JOHN BATES SERVICES

1 Pickler's Hill. ABINGDON OX14 2BA Tel: Abingdon (01235)-526702

Peer Review of NAPALM

prepared for International Networks Analysis and Support Division, Department for Transport

28 October 2010

Contents

1 Introduction

This Report relates to a commission for an Independent Transport Modelling Expert to peer review the **N**ational **A**ir **P**assenger **Al**location **M**odel (NAPALM), a key component of the DfT's suite of aviation forecasting and appraisal models. DfT has provided access to appropriate documentation and has facilitated communication with the consultants Scott Wilson who have designed and implemented the model.

While the model is complex and the timescale relatively short, I believe that I have been able to gain an appreciation of the model's strengths and weaknesses, though some misunderstandings may still remain. I would like to record the fact that the consultants have been extremely co-operative and constructive in answering questions and providing data.

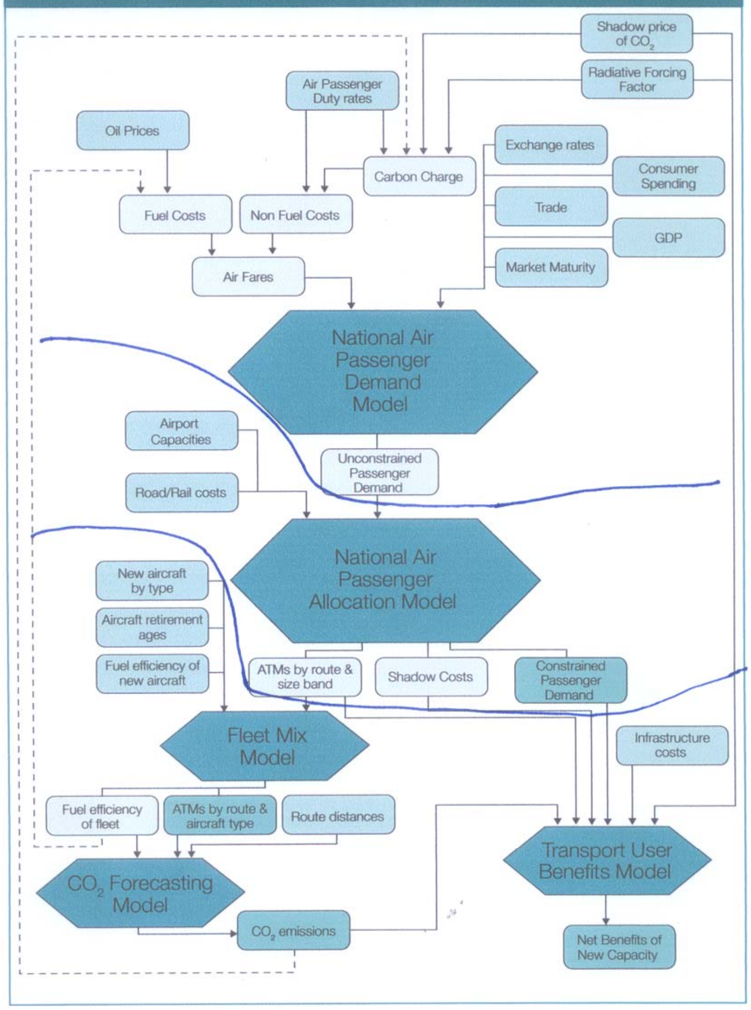
In essence, NAPALM is a model which takes a set of demand forecasts for air passenger travel between zones, and allocates it to UK airports, taking into account capacity restrictions. There are three principal components to the model:

- an airport choice model which, given access costs and service levels at each airport, allocates the input passenger demand proportionately to the various airports;
- a model of air traffic movements (ATMs) which takes the passenger demand between airports and, making assumptions about the aircraft mix, converts this into ATMs between airports (ie at the "route" level);
- a "capacity restraint" process which generates "shadow costs" for particular routes, based on runway and terminal capacity at the UK airports. These shadow costs are passed to the airport choice model in an iterative cycle. In addition, by means of a composite cost formulation, the shadow costs have an impact on total air passenger demand.

All three components are being investigated as part of the review. The review does not relate to any other components of the DfT Aviation forecasting system, and its scope is indicated in the annotated Figure 2.2, taken from the 2009 DfT document "UK Air Passenger Demand and CO2 Forecasts"

It should be noted that this review is being conducted on the understanding that the modelling work has a long history and that the most recent development work, being time-limited, has kept broadly within the spirit of the existing model. For this reason, the review adopts a practical approach in a) investigating whether the model can be described as fit for purpose in the light of these limitations while b) suggesting somewhat more radical departures which could be followed in a less constrained timescale.

Figure 2.2: Overview of UK air passenger demand and CO_2 forecasting framework



REMIT

2 The Airport Choice Model

2.1 Remarks on Choice Modelling

The theoretical analysis of "discrete choices" is well developed, and there are many practical applications. Though recent academic work has tended to more complex model structures, the bulk of practical work is confined to the so-called multinomial logit [MNL] model (and its slightly more complex relative, the nested logit model). The current model adopts an MNL approach, and its method of estimation (using the software BIOGEME) is in line with good practice.

Discrete choice models attempt to find a specification of "(indirect) utility" which best explains the observed choices among a set of options. "Utility" is usually specified as a linear combination of appropriate variables, but linearity is not essential, provided certain desirable properties are met. Setting the utility specification for any two possible options to be equal implies an "indifference curve" between the options, and from this "tradeoffs" (ie, marginal rates of substitution) may be deduced between the explanatory variables. When money costs are one of the elements in the utility function, it becomes possible to convert the utility function into a "generalised cost" equivalent, and this form is more familiar to transport professionals. If money cost enters the utility function in a linear form, then the conversion between utility and generalised cost is no more than a scaling factor[1]. This is the case with all the models that have been used in NAPALM.

The MNL model is an extension to multiple options of the simpler binary logit dealing with the choice between two alternatives. There are some important considerations in estimating choice models which are easiest to explain in the context of the binary logit model.

Generally, choice models may be thought of as an extension to standard regression models in that one is interested in explaining the choice between different alternatives simultaneously. Hence, a utility specification needs to be specified for each alternative. For example, in the binary case, we could have:

$$U_1 = a_1 + \beta_{time}\,\mathrm{Time}_1 + \beta_{cost}\,\mathrm{Cost}_1 + \varepsilon_1$$
$$U_2 = a_2 + \beta_{time}\,\mathrm{Time}_2 + \beta_{cost}\,\mathrm{Cost}_2 + \varepsilon_2$$

where 1 and 2 indicate the alternatives, the β's are the coefficients to be estimated which indicate the relative strength of the variables to be tested (here Time and Cost), and the a's are residual constants (known as "alternative-specific constants" [ASCs]), also to be estimated, which allow for additional "utility" over and above that conveyed by the proposed explanatory variables. As we will see, it is not in fact possible to estimate an ASC for every option – one ASC (arbitrarily chosen) must be preset (typically to zero) to

[1] An alternative metric is "generalised time" in which the units are in terms of time rather than money. Once again, if a single "value of time" is assumed (for any given demand segment), the conversion is achieved by a scaling factor.

ensure identifiability: other ASCs connote the residual utility relative to the selected "base" alternative. The ε terms are "error" terms allowing for a divergence between the modelled and "true" utility (giving rise to the term "Random Utility Model").

Note that in this specification the variables Time and Cost have the same coefficient for both alternatives – this is known as a "generic" specification. While this is not obligatory, allowing the coefficients to vary according to the alternative will typically require special justification (especially in the case of the cost variable).

We now postulate that alternative 2 will be chosen whenever $U_2 > U_1$. Taking the difference (and using Δ to indicate the difference operator for alternative 2 minus alternative 1) this implies that:

$$\Delta U = (a_2 - a_1) + \beta_{time}\,\Delta Time + \beta_{cost}\,\Delta Cost + \Delta\varepsilon > 0$$

It can be shown that in order to obtain robust estimates of the β parameters, we require data where the choices are close to indifference, so that ΔU is approximately zero. When one option is clearly preferable to another (eg if both time and cost are significantly lower), this tells us very little about the relative contributions of the different variables. This is a well-known issue for estimating the "value of time" but it applies generally to choice modelling.

Implications for the model estimation

Ideally, therefore, in estimating an airport choice model, we would require a set of observations in which the choice between at least some of the airports was marginal, in the sense that a very small worsening in the conditions for the currently chosen airport would induce a switch to the next preferred airport. It must be categorically stated that we do not have such data! The available (CAA) data is an assembly of separate airport-specific surveys, collected for a quite different reason. The chances of finding records relating to the same origin and destination zones but actually using different airports are very low. This key fact means that we should lower our expectations for the quality of the models which can be delivered.

Note that when we move from the binary case to the MNL model, it is also necessary to specify, for any given data record, which alternatives are in fact available.

2.2 Market Segmentation

The travel market has been divided into a large number of different segments, as follows:

Charter airlines, subdivided by destination airport = short-haul (SH), North America (US), or rest of world (LH)	3 models

Scheduled airlines to international destinations, similarly subdivided by destination airport, and further subdivided by passenger type (UK Resident/Foreign Resident and Business/Leisure, thus 4 combinations [UKB, UKL, FB, FL])	12 models
"No Frills Carriers" [NFCs] to international destinations (SH only), subdivided by the same 4 passenger types	4 models
The Domestic market is handled slightly differently, since allowance is made for the journey to be made by surface mode. Two purposes are distinguished (UKB, UKL)	2 models
Finally there is a separate market for the interliners [I to I], where no distinction is made either by purpose or destination	1 model

From this, it can be seen that there are a total number of 22 models. It is not possible, in the time available, to give a detailed assessment of all these models, and it has therefore been decided to concentrate on a small number, in particular the Scheduled UK Business North America segment, and the Scheduled UK Leisure Short-Haul segment, with a view to drawing some general conclusions. In addition, because of their inherently different nature, the Domestic and I to I models will be examined.

2.3 General estimation considerations

Given an appropriate choice data set, standard estimation practice choice models would be to define a likely set of explanatory variables, to specify the set of alternatives and indicate which is available, and then to experiment with model specifications. Typically, variables like cost and time would be presented in standard units (eg, £ and hours) and the ratio of the estimated coefficients would be interpreted as the marginal rates of substitution (eg, the value of time). However, it is open to the analyst to pre-specify some or all of the coefficients.

