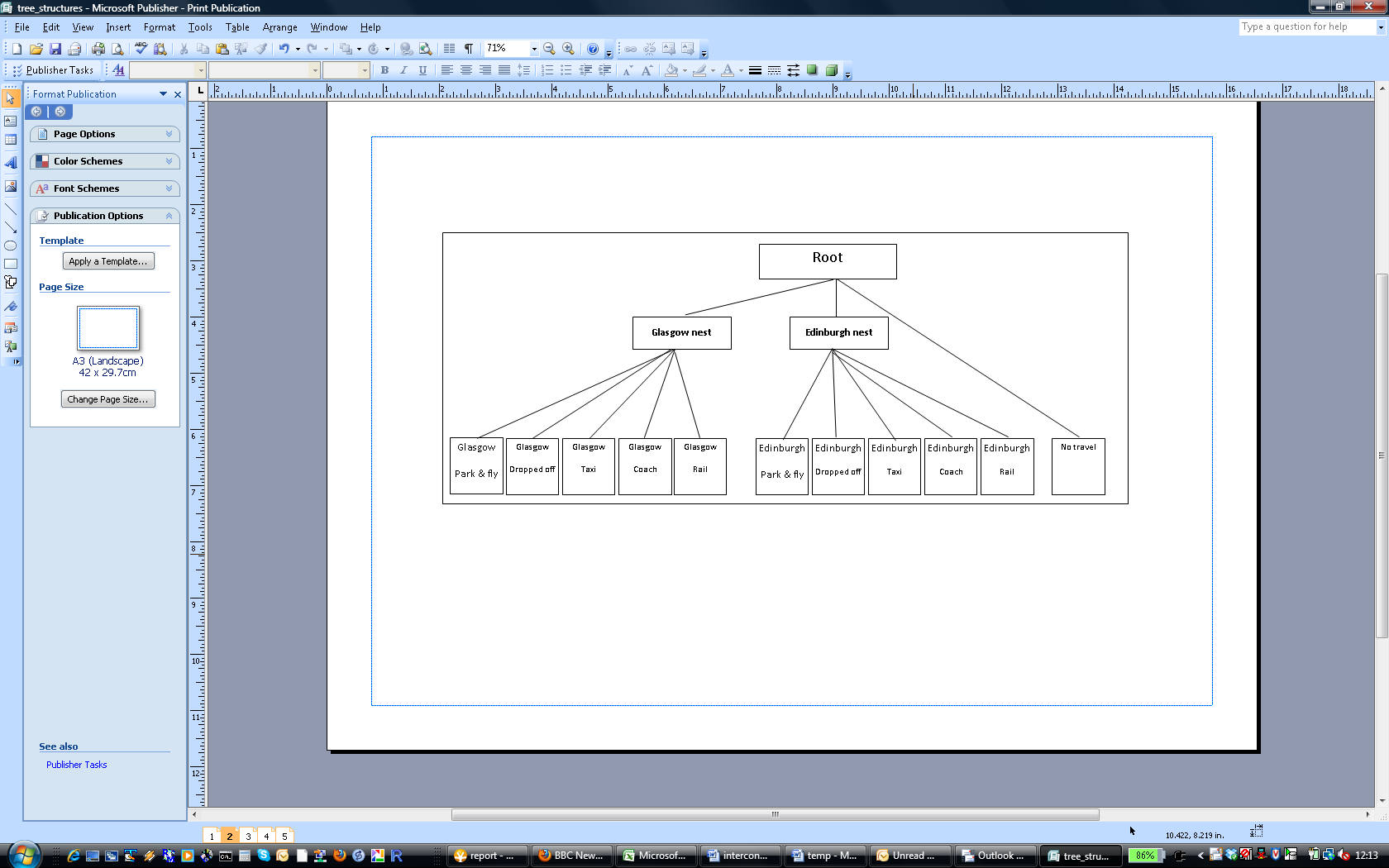
****

As a result of its independence assumption, this model thus cannot represent such substitution patterns as are likely to arise in this dataset, and there will be a proportional shift in probability towards all alternatives if one alternative becomes unavailable or reduces in attractiveness (e.g. due to increased cost). Crucially, this has implications not just in terms of the MNL model, producing incorrect substitution patterns and hence forecasts (less of an issue with SC data where such forecasts are inappropriate in any case), but also on the estimated parameters. I.e. the willingness-to-pay measures coming out of a MNL model may be biased due to a failure to account for inter-alternative correlation.

The typical approach for dealing with such an issue is estimating a Nested Logit (NL) model (cf. Daly and Zachary, 1978; McFadden, 1978; Williams, 1977)[[2]](#footnote-2). In this model, the error terms still follow an extreme value distribution, as in the simple MNL model, but the error terms of individual alternatives are no longer independently distributed. Any correlation between the error terms (or unobserved utility components) will lead to heightened substitution patterns between these two alternatives. Each alternative belongs to exactly one nest in a NL model, where a nest groups together alternatives that are closer substitutes for one another, for example public transport access modes, and where single alternative nests are used for any alternatives whose error terms are uncorrelated with those of any other alternatives. For each nest containing at least two alternatives, we estimate an additional model parameter λ, where this parameter is constrained between 0 and 1, with 1 reflecting an absence of correlation, and where the actual level of correlation between the errors is given by 1- λ2, so that decreasing values of λ lead to increased correlation[[3]](#footnote-3).

The first example of such a structure in the present study is shown in Figure 4.

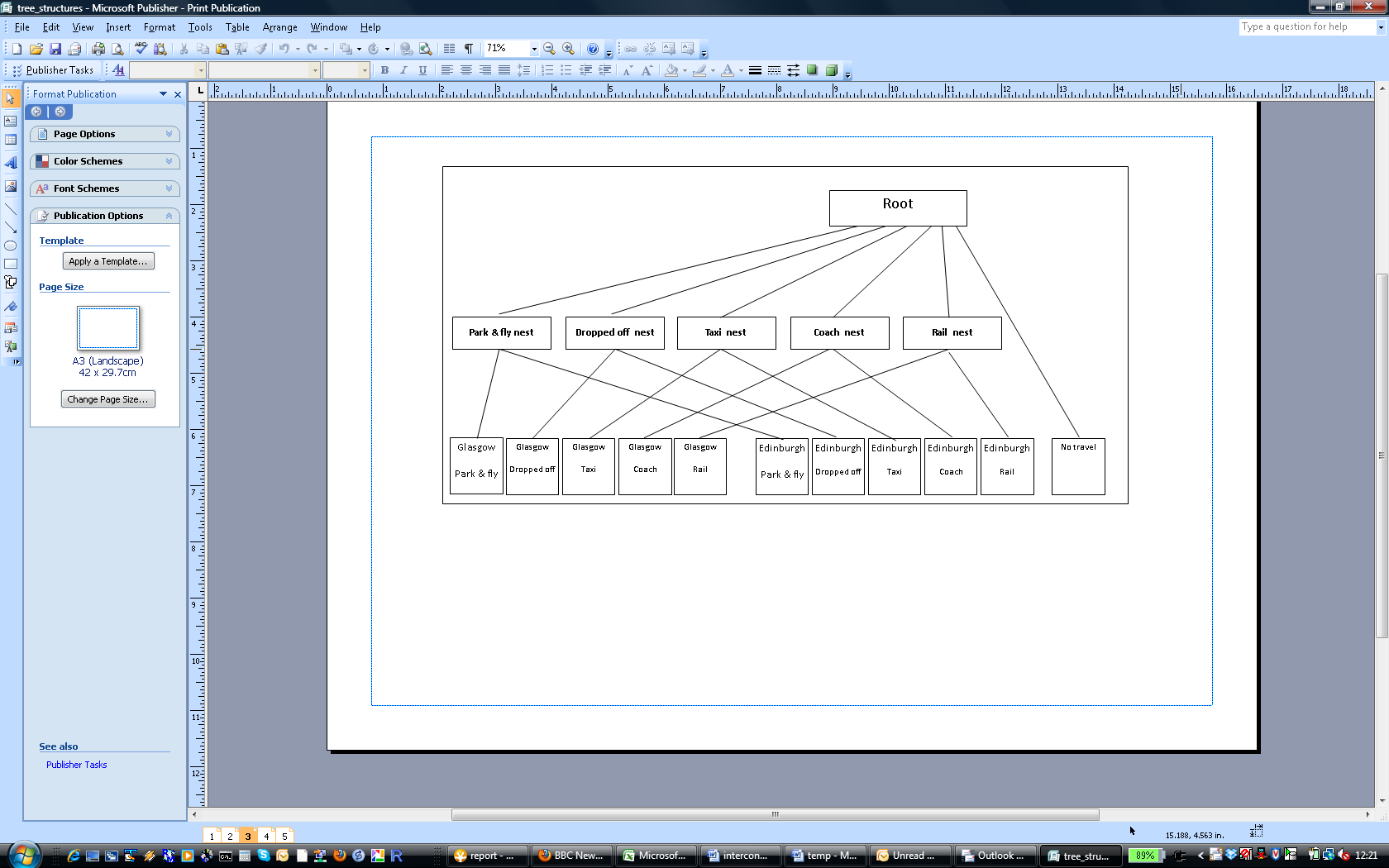
**Figure 4: Nested logit structure (nesting by airport)**