In earlier work (the NATS/CAA model SPAM), a pre-specified generalised cost formulation was used, in which all non-monetary variables were converted to money using appropriate valuations (values of time, etc). If this is done for all variables, then the utility of each alternative is essentially calculable in advance, and all that is required is to estimate a single scaling factor (which can be shown to be related to the variance of the error term ε). The modellers have described such models as "general models". Note that, with this pre-specification, there is still some flexibility as to what items to include in the utility (generalised cost).

In the recent estimation work, the more standard approach of allowing separate coefficients on each variable of interest to be estimated has been followed, in most cases: the modellers refer to such models as "individual models". These terms "general" and "individual" are not standard, but they will be retained for the purpose of this review.

Note that because of the "migration" over time from "general" to "individual" type models, the variables themselves, as presented to the estimation software, have all been transformed into money terms, ie, as if a "general" specification was to be adopted. If the general specification was in fact appropriate, one might expect that an individual specification would deliver coefficients all of the same size (essentially, equivalent to the scaling factor related to the variance of the error term ε). Some concern has been expressed that, in many cases, the results of the "individual" models deviate substantially from this position.

It may be noted, however, that even if the monetisation of the non-monetary variables is being done using "best practice" rules, there is much room for variation, and it would be sensible to allow for considerable leeway in this respect. This will emerge more clearly in the discussion of the explanatory variables below. Further, in terms of model fit, the individual specification will generally perform considerably better, by virtue of being less constrained.

2.4 Discussion of the data

It is helpful to discuss the data in the context of a specific example, and for this purpose the Scheduled UK Business North America segment [Sch_UKB_US] has been chosen. However, most of what is said will apply generally.

The choice data

The basic unit of observation is the number of UKB passengers travelling between one of 455 zones in GB and one of 19 airports in the North American continent (excluding Mexico) and using a scheduled service: within this, a separate record could be permitted for each airport surveyed as chosen. As noted, this data is obtained from the airport-specific CAA surveys carried out at different times between 2005 and 2008: note that the demand figures have not been adjusted to convert to a common year (though some of the explanatory variables are year-specific – see below).

The CAA data is of course sample data: each record is accompanied by a weight which gives a grossed up estimate of the number of passengers. Clearly, at this level of detail (GB zone, purpose, US airport) this cannot be considered as a reliable estimate. However, choice models are usually carried out on sample data, so as long as appropriate weighting is carried out this is not an issue. There is a more serious issue related to the nature of the sample: this could be described as a "choice-based sample". The theory of appropriate re-weighting for this kind of sample is quite complex, but within the confines of this review, this will be treated as an academic issue, at least in the short term.

The total grossed up annual demand for Sch_UKB_US is just under 1.3 million[2]. There are 1001 separate records in the file, though it is not known how large the underlying sample is. Weights have been applied in the estimation so that the total "estimation sample" is equivalent to 1001. This implies that each record is supported by a single sample observation, which is likely to be a conservative estimate. While ideally the weights should reflect the true sample size underlying the data (which cannot be less than 1001), the impact of the current procedure will be to underestimate the precision with which the coefficients are estimated, which may be viewed as a cautionary approach.

Of the 23 GB airports, only 12 provide direct flights to at least one of the US destination airports. In addition, a further 6 allow for connecting flights via a hub (for this segment, the hubs are restricted to LGW, LHR, CDG, AMS, FRA[3]). As noted, the estimation process requires that for each record the subset of available choices (airports) is defined. In terms of specifying the alternatives available for any particular record, a distinction is made between "direct" and "indirect" (connecting via a hub) airports. On this basis, there are 27 possible airport alternatives (6 for connecting flights only, 3 for direct flights only, and 9 airports offering both possibilities). Based on the actual services available to the US destination, however, the number of alternatives ranges from 8 to 22, and the number of direct alternatives ranges from 1 to 10.

The actual demand data at each airport for the Sch_UKB_US segment is given in the following table:

Airport	direct or indirect	Annual demand (PAX)	Percentage of total
BHX	direct	3223.247	0.25
BRS	direct	2658.683	0.21
CWL	direct	102.7308	0.01
EDI	direct	1606.53	0.12
EXT	direct	315	0.02
LGW	direct	90937.92	7.04
GLA	direct	4358.816	0.34
LHR	direct	1020209	78.95
LTN	direct	0	0
MAN	direct	36714.31	2.84
NCL	direct	150.4286	0.01
STN	direct	3524.113	0.27
ABZ	Indirect	8582.584	0.66
BRS	Indirect	1101.467	0.09
EDI	Indirect	19705.56	1.53
EXT	Indirect	0	0
LGW	Indirect	259.25	0.02
GLA	Indirect	11187.31	0.87
LHR	Indirect	6157.876	0.48

[2] Although most of this data refers to 2008, some of it comes from other years. Hence the data in the table is not strictly representative of any particular year.
[3] Dubai (DXB) and MAN were also possibilities, but in order to keep the process manageable, each UK airport was given a single "favourite hub" for indirect routes, and DXB and MAN were omitted on this basis.

HUY	Indirect	1313.025	0.10
INV	Indirect	1330.171	0.10
LBA	Indirect	2040.448	0.16
LCY	Indirect	575	0.04
MAN	Indirect	59671.28	4.62
NCL	Indirect	14749.52	1.14
STN	Indirect	0	0
MME	Indirect	1680.606	0.13

Although this is, of course, a rather specific long-haul segment, it is notable that it is completely dominated by Heathrow. Apart from Gatwick and Manchester, hardly any of the remaining airports even constitute 1% of the total. In terms of the concept of “marginality” discussed earlier, this is not encouraging for airport choice modelling. It is also worth noting that indirect airports constitute only 9.9% of the total demand.

“Level of Service” data

We now go on to consider the explanatory variables used. At the outset it is worth saying that all the variables which one would intuitively expect are present in the estimation data set. However, the form in which they are presented and the robustness of the data is open to some challenge.

The level of service data generally refers to 2008. However, for those survey points where the demand relates to different years, the frequency data is based on the year of survey, this being considered the variable most likely to have changed significantly.

The basic notion of airport choice implies that travellers would take into account the access considerations – both cost and time – and the nature of the flight offered, in particular the fare, the total flying time and the need for interchange. A more problematic variable is the flight frequency, which is subject to different kinds of interpretation, as we discuss below. Another possibility is some characteristic of the airport itself, in terms of the facilities etc which it offers.

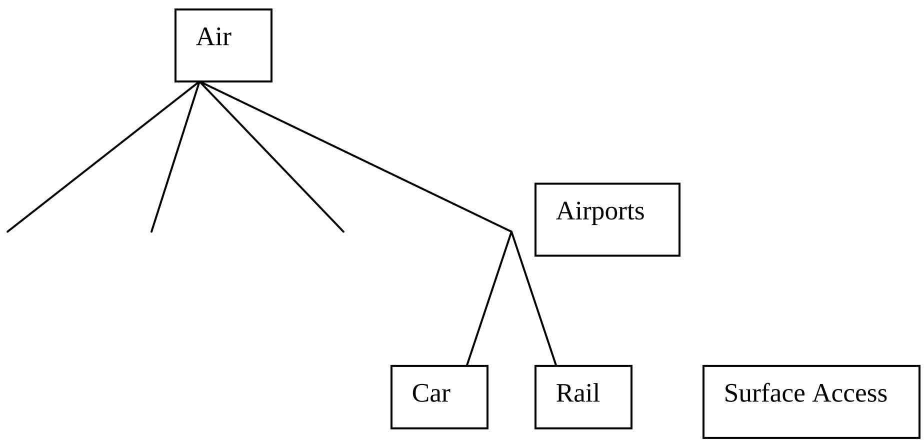
We will now describe the variables actually tested, and in the process provide some commentary. It should be clear from the earlier discussion that it is not the absolute values of these variables which is important, but the **difference** in values as it affects individual airports.

Once again, it should be noted that the current treatment of the explanatory variables typically derives from long-standing model conventions: while ultimately this should not justify it, there is a context in which continuity of model development is important. There are also short-term constraints related to the implementation of the models, in terms of modifications to the programming code (C++).

Surface Access (SA)

Time and Cost data has been derived, separately for road and rail, for each GB zone to each airport (one-way[4]), based on network data from the DfT's Long Distance Travel Model. In the case of road, the cost has been derived from the application of standard WebTAG operating cost formulae on a cost per Km basis. It may be assumed that this data is adequate.

A greater problem occurs in how it is transformed. Theoretically, the intention is to treat it as if it were a mode choice model nested below the main (airport) choice model, as illustrated in the diagram below.



In earlier versions of the model, the access mode data for calibrating such a model was not used, though it is available from CAA. For the recalibration, some attempts have been made to do this, but they have not been judged successful. As a result, it has been decided safest to stay with the earlier methodology, which

- combines the money cost and time data separately for each surface access mode by using segment-specific values of time [VoTs], thus producing a generalised cost in money units;
- assumes a general mode choice scaling parameter of $\lambda = -0.1$ for all segments;
- derives a "composite cost" (also in money units) using standard theory.

The **general** methodology is judged acceptable, and some support has been provided for all the individual assumptions. Nonetheless, this could be challenged on many grounds, *inter alia*:

- as admitted by the modellers, the λ value of -0.1 is high in comparison to the results of specific airport mode choice estimated models (some of which have allowed λ to decrease with distance);

[4] the data is provided as "tours" (round trips) but values are divided by 2 prior to use

- given that λ is in money terms, there is a case for allowing it to vary, at least between business and leisure segments;
- although the VoTs used have been justified at some length (see D0NFLSB/ 10/005), they are based on a number of assumptions which could be contested;
- even for the London airports, car is a dominant mode, and combining car and rail without allowance for mode constants will tend to overestimate the impact of rail cost changes

The preferred approach would be a simultaneous estimation of both mode and airport choice (this has been attempted elsewhere by Stephane Hess[5]). However, it is accepted that this is a more substantial task and is certainly not compatible with the current timescale.

It is important to note that the implicit assumption of a nested model implies that for the airport choice model, the absolute size of the SA coefficient should not exceed 0.1. For the Sch_UKB_US model, this requirement is met, as it is for the Business models generally. However, it turns out not to be the case for many of the other models. We will return to discuss this.