This groups together those alternatives representing flights at the same airport. This model has the capacity to allow, where appropriate, for heightened correlation between the error terms for alternatives sharing the same airport. Let us assume that in the present scenario, the model identifies such heightened correlation in the various airport nests. Now let us further assume that the coach option at Glasgow becomes unavailable for some reason. Given the heightened correlation within the nests grouping together alternatives sharing the same airport, there would in this case be a greater shift in probability towards the other options at Glasgow airport than towards the options at Edinburgh airport. This is clearly consistent with intuition; if a preferred flight at the current airport becomes unavailable, the respondent may be more likely to shift to a different flight at this airport than to shift to a flight at a different airport. Such substitution patterns would not be possible in a MNL structure; indeed, with any alternative becoming unavailable, there would be a proportional shift to all remaining alternatives, independently of which airport they are using.

The above discussion highlights the fact that the MNL may not be appropriate and suggests that gains in realism can be made by making use of a NL structure nesting together alternatives sharing the same airport. At the same time however, we should acknowledge that a corresponding structure can be produced for a model nesting access mode, i.e. grouping together those alternatives sharing the same access mode, a situation illustrated in Figure 5.

**Figure 5: Nested logit structure (nesting by access mode)**

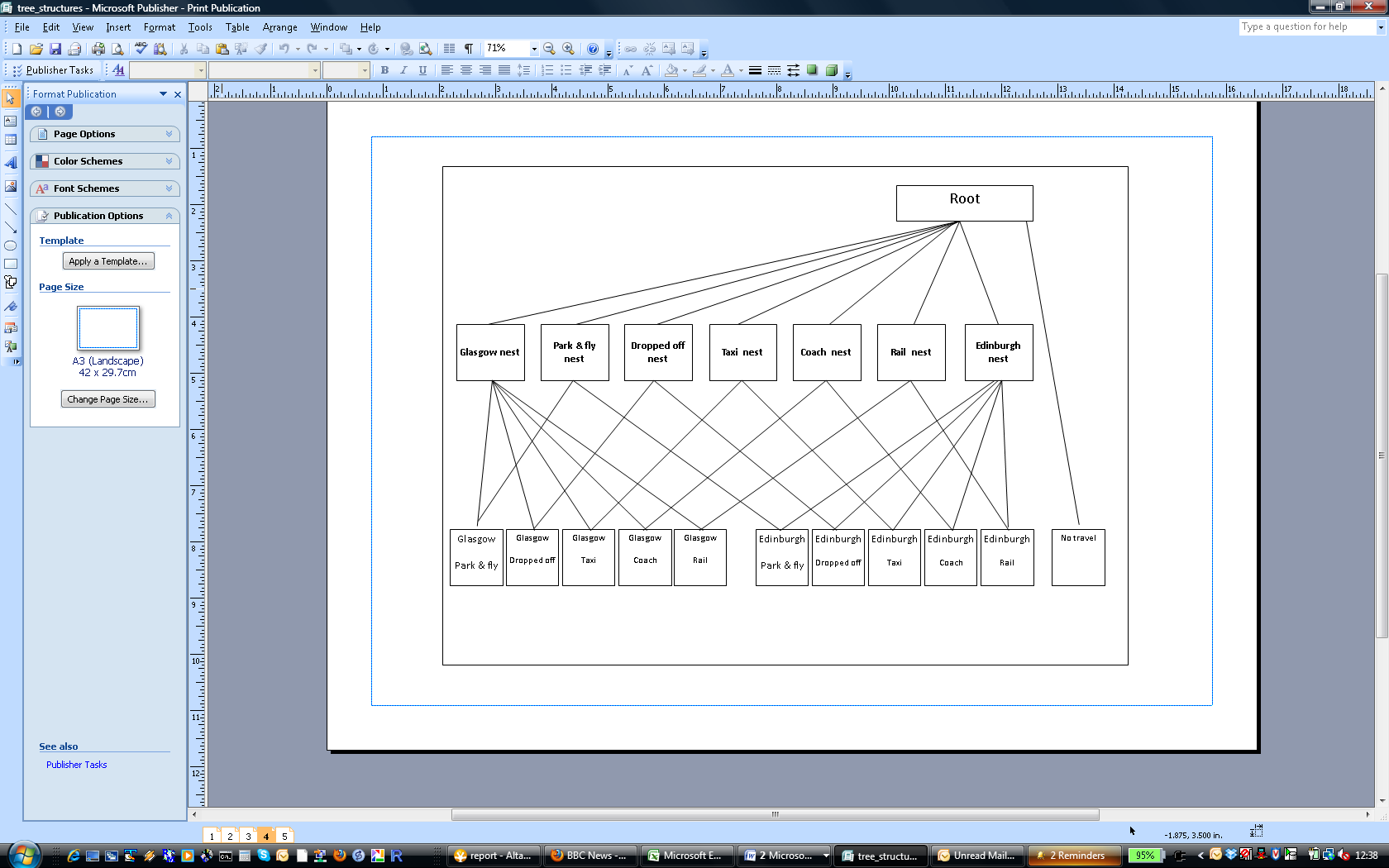


Crucially, either of these two models only allows us to represent correlation along a single dimension of choice, which may again lead to counterintuitive forecasts of behaviour in case correlation actually arises along multiple dimensions. In other words, the model in Figure 4 does not allow for correlation between two options with the same access mode but using different airports, while the model in Figure 5 disregards any correlation between two options at the same airport but using different access modes.