We illustrate the current data with reference to a record from the Sch_UKB_US dataset which refers to travel from Wakefield, Yorks to JFK. Confining ourselves to direct services, there are five possible airports (EDI, LGW, LHR, MAN, STN), and the nearest of these is MAN. The additional SA cost to reach the other four airports is calculated as:

Airport	Additional SA cost from Wakefield relative to MAN (£, one way)
EDI	176.11
LGW	157.99
LHR	124.46
STN	90.61

Wakefield to MAN is about 100 km, taking about 70 mins by car, while to LGW the distance is about 360 km, taking about 4 hours. With the assumed VoT of £51.80 per hour, these values do not seem unreasonable.

Note that as the destination airport is fixed for any single record, the surface access costs at the **destination** end do not vary by the UK airport chosen, so it is not needed in the data.

[5] See, for example, two papers by Hess S and Polak J W (2006): a) Airport, Airline and Access Mode Choice in the San Francisco Bay Area, Papers in Regional Science vol 85 no. 4, pp 543-67, and b) Exploring the Potential for Cross-Nesting Structures in Airport-Choice Analysis: A Case-Study of the Greater London Area, Transportation Research Part E Volume 42, Issue 2, March 2006, Pages 63-81

In Flight Time [IFT]

The data has been taken from the OAG timetable database, and thus relates to published times. As might be expected, the **difference** between flight times from the various GB airports is small. For example, for the Wakefield example given earlier, the values relative to MAN are:

Airport	Additional IFT to JFK relative to MAN (hours, one way)
EDI	0.103
LGW	0.655
LHR	0.041
STN	0.186

It is unlikely that these kinds of differences, which are generally far smaller than the access time differences, will be of consequence in affecting airport choice, especially as, being based on timetables, they are only nominal times anyway. For indirect flights, the times are of slightly more consequence (for the example given, they can be between 1 and 2.5 hours longer), but it will be difficult to distinguish this effect from that of other variables such as the need to interchange.

Hence, while the data itself raises no significant issues, it cannot be expected to play a significant role in explaining airport choice, at least in relation to long-haul movements.

Air Fare

The development of the air travel market and accompanying internet booking facilities and "yield management" makes the extraction of fares for particular journeys an extremely difficult process. The CAA surveys collect information about the fare paid, but there must be some doubt about a) whether respondents actually know the fare, b) the currency in which it is given, and c) whether it relates to the single/return/part journey[6]. In addition, the IPS collects similar information.

According to D0NFLSB/10/014 v1.2 (para 4.5):

> *It was quickly found that using only observed fares considerably restricted the number of valid observations for use in estimation, as fares are often not recorded and are not available for all the possible alternative routes. Restrictions on the depth of the observed and width of the observed choice dataset generally resulted in poorer model fits. Therefore to maximise the potential number of choice*

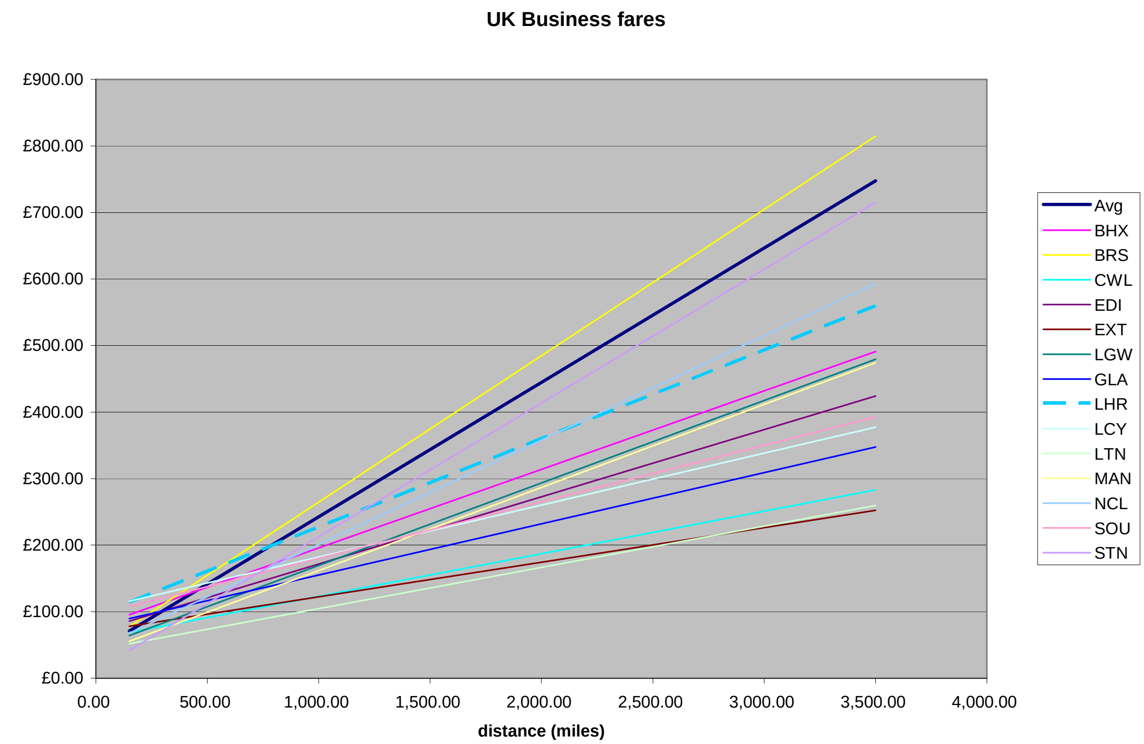
[6] CAA survey records currency and single/return fare information. Often (although increasingly less common) price is printed on the ticket / boarding card, but for business travel, there is the possibility of corporate discounts, etc., which may not be known by the passenger (if they did not book the ticket) or be captured on any documentation that they have. The CAA survey has also not historically tried to capture the ticket price of a passenger travelling on an inclusive tour or one taking a multi-leg journey (eg Paris-London-New York).

observations, work was initiated to generate modelled fares for use in further stages.

This is probably the only practical solution in the circumstances. Subject to sufficient data being available, linear models of fare against distance have been developed for each market (UKB, UKL, FB and FL) at each airport. Nonetheless, there must be considerable error about these estimates, which is particularly important given that, as noted earlier, it is the **difference** by airport which is of relevance. In addition, there is likely to be some endogeneity in the data, since there will be **some** tendency to use airports offering lower fares for the particular movement being made.

The following observations are based on documentation about the fare modelling, but they have not been discussed with the consultants.

It appears that separate fare models of the form Fare = a + b.Distance have been developed for the individual airports BHX, BRS, CWL, EDI, EXT, LGW, GLA, LHR, LCY, LTN, MAN, NCL, SOU, STN. The remainder have been combined (and are denoted by "Avg" in the Figures below). The resulting models for UK Business[7] are plotted (fare against distance) in the Figure below. Because of its special status, LHR is plotted with a dashed line. It should be borne in mind that there is likely to be considerable error in these individual models, so that the confidence interval around the line for any particular airport could be quite wide.



[7] These fares specifically exclude Low Cost Carriers

The modellers note a general tendency for LHR fares to be higher (though while this is true for distances up to 500 miles, for the longer distances higher fares are being modelled at BRS, "Avg", NCL and STN). Nevertheless, if we concentrate on the main alternatives to LHR for scheduled services, ie LGW and MAN, it is the case that LHR fares are about £50-70 higher across the distance range. Given the predominance of LHR as the chosen airport in many of the segments, this clearly presents a problem for the estimation.

It appears that observed fares were used where available, and were supplemented by the modelled values in other cases. This could introduce some bias to the estimation.

The table below presents, for the Wakefield example given earlier, the estimated fare values relative to MAN:

Airport	Estimated Additional Fare to JFK relative to MAN (£, return)
EDI	-291.32
LGW	-211.27
LHR	561.78
STN	318.07

This does demonstrate a sizeable variation, and some of the indirect fares offer even greater savings (of the order of £500 less than the Manchester fare). Even though the long-haul business segment are likely to be rather less price sensitive, such variation – if reliable – might be expected to be of influence on airport choice. The actual choice data, on the other hand, suggests that it is not. It is difficult to assess how much this is a real issue and how much it is a data problem.

<u>Frequency</u>

Conceptually, this is probably the most difficult variable to deal with. In the conventional treatment of urban public transport, service frequencies [Freq] are typically converted to "waiting times", usually by halving the (average) headway (= 1/Freq) on the assumption of random arrivals at the stop/station. On this basis, conventional wisdom is that the value of **headway** is more or less equal to the value of travel time (ie waiting time has double the value of in-vehicle time).

These assumptions are much less appropriate in the case of long distance transport in general, and air travel in particular. As services become more infrequent, travellers will not turn up at random, so that "half headway = waiting time" is not appropriate. Very often the services are not scheduled at regular headways, so the average headway may not be representative. And there are other aspects of waiting time (check-in, security considerations etc) which are more or less independent of the scheduled frequency. At the same time, the frequency does retain some aspect of reflecting travel possibilities, especially the likelihood that a particular departure time will be conformant with individual travel plans. Hence, we may say that “utility” is likely to be positively related to frequency, but the functional form is unclear.

The modellers’ approach is generally in line with these observations but (in my view) they are more influenced by the “half headway” rule than is really appropriate. They have developed a preferred measure referred to as the “empirical frequency formulation” with the form:

$$W = \frac{1}{2} \cdot \frac{16}{F} \left(1 - (1-a)^F\right)$$

where a (referred to as “little_a”) is in the range [0,1]. This can be described as half the headway (if F is the daily frequency, and assuming a 16 hour day) with a damping factor as frequencies decrease. This is then being interpreted literally as a waiting time, and multiplied by the value of in-vehicle time to bring it to money units (NB **not** twice the value of time). The justification for this is set out at length in the SPASM documentation (paras E6-E21), and in D0NFLSB/10/014 v1.2 (paras 5.25-32) where some alternative formulations are also discussed.

An alternative approach would be to make use of a “size variable” to indicate the increased opportunities for the route at the airport. This would be done by taking the natural logarithm of frequency. It turns out that the resulting functions ln (F) and W are very highly correlated! Further discussion will be given below.

As regards the data for frequency, this is obtained from CAA route statistics about departing and arriving flights at each airport. The table below presents, for the Wakefield example given earlier, the daily frequencies for the five direct airports

Airport	Daily 2-way ATMs to JFK
EDI	1.28
LGW	2.15
LHR	38.83
MAN	4.27
STN	2.43

Once again, the dominant position of LHR can be seen.

A greater difficulty is associated with transfer flights. Here there is a more significant difference in interpretation. In terms of the “opportunity” interpretation, it would make sense to take the **minimum** of the two frequencies. In terms of waiting time, it would seem reasonable to calculate the two legs separately and add them: there is also the question of transfer time (referred to as the “interchange penalty”) which appears to be the (one-way) average connection time.

We may note that there has been some potential confusion here between one-way and two-way quantities. For variables which enter the utility function linearly, this does not matter greatly, since the coefficients can be corrected after estimation. However, for the frequency variable, there is a potential problem. If we are talking about one-way "waiting time", then the 2-way frequency should be halved before entering the non-linear "empirical" function. [NB In later documentation these issues have generally been resolved]

In the light of all these remarks, we now go on to discuss the Sch_UKB_US model in some detail.

2.5 Airport choice model for Sch_UKB_US

While a number of models have been estimated for this segment, we will concentrate on the recommended model (no. 41). We have reproduced this model, and have also estimated a small number of variants. The aim has not been to carry out a substantial model estimation, but merely to understand the properties of the recommended model. Note also that since these models were run, minor changes have been made to the data set. Although these will not affect the conclusions, there may be slight differences in the coefficients.

The recommended model is summarised in the table below. It can be seen that all the variables are expressed in £ so that (apart from some concerns about one-way or two-way variables) if a "general" model was appropriate, all the coefficients should be of a similar size. In fact, with the exception of the Air Fare, the coefficients **are** all of the same order of magnitude, though it should be noted that the corresponding "general" model fitted the data much less well.

Number of observations	1001		
Final log-likelihood:	-293.12		
Adjusted rho-square:	0.887		
Variable (description)	**Formula**	**Coefficient**	**T-ratio**
Composite Surface Access (1-way) (£)	SA	-0.0303	-16.26
In-Flight Time Value (£)	VOT*IFT	not included	
Air Fare (2-way) (£)	Fare	-0.000449	-1.90
Direct Waiting Time Value (£)	VOT* $\frac{1}{2}.\frac{16}{F_1}\left(1-(1-a)^{F_1}\right)$	-0.0900	-15.71
Indirect Waiting Time Value (£) – leg 1	VOT* $\frac{1}{2}.\frac{16}{F_2}\left(1-(1-a)^{F_2}\right)$	-0.0574	-5.69
Indirect Waiting Time Value (£) – leg 2	VOT* $\frac{1}{2}.\frac{16}{F_3}\left(1-(1-a)^{F_3}\right)$	-0.0408	-4.19
Transfer Time Value (£)	VOT*IcPen	-0.0503	-9.50

The next thing to note is the extremely high rho-squared. While this connotes a very good model fit, it needs to be viewed with some scepticism, as it actually reveals the lack of choice variation: as noted earlier, the data is dominated by LHR. On this basis, a model which gave a high utility to LHR and zero to all other airports would still score well.

It is generally standard practice to include ASCs in choice model utility functions: this has the advantage of ensuring that the **overall** model shares are exactly reproduced by the model. At the same time, there is a disadvantage that, if the ASCs dominate the estimation, there will be little true explanation, and difficulty in forecasting.

Out of interest, two additional models were run, one including only a set of ASCs, and one adding the set of ASCs to the existing model. The "constant only" model produced a log-likelihood of -799.09 and an adjusted rho-squared of 0.687. This shows that the recommended model is making a very substantial improvement in fit, but also puts the high rho-squared in some perspective – it would be more appropriate to report the improvement relative to the constant only model, ie 0.2.

The model which added ASCs to the existing variables was also encouraging, in two respects. Firstly, the improvement in log-likelihood was not great (-262.80 compared with -293.12) considering the large number of additional parameters required (26), and the great majority of the ASCs were not significant. Secondly, although the inclusion of the ASCs had some impact on the model coefficients and their significance, the key variables (SA and direct Waiting time) held up well, and there were no order of magnitude changes in the other coefficients.

This suggests, therefore, that despite the dominance of LHR, the model variables are playing a useful role. This was also demonstrated in a very simple way by running a model in which the weights were inverted, thus biasing the data away from LHR. Although these results are of no direct interest, the estimated model coefficients were again relatively unchanged. Hence it may generally be concluded that robust effects are being identified.

The next thing investigated was the distinction between direct and indirect flights. Separate models were run so that the choice set was restricted in the first place to direct flights only, and in the second to indirect only. The results are presented below:

	All obs		Direct only		Indirect only	
N of observations	1001		830		171	
Final log-lik:	-293.12		-133.03		-51.99	
Adjusted rho^2:	0.887		0.836		0.753	
Variable	**Coeff**	**T**	**Coeff**	**T**	**Coeff**	**T**
SA	-0.0303	-16.26	-0.0235	-9.83	-0.0368	-7.94
Fare	-0.000449	-1.90	-0.000476	-1.24	+0.000173	0.39
Direct Wait	-0.0900	-15.71	-0.0828	-10.63	n/a	
Indir Wait – leg 1	-0.0574	-5.69	n/a		-0.0693	-4.48
Indir Wait – leg 2	-0.0408	-4.19	n/a		+0.0159	0.60
IcPen	-0.0503	-9.50	n/a		-0.00581	-0.08

This again is generally encouraging, in that it does not suggest that the coefficients are being unduly influenced by the generally negative comparison between the direct and indirect alternatives. In particular, the SA coefficient is reasonably similar between both data sets, and the key Direct Wait and Indirect wait leg 1 coefficients are not strongly influenced by the removal of a subset of the choices. Some doubt is cast, however, on the value of the 2nd leg frequency and the connection time.
Two models were then run to test the frequency specification. The value of little_a for this segment has been fixed at 0.2, but it is also possible to estimate it. Alternatively, the waiting variables can be replaced by a frequency "size" term, given as ln ($Freq_1$) for direct, and ln [min ($Freq_2$, $Freq_3$)] for indirect flights. These two models are shown below.

	Recommended model		estimating little_a		Size effect	
N of observations	1001		1001		1001	
Final log-lik:	-293.12		-281.95		-284.96	
Adjusted rho^2:	0.887		0.891		0.890	
Variable	**Coeff**	**T**	**Coeff**	**T**	**Coeff**	**T**
SA	-0.0303	-16.26	-0.0308	-16.53	-0.0298	-16.60
Fare	-0.000449	-1.90	-0.000704	-2.48	-0.000570	-2.05
Direct Wait	-0.0900	-15.71	-0.0408	-7.02	1.96	15.21
Indir Wait – leg 1	-0.0574	-5.69	-0.0397	-5.23		
Indir Wait – leg 2	-0.0408	-4.19	-0.0239	-4.13	1.83	6.12
IcPen	-0.0503	-9.50	-0.0461	-10.72	-0.0406	-6.55
little_a	0.2	fixed	0.413	7.90	n/a	

Both these models produce useful improvements relative to the recommended model.

Note that the principal effect of the estimated little_a is to reduce the Direct Wait coefficient to a value more in line with the other coefficients (excluding Fare). The Wait coefficients in the "size" model are not scaled, but they indicate that for both direct and indirect choices there is an increased likelihood of going to a particular airport proportional to (approximately) the square of the frequency.

Overall, this brief analysis indicates that the recommended model is generally robust, but there is some room for improved specification of the frequency effect. It is not known whether any examination of residuals has been carried out – casual inspection indicates a limited number of observations which are poorly modelled (the appropriate indicator is the natural logarithm of the modelled probability of the chosen alternative). It is recommended that this is further investigated.

The final investigation is into the overall fit of the model at the airport level: this is shown in the Table below (errors of greater than 30% are highlighted):

Airport	direct or indirect	Annual demand		
		observed (PAX)	modelled (PAX)	mod – obs
BHX	direct	3223.247	2814.219	-409.027
BRS	direct	2658.683	650.500	-2008.180
CWL	direct	102.7308	166.533	63.803
EDI	direct	1606.53	2701.176	1094.646
EXT	direct	315	233.078	-81.922
LGW	direct	90937.92	74605.830	-16332.100
GLA	direct	4358.816	5028.245	669.429
LHR	direct	1020209	1021277.000	1067.768
LTN	direct	0	635.681	635.681
MAN	direct	36714.31	54257.550	17543.240
NCL	direct	150.4286	240.790	90.361
STN	direct	3524.113	1452.225	-2071.890
ABZ	Indirect	8582.584	8835.190	252.607
BRS	Indirect	1101.467	861.488	-239.979
EDI	Indirect	19705.56	19483.130	-222.432
EXT	Indirect	0	12.625	12.625
LGW	Indirect	259.25	176.198	-83.052
GLA	Indirect	11187.31	15171.370	3984.055
LHR	Indirect	6157.876	10545.030	4387.155
HUY	Indirect	1313.025	2538.296	1225.271
INV	Indirect	1330.171	997.222	-332.949
LBA	Indirect	2040.448	7249.557	5209.109
LCY	Indirect	575	3708.842	3133.842
MAN	Indirect	59671.28	43720.630	-15950.600
NCL	Indirect	14749.52	11104.280	-3645.250
STN	Indirect	0	0.865	0.865
MME	Indirect	1680.606	3684.412	2003.805

The worst error is at MAN, but here the overall effect combining direct and indirect flights more or less cancels out. There is also a reasonably large underprediction at LGW. Apart from this, none of the errors is likely to be significant, though the special nature of this segment should be borne in mind.