The solution to this problem is to follow the suggestion of Hess & Polak (2006) and specifically a CNL structure, as illustrated in Figure 6. In a CNL model, we avoid the restriction of making the nests mutually exclusive, meaning that an alternative can belong to multiple nests, leading to more flexible substitution patterns. As an example, imagine the situation where we want to have correlation between alternatives A and B, and between alternatives B and C, without correlation between alternatives A and C. Such a scenario cannot be accommodated in a NL structure, as we would have to group alternative B both with A and with C, thus also introducing correlation between A and C. In the CNL structure, we would have two separate nests, grouping together A and B, and B and C respectively. In a CNL model, an alternative is allowed to belong to more than one nest, thus allowing for far greater flexibility in the specification of the correlation structure. As an example, we can allow for situations in which we have correlation between alternatives A and B, and between alternatives A and C, with no correlation between alternatives B and C. The CNL model has its origins in the work of Vovsha (1997). Various alternative versions of the CNL model have been proposed by Vovsha & Bekhor (1998), Ben-Akiva & Bierlaire (1999) (further expanded by Bierlaire 2006), Papola (2004), and Wen & Koppelman (2001). The differences between the models arise primarily in the specification of the allocation parameters and the conditions associated with these parameters. The role of the allocation parameters is to explain the membership of an alternative in the different nests of the model, where these parameters are required given that we are no longer operating under the strict single nest membership condition of the simple NL model[[4]](#footnote-4).

In the present context, the CNL structure makes use of one nest for each airport and one nest for each access mode, i.e. 7 nests in total. Each alternative in this model is still made up of an airport and access mode pair, but now belongs to two nests, one airport nest and one access mode nest[[5]](#footnote-5). This creates correlation between two alternatives sharing the same airport, and between two alternatives sharing the same access mode. The model allows simultaneously for the correlation along each of the two dimensions of choice.

**Figure 6: Cross nested logit structure**



# Model results

## Main estimation results

All model estimation and forecasting work reported in this paper was carried out using BIOGEME (Bierlaire, 2005), where the standard errors were corrected to account for the repeated choice nature of the data used in the analysis by using the panel specification of the sandwich estimator (cf. Daly & Hess, 2010).