2.6 Airport choice model for Sch_UKL_SH

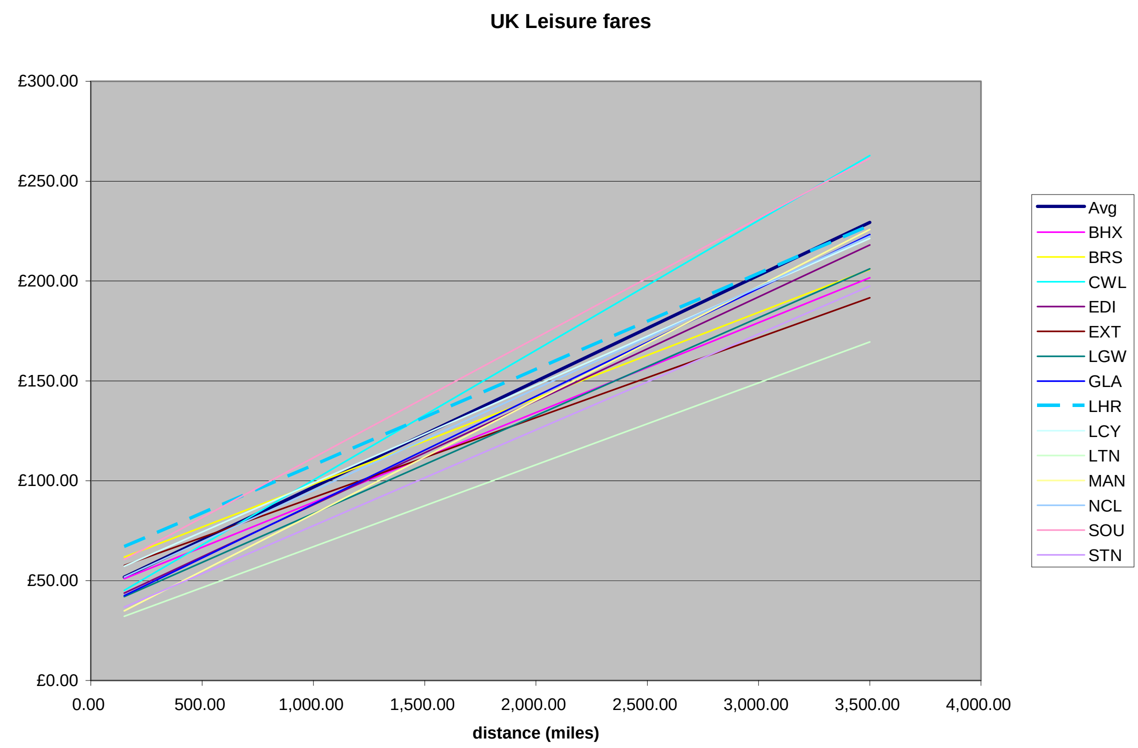
This is a rather larger dataset, and there has been less time to look at it in the same amount of detail. Nevertheless, because it has some rather different features, some comments are worthwhile.

The basic unit of observation is the number of UKL passengers travelling between one of 455 zones in GB and one of 56 short-haul international airports broadly in the European continent (including Israel) and using a "full service" scheduled flight.

The total grossed up annual demand[8] for Sch_UKL_SH is just around 13.4 million, and there are 8539 separate records in the file. All 23 GB airports provide direct flights to at least one of the SH destination airports. In addition, 15 of these allow for connecting flights via a hub (for this segment, the hubs are restricted to LGW, LHR, MAN, CDG, AMS, FRA). On this basis, there are 38 possible airport alternatives.

The actual demand data at each airport for the Sch_UKL_SH segment is given in the table on the next page. While LHR is still the largest airport for this segment, LGW and MAN are not that far behind, with Birmingham and Luton some way further behind. Apart from these, hardly any of the remaining airports even constitute 1% of the total (it should be remembered that the majority of flights from STN are not classed as "full service" scheduled). Indirect airports constitute only 3.4% of the total demand.

The fares model for UK Leisure is illustrated in the following figure:



[8] the remarks made in footnote 2 apply here as well

Airport	direct or indirect	Annual demand	Percentage of total
		(PAX)	
ABZ	direct	31041.41	0.23
BHX	direct	1169852	8.71
BOH	direct	1520.382	0.01
BRS	direct	150273.4	1.12
CWL	direct	69829.01	0.52
EMA	direct	6230	0.05
EDI	direct	241601.6	1.80
EXT	direct	229722.3	1.71
LGW	direct	2713314	20.20
GLA	direct	196146.6	1.46
LHR	direct	4083479	30.40
HUY	direct	23863.19	0.18
INV	direct	0	0
LBA	direct	63087.87	0.47
LPL	direct	99612.57	0.74
LCY	direct	441913.9	3.29
LTN	direct	1069994	7.96
MAN	direct	1839530	13.69
NCL	direct	109422.1	0.81
STN	direct	379092.8	2.82
MME	direct	14432.04	0.11
DSA	direct	38742.37	0.29
PIK	direct	0	0
ABZ	Indirect	54324.02	0.40
BRS	Indirect	15163.5	0.11
EDI	Indirect	109410	0.81
EXT	Indirect	555.9069	.004
LGW	Indirect	0	0
GLA	Indirect	72258.66	0.54
LHR	Indirect	52953.94	0.39
HUY	Indirect	12375.86	0.09
INV	Indirect	13756.93	0.10
LBA	Indirect	16025.61	0.12
LCY	Indirect	4377	0.03
MAN	Indirect	49343.28	0.37
NCL	Indirect	48404.43	0.36
MME	Indirect	12911.07	0.10
PIK	Indirect	0	0

Again there is a general tendency for LHR fares to be higher (though for the longer distances higher fares are being modelled at CWL and SOU). Nevertheless, relative to the next largest alternatives to LHR for scheduled services, ie LGW, it is the case that LHR fares are about £20 higher across the distance range.

The recommended model is summarised in the following table:

Number of observations	8539		
Final log-likelihood:	-5540.201		
Adjusted rho-square:	0.757		
Variable (description)	**Formula**	**Coefficient**	**T-ratio**
Composite Surface Access (1-way) (£)	SA	-0.134	-67.22
In-Flight Time Value (£)	VOT*IFT	not included	
Air Fare (2-way) (£)	Fare	-0.0151	-10.97
Direct Waiting Time Value (£)	$VOT^{*}\frac{1}{2}.\frac{16}{F_1}\left(1-(1-a)^{F_1}\right)$	-0.103	-49.31
Indirect Waiting Time Value (£) – leg 1	$VOT^{*}\frac{1}{2}.\frac{16}{F_2}\left(1-(1-a)^{F_2}\right)$	-0.0812	-5.10
Indirect Waiting Time Value (£) – leg 2	$VOT^{*}\frac{1}{2}.\frac{16}{F_3}\left(1-(1-a)^{F_3}\right)$	-0. 169	-9.61
Transfer Time Value (£)	VOT*IcPen	-0. 187	-21.60

Once again, apart from the Air Fare, the coefficients are all of the same order of magnitude. It is noteworthy, however, that the coefficient on SA is higher than the λ parameter used for the composite cost formulation, which is not acceptable.

As with the previous segment, there is an extremely high rho-squared value, which again should be viewed with some scepticism.

The next thing investigated was the distinction between direct and indirect flights. A model was fitted with the choice set restricted to direct flights only (in view of the relative unimportance of indirect flight a separate model was not estimated for this set). The results are presented below:

	All obs		Direct only	
N of observations	8539		7612	
Final log-lik:	-5540.201		-4693.401	
Adjusted rho^2:	0.757		0.685	
Variable	**Coeff**	**T**	**Coeff**	**T**
SA	-0.134	-67.22	-0.141	-62.14
Fare	-0.0151	-10.97	-0.0156	-10.57
Direct Wait	-0.103	-49.31	-0.105	-47.20
Indir Wait – leg 1	-0.0812	-5.10	n/a	
Indir Wait – leg 2	-0. 169	-9.61	n/a	
IcPen	-0. 187	-21.60	n/a	

This again is encouraging, in that the coefficients relevant for the direct choices are not being unduly influenced by the generally negative comparison between the direct and indirect alternatives.

As with the previous segment, two models were run to test the frequency specification, one estimating little_a (fixed at 0.4 for this segment) and the other using the frequency "size" term. These two models are shown below.

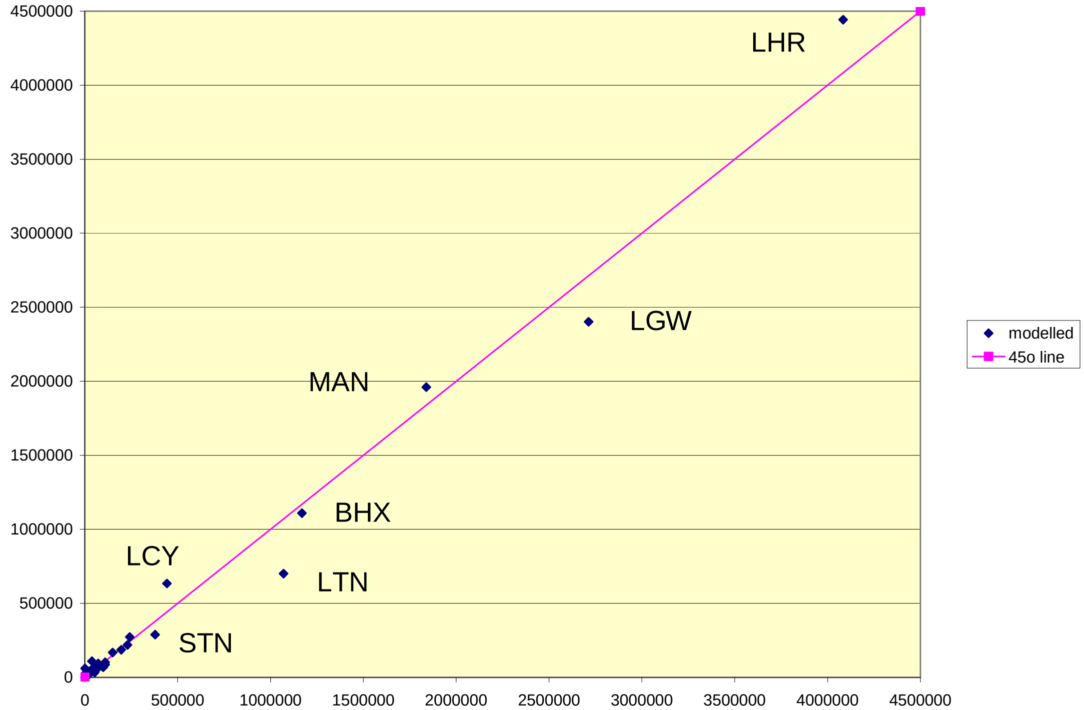
	Recommended model		estimating little_a		Size effect	
N of observations	8539		8539		8539	
Final log-lik:	-5540.201		-5518.188		-5524.513	
Adjusted rho^2:	0.757		0.758		0.758	
Variable	**Coeff**	**T**	**Coeff**	**T**	**Coeff**	**T**
SA	-0.134	-67.22	-0.135	-66.78	-0.134	-67.79
Fare	-0.0151	-10.97	-0.0150	-10.67	-0.0147	-10.79
Direct Wait	-0.103	-49.31	-0.0734	-19.01	1.02	48.82
Indir Wait – leg 1	-0.0812	-5.10	-0.0673	-4.62		
Indir Wait – leg 2	-0. 169	-9.61	-0.132	-8.24	0.845	7.34
IcPen	-0. 187	-21.60	-0.194	-22.55	-0.210	-16.73
little_a	0.4	fixed	0.539	24.30	n/a	

Again, as with the previous segments, both these models produce useful improvements relative to the recommended model. Both models demonstrate that frequency is less important for this segment than for the UKB_US segment. The Wait coefficients in the "size" model indicate that for both direct and indirect choices there is an increased likelihood of going to a particular airport (approximately) proportional to the frequency (as opposed to the square of the frequency in the UKB_US segment).

As before, this brief analysis indicates that the recommended model is generally robust, but there is some room for improved specification of the frequency effect. Examination of residuals is strongly recommended.

The final investigation is into the overall fit of the model at the airport level: this is shown in the Table on the next page (errors of greater than 30% are highlighted):

There are some substantial errors at some of the smaller airports (ABZ, BOH, EMA, LPL, DSA), as well as LCY and LTN. The overall fit is shown in the Figure below. Potentially worrying is the model overprediction at LHR and the underprediction at LGW, in both cases by over 300,000 PAX, although in percentage terms the errors are only 9% and 11% respectively. This does suggest that the allocation between these two London airports could be improved.



Apart from this, the model fit seems reasonable, given the nature of the exercise, but it needs to be seen in the context that the figures will be used. *A priori*, there could be a case for applying the model incrementally at some level.

Airport	direct or indirect	Annual demand		
		observed (PAX)	modelled (PAX)	mod – obs
ABZ	direct	31041.41	45646.76	14605.35
BHX	direct	1169852	1108828	-61024.1
BOH	direct	1520.382	59687.33	58166.95
BRS	direct	150273.4	167882	17608.57
CWL	direct	69829.01	71325.56	1496.553
EMA	direct	6230	29787.22	23557.22
EDI	direct	241601.6	272474.9	30873.33
EXT	direct	229722.3	219099.7	-10622.6
LGW	direct	2713314	2401947	-311367
GLA	direct	196146.6	186583.1	-9563.46
LHR	direct	4083479	4441746	358267.2
HUY	direct	23863.19	23183.58	-679.617
INV	direct	0	0.045184	0.045184
LBA	direct	63087.87	52278.57	-10809.3
LPL	direct	99612.57	67752.76	-31859.8
LCY	direct	441913.9	633312.6	191398.8
LTN	direct	1069994	700883.5	-369110
MAN	direct	1839530	1960967	121436.1
NCL	direct	109422.1	102258.9	-7163.2
STN	direct	379092.8	288119.1	-90973.7
MME	direct	14432.04	19326.7	4894.667
DSA	direct	38742.37	108627.6	69885.25
PIK	direct	0	128.3105	128.3105
ABZ	Indirect	54324.02	54667.52	343.4989
BRS	Indirect	15163.5	10652.38	-4511.12
EDI	Indirect	109410	87075.21	-22334.7
EXT	Indirect	555.9069	2121.819	1565.912
LGW	Indirect	0	798.2548	798.2548
GLA	Indirect	72258.66	94277	22018.35
LHR	Indirect	52953.94	32095.57	-20858.4
HUY	Indirect	12375.86	14652.27	2276.414
INV	Indirect	13756.93	13450	-306.933
LBA	Indirect	16025.61	23769.3	7743.689
LCY	Indirect	4377	10560.56	6183.56
MAN	Indirect	49343.28	67845.41	18502.13
NCL	Indirect	48404.43	43402.43	-5002.01
MME	Indirect	12911.07	17339.62	4428.543
PIK	Indirect	0	7.271406	7.271406

2.7 Conclusions on the "standard" allocation models

The remarks made here need to acknowledge that only two of the 19 "standard" models have been examined in any detail. Nevertheless, the conclusions appear to be reasonably robust. I deal separately with issues which impact on the immediate short term use of the models, and those which could aid further model development.

While I have made some comments on the data problems, I am satisfied that in this respect no useful improvements can be made in the short term, so that the data set being used is the best available. I am also entirely satisfied with the general approach to estimation, having examined the data set-up in some detail. I would only note that, in presenting the models, both the coefficients **and** the estimated t-ratios should be given: the t-ratios are the simplest way of conveying the relative strength of the effect of the variable.

For all the segments, surface access and (direct) frequency emerge strongly as important variables. I have made some comments on the nature of both these variables which will be expanded on below, but in the short term (bearing in mind both the timescale and the restrictions of the existing model code) I consider that these effects are robustly determined. The frequency coefficient is also applied, for the charter segments, to a "charter attractiveness" variable which appears to be a relic of an early model: while I note that this improves the model fit I have not given priority to examining this further.

For the scheduled models, additional variables (frequency of first and second leg, interchange penalty) have been brought in to explain the choice between direct and indirect (hub) flight options between airports. While these are less robustly estimated, they are generally convincing, and in the light of the more limited role that they need to play, I consider that they are satisfactory.

The two other variables which might, in principle, be expected to play a role in airport choice are in-flight time and fare. Neither of these variables have generally produced convincing results. As the implications are rather different, it is worth discussing each variable separately.

It was noted earlier that the **differences** in in-flight times between airports offering the same destination were (as could be expected) very small, making it very difficult to identify an effect in the estimation. This applies particularly to the long-haul traffic: in fact, reasonably convincing coefficients have been obtained for the short-haul models, both NFC and scheduled. From a policy point of view, it would seem unlikely that variations in IFT will play an important role, so that the omission of the variable in the model is probably not critical. In spite of this, there would seem to be some case for including IFT in the preferred short-haul models

The Fare variable is clearly more problematic. Some of the identified differences in fare certainly seem large enough to influence airport choice, but the estimation has not been able to identify this. The estimated coefficients are rarely significantly different from zero, and often have the wrong sign. Even when significant coefficients are obtained, their order of magnitude (compared with, for example, the surface access variable) is far too low. I attribute this to data problems, and indeed other research in this area has concluded similarly. Despite the commendable efforts to interrogate existing fare data and to deduce relationships, I see no immediate possibility for resolving this in the short term.

Since the surface access variable includes an explicit money component (the cost of accessing a particular airport), and it seems reasonable to consider that in choosing an airport, passengers could consider a direct tradeoff between fare and surface access cost, there is an *a priori* expectation that the coefficients on SA and fare should be the same

Whereas it was argued above that IFT is unlikely to be a significant policy variable, this is certainly not true of fare, and hence the absence of a convincing fare coefficient is a significant drawback for the model. With this in mind, the modellers appear to have taken the view that they will keep fare in the model whenever it has a significant coefficient of the right sign (despite the "order of magnitude" problem). I am not convinced that this is the right approach and would recommend that, in the short term, fare is withdrawn.

The solution to this problem, in my opinion, is to adopt an "incremental" approach to the airport allocation model, and to force the model to respond to future **changes** in fare in an appropriate way. However, this must be viewed as a medium term, rather than short term, improvement. I discuss this further below.

These conclusions for the model in the short term imply that

- there will be difficulties in testing fares policy as applied to specific airports or routes (of course, **global** fares policies can still impact (outside NAPALM) on overall air passenger demand);
- given some concerns about the specification of the surface access variable, the model's prediction of the impact of specific **modal** changes in particular corridors on airport choice should be treated with caution.

Turning now to the longer term, the following specific points are identified as worthy of investigation:

re-formulation of the surface access variable, in particular bringing the λ parameter more into line with other studies of airport access, allowing for variation between business and leisure purposes, and investigating how far the current composite variable reflects the balance between the highway and rail costs;

re-consideration of the most appropriate way to incorporate frequency into the model, and relaxing the general "waiting time" interpretation;

considering whether some of the variables might be better included in logarithmic form, as has been found in other studies;

giving some consideration as to whether some "nesting" of alternatives might be appropriate (in particularly between direct and indirect airport choices for the scheduled market);

considering operating the model on an incremental basis

consideration of the most appropriate weighting to deal with estimation using "choice-based" samples.

While each of these would require a more detailed specification than can be offered here, it seems worthwhile to say a little more about the incremental model. Writing the existing model in the form

$$p^k_{A|ij} = \frac{\exp(U^k_{A|ij})}{\sum_{A' \in C^k_j} \exp(U^k_{A'|ij})}$$

where k is the segment, A is a particular airport, i is the passenger origin, j is the foreign destination, and C^k_j is the set of airports providing service to j for segment k, the probability of choosing airport A is given by the logit function based on the estimated utility for each available airport.

The corresponding incremental version of the logit model has the form:

$$p^k_{A|ij} = \frac{D^k_{ijA} \exp(\Delta U^k_{A|ij})}{\sum_{A' \in C^k_j} D^k_{ijA'|} \exp(\Delta U^k_{A'|ij})}$$

where D^k_{ijA} represents a base year allocation of total demand between i and j to each airport, and $\Delta U^k_{A|ij}$ represents the **change** in utility relative to the base year. With this formulation, changes in airport choice are brought about by changes in the relevant explanatory variables affecting utility, and it becomes possible to include components (such as fare) which could not be introduced into the estimated model.

Clearly, moving to this formulation requires the base year matrix D^k_{ijA} to be explicitly derived. In practice, this would be done by a combination of the estimated model and the observed (CAA) data applied at a level at which it could be considered reliable. In practice, this would mean aggregating the CAA data at some geographical level, using local data (eg TEMPRO) to disaggregate to the origin zone, and the estimated model to disaggregate to airports. The process would ensure that in the base year the overall allocation to airports was correct for each segment.

In point of fact, some movements in this direction have already been made. Base Year matrices are created for air travel between ultimate origin zone (i) and ultimate destination zone (j): however, at an interim stage they take account of the airport used (A) and are corrected to achieve some reconciliation with the A-j demand in CAA statistics for the survey year. The adjustment is achieved by means of generalised cost constants added at the route (A-j) level, somewhat poorly named "fare adjusters" as briefly referred to in paragraphs E.53 & I.209 of “Rules and Modelling”. In reality they are added into surface access cost in a manner similar to that of shadow costs (see section 4 below). Obviously such constants could be dispensed with if the model became incremental.

2.8 The Domestic Models

The domestic models differ from the standard allocation models in that they allow for the possibility of a surface mode choice, in addition to the choice of airport. Only direct flights are considered. Separate models are estimated for business and leisure passengers, but the model relates to both NFC and scheduled flights.