During the estimation work, we estimated MNL, NL and CNL models. In line with the above discussions, our expectation was that superior performance would be obtained by the CNL model, and this was the case. The model also yielded more reasonable willingness-to-pay (WTP) patterns. As a result, we will now focus solely on the outputs from this model.

The final estimation results for the CNL model are summarised in Table 1.

**Table 1: CNL estimation results**

|  |  |  |
| --- | --- | --- |
| **Respondents** | 303 | |
| **Observations** | 3030 | |
| **Model parameters** | 31 | |
| **Null Log-likelihood** | -6961.607 | |
| **Final log-likelihood** | -5662.11 | |
| **Ajdusted ρ2** | 0.182 | |
|  |  |  |
| **Utility parameters** | **est.** | **t-rat. (0)** |
| b\_ac | -0.0232 | -5.85 |
| b\_at | -0.00691 | -3.65 |
| b\_ctime\_120 | -0.415 | -6.87 |
| b\_ctime\_180 | -0.481 | -5.13 |
| b\_ctime\_45 | -0.438 | -3.32 |
| b\_ctime\_60 | -0.324 | -4.09 |
| b\_ctime\_90 | -0.398 | -4.71 |
| b\_fare | -0.00154 | -7.83 |
| b\_sde | -0.00043 | -0.76 |
| b\_sdl | -0.00052 | -1.52 |
| d\_coach | -1.3 | -5.68 |
| d\_glas | 0.114 | 1.98 |
| d\_home | 0.221 | 3.55 |
| d\_hsr | 0.0437 | 0.37 |
| d\_kf | -0.436 | -3.09 |
| d\_nofrills | 0.13 | 2.89 |
| d\_notr | -2.97 | -9.05 |
| d\_rail | -0.171 | -1.55 |
| d\_taxi | -0.718 | -4.3 |
| lambda\_ac\_inc | -0.0345 | -0.15 |
| lambda\_at\_ft | -0.517 | -2.07 |
| lambda\_conn\_ft | -0.511 | -3.43 |
| lambda\_fare\_inc | -0.427 | -3.37 |
| lambda\_sde\_ft | -0.293 | -0.27 |
| lambda\_sdl\_ft | -0.6 | -0.99 |
| scale2 | 1.26 | 15.19 |
|  |  |  |
| **Airport nesting parameters** | **est.** | **t-rat. (1)** |
| Edinburgh | 1 | - |
| Glasgow | 0.6993 | 1.83 |
|  |  |  |
| **Access mode nesting parameters** | **est.** | **t-rat. (1)** |
| Park & fly | 0.3534 | 1.2 |
| Dropped off | 0.0461 | 3.41 |
| Taxi | 0.4762 | 1.13 |
| Coach | 0.8696 | 0.16 |
| Rail | 1 | - |

We observe the expected negative and significant sensitivities to increases in access cost, access time, and air fare. We also observe a negative response to connecting flights. Here, longer connections are penalised more than connections of one hour or ninety minutes, but there is also a higher dislike of very short (i.e. 45 minutes) connections. The sensitivity to early and late schedule delay is not significant at the usual levels of confidence. In terms of constants, we observe a slight preference for Glasgow over Edinburgh, with a higher preference for the *home* airport. For access modes, we see that all else being equal, *drive and park* (the base) has the highest utility, along with high speed rail. The lowest utility is obtained for *dropped off* and *taxi* (where these effects are net of cost and time). Surprisingly, the models show a very slight preference for *no frills* airlines. In terms of interaction effects, we see decreasing sensitivity to access time and connections on longer flights, along with decreasing fare sensitivity with higher income. The other interactions are not statistically significant, although they are all of the expected sign. Finally, we note that the scale parameter for the second game is greater than 1, indicating a somewhat reduced error variance for responses in game 2.

In terms of nesting parameters, an estimate less than 1 shows heightened correlation. Here, we observe heightened correlation between flights at Glasgow airport, but not between flights at Edinburgh airport; this could suggest higher allegiance by passengers to Glasgow than to Edinburgh. For access modes, no correlation is observed between rail alternatives, but this is maybe not surprising in the context of a *new* mode. On the other hand, we see high correlation (and hence mode allegiance) for all other access modes, although this is lower for coach (higher λ meaning lower correlation).