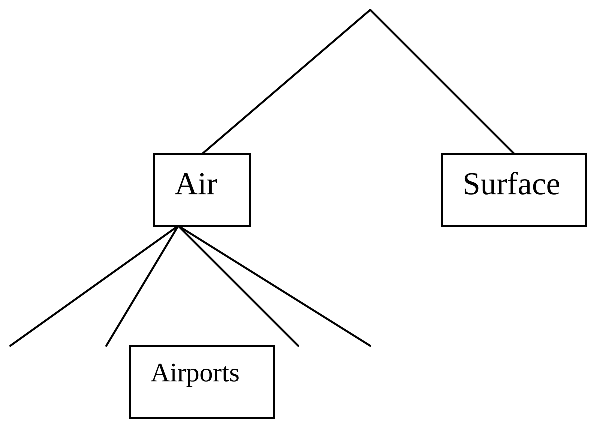
In contrast to the “standard” models which are all of the “individual” type, the domestic models are of the “general” type: thus they involve a pre-specified utility function, and only the “spread” parameter is estimated.

Because the model is essentially a form of mode choice model between a UK origin and destination, the surface mode option represents a (composite) cost for the whole journey, and the airport costs need to reflect the egress cost as well as the access cost. In all other respects, the variables are as defined for the “standard” models.

In order to estimate the model including the (composite) surface mode, it is necessary to have choice data for that mode as well: this is of course not available from the CAA data. The surface mode choice data has therefore been taken from the current version of the long-distance model.

It is important to realise that this data is quite different in kind from the CAA data: it is essentially synthetic data, and its relation to sample information is unknown. It appears that the data has been selected to coincide with the observed origin-destination pairs in the CAA data, though not all the O-D pairs are represented for the surface mode (for example, in the Leisure data, there are 2284 records relating to air travel, and allowing for some multiple occurrences of O-D pairs this gives 2136 O-D pairs, but only 782 of these have associated surface records). Clearly, the choice data needs to be weighted for representativity, and it is not clear how much reliance can be put on this.

However, even if we accept the surface mode choice data at face value, the assumption that the error term for each airport utility is of the same magnitude as that for the surface mode (as is implied by the MNL) does not seem very credible. Indeed, such a structure more or less cries out for a nested model approach, along the lines of the Figure below.



My attempts to estimate such a model using BIOGEME produced, for both purposes, a huge improvement in log-likelihood, indicating that the MNL model is strongly rejected. The results are shown in the Table below (note that the t-statistics for the "scale" coefficients test the significance of the difference from 1.0, which is the value implied by the MNL model):

	Domestic Business			
	MNL		Nested	
N of observations	3587		3587	
Final log-lik:	-4064.52		-3588.03	
Adjusted rho^2:	0.649		0.689	
Variable	**Coeff**	**T**	**Coeff**	**T**
Utility	-0.0525	-52.92	-0.0146	-18.18
Scale	1	n/a	5.22	13.49

	Domestic Leisure			
	MNL		Nested	
N of observations	3066		3066	
Final log-lik:	-4266		-3577.86	
Adjusted rho^2:	0.595		0.659	
Variable	**Coeff**	**T**	**Coeff**	**T**
Utility	-0.1620	-47.37	-0.0317	-16.30
Scale	1	n/a	8.23	13.13

The implication of the nested model is that, in both cases, the choices between airports are far more sensitive (by a factor of 5.22 for business and 8.23 for leisure) than the choice between air and surface. The result is that the MNL model will both underestimate the amount of substitutability between airports and overestimate the substitutability between air and surface modes.

It is desirable to correct this as soon as possible. Note in addition that, as with most of the international models, the utility coefficients for the **Leisure** segment are incompatible with the definition of the Surface Access composite cost.

The general remarks made in Section 2.7 apply in most cases to the Domestic model also.

2.9 The I to I Model (interlining)

The I to I model relates to a limited and somewhat artificial choice: essentially, of those international passengers for whom neither end of the journey is in the UK **and** who elect to interline at a UK airport, what proportion will choose each available UK airport? The choice data is gathered from CAA interviews carried out at departure gates.

Only 5 of the airports are of relevance here (LGW, LHR, STN, LTN, MAN), and the extent of interlining at LTN and MAN is trivial. No distinction is made by purpose. The definition of interlining includes "informal" transfers between NFC flights, since a considerable number are observed (particularly at STN) in the CAA surveys.

The variables tested are Fare[9], IFT and Frequency (for both legs): surface access is, of course, not relevant for this segment. Attempts were made to estimate an interchange penalty, but this is not feasible because it applies to all airports.

The recommended model has the following form:

Number of observations	1769		
Final log-likelihood:	-507.11		
Adjusted rho-square:	0.693		
Variable (description)	**Formula**	**Coefficient**	**T-ratio**
In-Flight Time Value (£)	VOT*IFT	-.0402	-1.36
Air Fare (2-way) (£)	Fare	not included	
Indirect Waiting Time Value (£) – leg 1	$\text{VOT*}\frac{1}{2}.\frac{16}{F_2}\left(1-(1-a)^{F_2}\right)$	-0.0927	-23.55
Indirect Waiting Time Value (£) – leg 2	$\text{VOT*}\frac{1}{2}.\frac{16}{F_3}\left(1-(1-a)^{F_3}\right)$	as above	as above

Note that a model with ASCs only produced a log-likelihood of -559.05 (adjusted rho-square = 0.660), so that the variables are not making a major contribution. In any case, the IFT coefficient is not significant.

A small number of additional models were tried. In the first place, in line with previous work, two models were run to test the frequency specification, one estimating little_a

[9] A special "hub" fare model was designed for this purpose

(fixed at 0.35 for this segment) and the other using the frequency "size" term. These two models are shown below.

	Recommended model		estimating little_a		Size effect	
N of observations	1769		1769		1769	
Final log-lik:	-507.11		-468.819		-524.88	
Adjusted rho^2:	0.693		0.715		0.682	
Variable	**Coeff**	**T**	**Coeff**	**T**	**Coeff**	**T**
IFT	-.0402	-1.36	-0.0313	-0.92	-0.0508	-1.86
Indir Wait – leg 1 Indir Wait – leg 2	-0.0927	-23.55	-0.234	-6.57	1.78	26.13
little_a	0.35	fixed	0.110	6.75	n/a	

As with the previous segments, estimating "little_a" improves the model fit substantially, though it reduces the significance of IFT still further: in this case, little_a is much lower than the assumed value, implying that "waiting times" are heavily damped. The "size" model does not, however, produce an improvement, though the importance of the alternative frequency variable is still supported.

A final model was attempted, which estimated separate coefficients on the two frequency variables.

	Recommended model		separate frequencies	
N of observations	1769		1769	
Final log-lik:	-507.11		-496.40	
Adjusted rho^2:	0.693		0.699	
Variable	**Coeff**	**T**	**Coeff**	**T**
IFT	-.0402	-1.36	-0.0406	-1.35
Indir Wait – leg 1	-0.0927	-23.55	-0.0732	-13.68
Indir Wait – leg 2			-0.118	-16.03
little_a	0.35	fixed	0.35	fixed

This suggests, not implausibly, that the frequency effect is more important for the second leg, and delivers some improvement to the model.

Overall, the model is not very impressive, but is probably acceptable in the circumstances. Apart from reflecting on the treatment of frequency, no particular recommendations are made for its improvements. However, the general remarks made in Section 2.7 apply in most cases to this model as well.

3 Converting to air traffic movements

[Note: in the time available to date not all the issues have been resolved. In addition, the process for allocating new routes in future years has not been examined.]

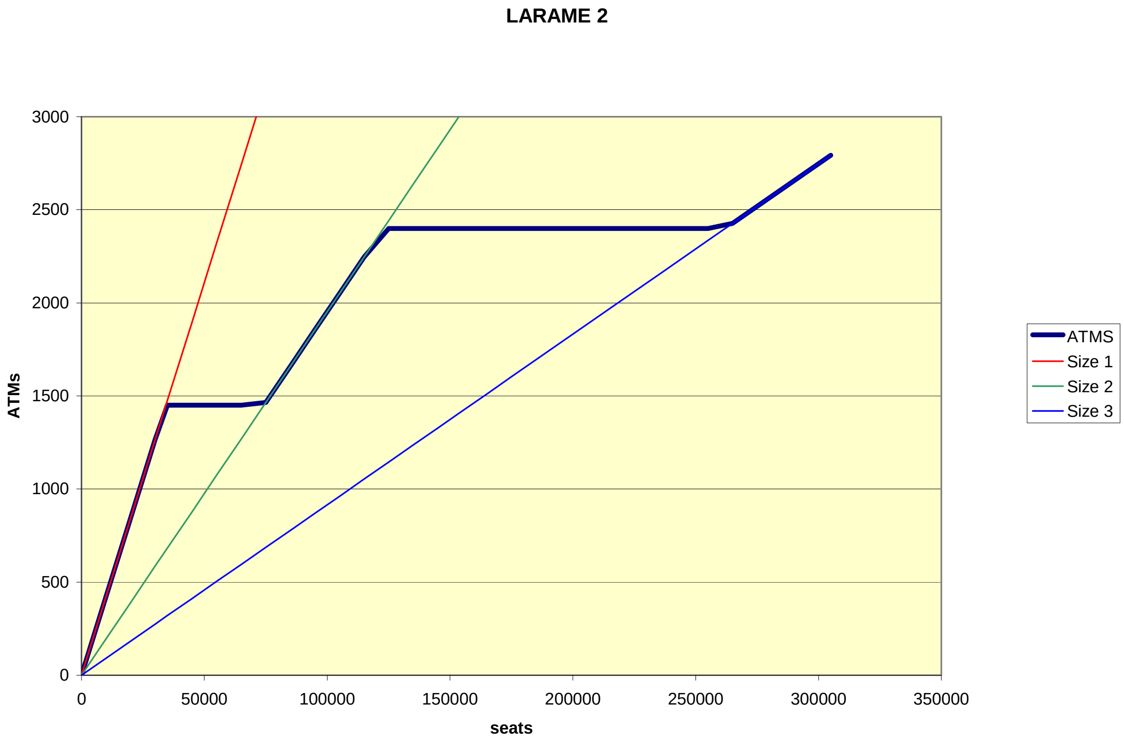
3.1 Larame Graphs

According to "Rules and Modelling" [para C.27], "Larame graphs are a well established procedure to calculate the number and size of planes on groups of routes taking into account numbers of seats available per plane and seat factors which give the number of seats occupied relative to available seats." I should note, however, that I have not been able to find any independent reference to Larame graphs. Para C.27 makes reference to an internal document "NATS FAG Paper 1, SPAM Larame Graphs", a copy of which has been provided, but this also gives no backward reference. It seems, therefore, that this is a terminology developed for the original CAA version of the model.