## Willingness-to-pay patterns

WTP values vary as a function of income and journey time. WTP indicators for a range of 9 representative individuals with a ‘short’, ‘medium’ and long flight times of 60, 180 and 420 minutes respectively and ‘low’, ‘medium’ and ‘high’ annual incomes of £20,000, £40,000 and £80,000 respectively. are shown in Table 2, An illustration of the impact of flight duration and income on WTP for access time reductions is also given in Figure 7.

The WTP for access time reductions is given by the ratio between the access time and access cost coefficients. Here, we observe decreasing WTP measures with increasing flight duration, while the impact of income is negligible (and not statistically significant). The actual values are higher than value of time measures for example in a commuting context, but in line with values from other airline studies – such high values are to be expected given the greater financial penalty associated with missing a flight.

When looking at the WTP measures for avoiding early (SDE) or late schedule delay (SDL) as shown in Table 2, we need to take into account the high associated standard errors. Nevertheless, we observe a higher WTP for avoiding late schedule delay than for avoiding early schedule delay, with a strong associated income effect, and a less strong impact of flight time.

Finally, the WTP measures for avoiding connecting flights are very high, especially for very short and long connections. We see a reduction in the WTP on longer flights, which is to be expected, while the WTP increases substantially with income. All in all, this shows that when faced with a choice between a direct and a connecting flight, travellers need substantial incentives to accept the indirect option.

**Table 2: WTP indicators for range of representative individuals.**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Flight time (direct, minutes) | **Short**  **60** | **Medium**  **180** | **Long**  **420** | **Short**  **60** | **Medium**  **180** | **Long**  **420** | **Short**  **60** | **Medium**  **180** | **Long**  **420** |
| Annual Income | **Low**  **20** | **Low**  **20** | **Low**  **20** | **Medium**  **40** | **Medium**  **40** | **Medium**  **40** | **High**  **80** | **High**  **80** | **High**  **80** |
| WTP for reducing access time (£/hr) | 35.5 | 20.12 | 12.98 | 36.36 | 20.61 | 13.3 | 37.24 | 21.11 | 13.62 |
| WTP for reducing SDE (£/hr) | 17.42 | 12.63 | 9.85 | 23.43 | 16.98 | 13.25 | 31.5 | 22.83 | 17.81 |
| WTP for reducing SDL (£/hr) | 31.9 | 16.5 | 9.93 | 42.89 | 22.19 | 13.35 | 57.67 | 29.83 | 17.94 |
| WTP for avoiding 45 minute connections | 397.43 | 226.7 | 147.03 | 534.31 | 304.78 | 197.67 | 718.35 | 409.76 | 265.76 |
| WTP for avoiding 60 minute connections | 293.99 | 167.69 | 108.76 | 395.25 | 225.45 | 146.22 | 531.38 | 303.11 | 196.59 |
| WTP for avoiding 90 minutes connections | 361.13 | 205.99 | 133.6 | 485.52 | 276.95 | 179.62 | 652.75 | 372.34 | 241.49 |
| WTP for avoiding 120 minutes connections | 376.56 | 214.79 | 139.31 | 506.26 | 288.78 | 187.29 | 680.63 | 388.24 | 251.8 |
| WTP for avoiding 180 minute connections | 436.44 | 248.95 | 161.47 | 586.77 | 334.7 | 217.08 | 788.87 | 449.98 | 291.85 |

**Figure 7: WTP for reducing access time as a function of income and flight time (£/hr)**

It is also possible to compute WTP measures for each individual in the sample, hence taking into account the sample distribution in terms of income and flight duration. The results from this calculation are shown in Table 3, once again showing very high WTP measures.

**Table 3: Sample level WTP indicators**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | WTP for access time (GBP/hr) | WTP for SDE (GBP/hr) | WTP for SDL (GBP/hr) | WTP for avoiding conn with 45 (GBP) | WTP for avoiding conn with 60 (GBP) | WTP for avoiding conn with 90 (GBP) | WTP for avoiding conn with 120 (GBP) | WTP for avoiding conn with 180 (GBP) |
| **min** | 31.25 | 9.21 | 16.09 | 203.22 | 150.33 | 184.66 | 192.55 | 223.17 |
| **mean** | 33.62 | 23.42 | 40.90 | 516.46 | 382.04 | 469.30 | 489.34 | 567.17 |
| **sd** | 0.70 | 5.40 | 9.44 | 119.17 | 88.15 | 108.28 | 112.91 | 130.86 |
| **max** | 34.53 | 31.66 | 55.28 | 698.18 | 516.46 | 634.42 | 661.52 | 766.72 |

## Trade-offs against access time

The main finding from the analysis so far is that travellers have a very strong aversion to connecting flights, especially in the case of short haul routes. The model estimates show a very high willingness by respondents to pay higher fares in return for direct flights. Similarly, it is possible to compute trade-offs between access times and connecting flights. Looking at the sample level distribution of these indicators, we obtain the results in Table 4. The resulting trade-offs show the amount of extra access time respondents are willing to incur in return for avoiding a connecting flight.

**Table 4: Willingess to accept longer access time in return for avoiding connections (sample level)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Willingess to accept longer access time in return for avoiding conn with 45 (mins)** | **Willingess to accept longer access time in return for avoiding conn with 60 (mins)** | **Willingess to accept longer access time in return for avoiding conn with 90 (mins)** | **Willingess to accept longer access time in return for avoiding conn with 120 (mins)** | **Willingess to accept longer access time in return for avoiding conn with 180 (mins)** |
| **min** | 51.49 | 38.09 | 46.79 | 48.79 | 56.54 |
| **mean** | 105.01 | 77.68 | 95.42 | 99.50 | 115.32 |
| **sd** | 48.70 | 36.03 | 44.26 | 46.15 | 53.48 |
| **max** | 215.16 | 159.16 | 195.51 | 203.86 | 236.29 |

# Conclusions

The aim of this study was to conduct an analysis of air travel behaviour using stated choice data. Specifically, the main focus was on studying travellers’ preferences in relation to connecting flights, and how they might react to a situation in which only connecting flights were offered from their *home* airport, while direct flights were offered from an alternative airport. The results highlight the strong aversion to connecting flights. In particular, we observe very high willingness to accept higher air fares in return for direct flights. Similarly, we observe a strong willingness to accept increased access time in return for a direct flight. Here, the values observed in Table 4 show that the aversion to connecting flights is so high that in a scenario where direct flights were only available at the *non-home* airport, the access time to this alternate airport would be sufficiently low to guarantee a high level of interconnecting passengers, independently of the existence of a direct rail link, and even when taking into account the higher baseline preference for the *home* airport.

# References

Barrett, S.D., (2004), How do the demands for airport services differ between full-service carriers and low-cost carriers?, Journal of Air Transport Management 10 (2004) 33–39

Ben-Akiva, M. & Bierlaire, M. (1999), Discrete choice methods and their applications to short term travel decisions, in R. Hall, ed., ‘Handbook of Transportation Science’, Kluwer Academic Publishers, Dordrecht, The Netherlands, chapter 2, pp. 5–34.

Bierlaire, M., 2005. An introduction to BIOGEME Version 1.4. biogeme.ep.ch.

Bierlaire, M. (2006). A theoretical analysis of the cross-nested logit model. Annals of Operations Research, Vol. 144, 287-300.

Button, K. J. (2002), ‘Debunking Some Common Myths about Airport Hubs’, Journal of Air Transport Management, May:

Daly, A.J. & Hess, S. (2010), Simple Approaches for Random Utility Modelling with Panel Data, paper presented at the European Transport Conference, Glasgow, October 2010.

Daly, A. and S. Zachary (1978), Improved multiple choice models, in D.Hensher and M.Dalvi, eds, ‘Determinants of Travel Choice’, Saxon House, Sussex.

Gosling, G.D. (2008) Airport Ground Access Mode Choice Models, a synthesis of airport practice, Airport Cooperative Research Program synthesis 5, Transport Research Board, available via <http://onlinepubs.trb.org/onlinepubs/acrp/acrp_syn_005.pdf>

Graham, B. and C. Guyer. (2000) “The Role of Regional Airports and Air Services in the United Kingdom”, *Journal of Transport Geography*, 8, 249-262.

Hess, S., Ryley, T., Davison, L.J., and Thomas Adler, (20XX), Improving the quality of demand forecasts through cross nested logit: a stated choice case study of airport, airline and access mode choice, ??

Hess, S., Polak, J.W., 2005a. Mixed Logit modelling of airport choice in multi-airport regions. Journal of Air Transport Management 11

(2), 59–68.

Hess, S. & Polak, J.W. (2006), Exploring the potential for cross-nesting structures in airport-choice analysis: a case-study of the Greater London area, *Transportation Research Part E*, 42, pp. 63-81.

Mason, K. (2001) “Marketing Low-Cost Airline Services to Business Travellers”, *Journal of Air Transport Management*, 7, 103-109.

McFadden, D. (1974), Conditional logit analysis of qualitative choice behavior, in P.Zarembka, ed., ‘Frontiers in Econometrics’, Academic Press, New York, pp. 105–42.

McFadden, D. (1978), Modeling the choice of residential location, in A.Karlqvist, L.Lundqvist, F.Snickars and J.Weibull, eds, ‘Spatial Interaction Theory and Planning Models’, North-Holland, Amsterdam, pp. 75–96.

O’Connell, J.F., and Williams, G., (2005), Passengers’ perceptions of low cost airlines and full service carriers: A case study involving Ryanair, Aer Lingus, Air Asia and Malaysia Airlines, Journal of Air Transport Management 11.

Papola, A. (2004). Some developments on the Cross Nested Logit model.

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, pp.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 A 37, 71–83.

Skinner, R. E. Jr. (1976) Airport choice: an empirical study *Transportation Engineering Journal,* 10, pp. 871-883

Transportation Research Part B, 38, 833-851.

Train, K.E. (2003), Discrete choice methods with simulation, Cambridge University Press, Cambridge, MA.

Wen, C.-H. & Koppelman, F. S. (2001), ‘The Generalized Nested Logit Model’, Transportation Research Part B: Methodological 35(7), 627–641.

Whitaker, B., Terzis, G., Soong, E., Yeh, W., (2005), Stated Preference as a Tool to Evaluate Airline Passenger Preferences and Priorities, Transportation Research Record: Journal of the Transportation Research Board, 1915, pp 55-61.

Williams, H. (1977), ‘On the formation of travel demand models and economic evaluation measures of user benefits’, Environment and Planning A 9, 285–344.

Windle, R. and Dresner, M. (1995) Airport choice in multi-airport regions *Journal of*

*Transportation Engineering,* 121, pp. 332-337

Vovsha, P. (1997), ‘Application of a Cross-Nested Logit model to mode choice in Tel Aviv, Israel, Metropolitan Area’, Transportation Research Record 1607, 6–15.

Vovsha, P. & Bekhor, S. (1998), ‘The Link-Nested Logit model of route choice: overcoming the route overlapping problem’, Transportation Research Record 1645, 133–142.

1. With Vi giving the modelled utility of alternative i out of J alternatives, the MNL probability of choosing alternative i is given by . Here, Vi is a function of the attributes of alternative i and *estimated* parameters which include the various constants and marginal utility coefficients listed above. [↑](#footnote-ref-1)
2. Here, we focus on the use of NL models for inter-alternative correlation, rather than NL models applied for estimating models on mixed data sources (cf. Bradley & Daly, 1996; Wen, 2009). [↑](#footnote-ref-2)
3. In a two level NL model with M different nests, where defines the set of alternatives contained in nest m, the probability of choosing alternative i (where i is contained in nest k) is given by , with . [↑](#footnote-ref-3)
4. In the present paper, the general specification also given in Train (2003) is used. Again using different nests, with αjm describing the allocation of alternative j to nest m, we have that

   . Here, the extra summation in comparison with the NL formula ensures that each alternative can potentially belong to each nest. In the present specification, we have two conditions for the allocation parameters, namely , and . [↑](#footnote-ref-4)
5. On a technical aside, the CNL specification works by allocating an alternative by different proportions into different nests, collapsing back to a NL model when all allocation parameters are equal to 1, i.e. an alternative belongs into one nest. In the present context, the allocation parameters were all fixed to a value of 1/2, meaning that an alternative belongs to one airport nest and one access mode nest. The estimation of actual values for the two non-zero allocation parameters for each alternative would have been very difficult due to the high number of parameters and would arguably not have provided any further benefits from an interpretation perspective. [↑](#footnote-ref-5)