In spite of this, the aim and general methodology are relatively straightforward. Essentially we require a conversion between passengers and planes, and clearly there are two key factors – plane size (seats), and the level of utilisation (proportion of seats occupied). There are of course commercial/cost considerations here: the level of utilisation will be related to service frequency, and this has to be reasonably conformant with the pattern of demand (in terms of preferred timing). There may also be limitations on plane size at some airports.

The Larame Graph converts between (total available) seats and ATMs, on a yearly basis. The general form is that initially a constant average plane size is assumed, so that ATMs are directly proportional to seats. However, when a defined threshold is reached, there is a transition phase to a larger plane, and during this phase, it is assumed that growth in seats can be achieved without an increase in ATMs. Once the appropriate seat level for the larger plane is reached, ATMs again increase proportionally to seats. The pattern may repeat with several sizes of planes.

An example is given below. Each successive plane size is denoted by a "ray" from the origin, converting from seats to ATMs at a constant rate, with horizontal sections denoting the transitions from one plane size to the next. Thus the graphs require only the number of seats for each plane size, together with the "threshold" in terms of ATMs at which the transition begins.



According to “Rules and Modelling”, “Larame graphs are based on extensive statistical analysis of plane size mixes and seat demand [??] on routes made by the CAA.” It is also noted that “Each Larame graph contains a minimum passenger threshold, which by default applies to each route to which the graph is allocated.” However, in implementation, these values can be overridden (paras. C.18, C.32 and I.202 of “Rules & Modelling” refer)

According to FAG Paper 1, Larame graphs were (originally) differentiated by the following variables:

- Flight type
- Distance
- UK Airport
- Competition Level
- Plane type (Jets and Turboprops)
- Routes (Individual and Grouped)

In the latest version, there are 244 graphs, and the differentiation appears to be along the lines of:

- Long/Short some general description (eg. W Europe) and some very detailed
- Type Charter, LCC, Sch, Scheduled, Blank (with notes)
- Airports various combinations, and some individual
- Airlines some specific references (eg to RyanAir)

While more documentation would be welcome, and in particular a justification of the number of different graphs, and the way they are allocated to particular routes, the process itself does not raise any technical issues of concern.

3.2 Converting from Passenger Demand to Seats Available

To achieve the conversion from **passengers** to ATMs, we also require information about the level of utilisation (average load factors – ratio of passengers to available seats). This is available as a separate element defined for each Larame graph.

In some cases, the graphs are duplicated, and only the load factors are different. For example, graphs 34-36 all relate to "LGW Trunk International LF" with identical parameters, but they are differentiated by three levels of load factors – low (0.65), mid (0.7) and high (0.76). There is no automatic transition linking two similar graphs with different load factors. However, the opportunity to redefine a new graph every forecast year can be used to increase load factor by graph (where appropriate) year on year in line with recent trends for the routes represented and a view on how close the load factor is to the practical maximum for this segment – this is a matter for expert user judgment. There is also a mechanism whereby load factor/aircraft size is effectively increased, within user defined limits, when an ATM shadow cost is first imposed at an airport (see "Rules and Modelling" para H.35).

Otherwise, given the appropriate Larame graph, the total passenger demand is simply converted to available seats by dividing by the load factor. The number of ATMs is then read off from the graph.

Given that the load factors are an essential element in the conversion between passengers and ATMs, some support for the values should be provided: if there is significant uncertainty about some of them, then sensitivity tests should be carried out.

4 Capacity Constraint and Shadow Costs

4.1 The basic approach: suppression and re-allocation

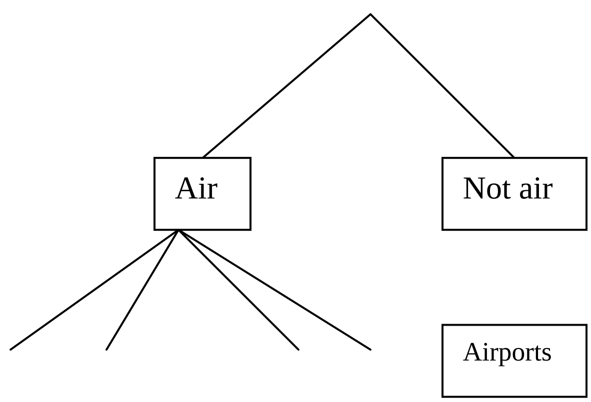
There are two capacity constraints active in NAPALM, one relating to aircraft movements ("runway") and one to passenger movements ("terminal"). In the case of runway constraints, the passenger movements need to be filtered, via the Larame curves, to create the corresponding estimates of ATMs. Nonetheless, the principal mechanism for dealing with the constraint is, in both cases, the allocation of **passengers** to airports.

There are various ways in which the algorithm might proceed to ensure that the constraints are met, and it is not intended to discuss the details of the search procedure. However, it should be noted that the current approach essentially selects an initial per passenger “shadow cost” for each affected airport which is added to the generalised cost (in £): the shadow cost estimate is then iteratively updated until the constraints are met.

We will first briefly explain and comment on the mechanism itself, and then discuss the issue of “shadow costs” in more detail.

The addition of the shadow cost to the generalised cost for any given airport will straightforwardly affect the allocation to that airport within the choice model: a positive shadow cost will divert passengers away to alternative airports. As is usual within choice models, the extent to which this diversion occurs will depend on how “close” (in generalised cost terms) the alternatives are.

In addition to the application of the airport choice, the model is used to reduce the overall demand for air travel (“suppression”). This is done by calculating the impact on the (higher level) composite cost taken over all airports: this should not be confused with the use of composite cost for the surface access variable discussed in Section 2. Essentially, the model has the form below, where the “Not Air” box refers to any conceivable option to making the journey by air (which will usually mean not making the journey).



As before, the existing model for the probability of choosing airport A can be written in the form

$$p^k_{A|ij} = \frac{\exp(U^k_{A|ij})}{\sum_{A' \in C^k_j} \exp(U^k_{A'|ij})}$$

where k is the segment, A is a particular airport, i is the passenger origin, j is the foreign destination, and C^k_j is the set of airports providing service to j for segment k. The composite utility over all airports, for a particular (i j k), is given by the natural logarithm of the denominator, ie:

$$U*^k_{ij} = \ln \sum_{A' \in C^k_j} \exp(U^k_{A'|ij})$$

and can be interpreted as the expected maximum utility from the airport choice.

Any **change** in this composite utility can be used to influence the overall demand for air travel, for a particular (i j k), by means of an implied higher level choice process (between air travel and some alternative). There are two reasonably standard ways of doing this.

The first is to posit an incremental choice structure between "air" and "not air", so that the revised demand for "air" D′ is given as:

$$D'^{k}_{air|ij} = D^{k}_{air|ij} \cdot \frac{D^{k}_{air|ij} \exp(\theta.\Delta U *^{k}_{ij})}{D^{k}_{notair|ij} + D^{k}_{air|ij} \exp(\theta.\Delta U *^{k}_{ij})}$$

under the assumption that the utility for the "not air" does not change. The problem with this formulation is that it is more or less arbitrary how the base demand for "not air" is defined.

The second method avoids this problem, by ignoring the denominator, so that:

$$D'^{k}_{air|ij} = D^{k}_{air|ij} \exp(\theta.\Delta U *^{k}_{ij})$$

In both cases, an extra parameter θ is required, which should be in the range $0 \le \theta \le 1$. Essentially, this can be calibrated to achieve a desired elasticity (provided the condition is met).

The methodology described in section H of the core document "Rules and Modelling" is not completely consistent with this. In the first place, the composite utility is converted to composite generalised cost in monetary units: this raises no issues when dealing with a "general" type of choice model, but – as we shall see below – causes some difficulty when using an "individual" type model.

Secondly, rather than using the difference in composite generalised cost with and without the shadow cost element, the ratio is used. In other words, using comparable notation:

$$D'^{k}_{air|ij} = D^{k}_{air|ij} \cdot \left(\frac{C' *^{k}_{ij}}{C *^{k}_{ij}} \right)^{\eta}$$

This has the advantage of allowing the chosen elasticity (η) to be used directly, but there are dangers in using "generalised cost elasticities" in this way, since the ratio is sensitive to the number of components included in the generalised cost formula. It would be better to change to the formula based on the cost (or utility) difference, given earlier.

It should be noted that there is an odd reference to "borrowing" the surface access composite cost scaling parameter of –0.1 (p. 83 of "Rules and Modelling"): this would be illogical. However, the calculations illustrated in the example which follows on the same page do not support this. It is recommended that the actual code is carefully checked.

The suppression elasticities are provided on p. 84 of "Rules and Modelling", and are implicitly treated as fare elasticities. At the time of writing, it is not known whether these have been updated: the values given for international scheduled (not I to I) passengers are based on those “used by the DfT in their sensitivity analysis of air traffic forecasts, contained in DETR 2000 Air Traffic Forecasts.” There is likely to be a considerable element of judgment here. Nor do I find the argument about the similarity between domestic and international elasticities very convincing (it appears that the surface mode component is excluded from the calculation of the composite generalised cost for domestic, though it could be argued that at least some of the suppression would be achieved by switching to surface modes). A clearer account is needed.

Whichever approach is taken to the calculation of the revised total demand in the context of capacity constraints, the resulting figure then needs to be multiplied by the airport shares as modified by the inclusion of the shadow cost. The mechanism for doing this appears to be correct.

4.2 “Shadow Costs” in the context of NAPALM

4.2.1 Introduction

Having described the suppression mechanism in outline, we now turn to a detailed discussion of shadow costs.

The notion of shadow prices is fundamentally associated with capacity constraints, and is probably most commonly encountered with respect to production functions in economic theory. Formally, it may be related to Lagrangean multipliers in a constrained optimisation program. As the objective functions in economic theory are typically expressed in money terms, the shadow price has a natural interpretation in terms of the marginal benefit of relaxing the constraint.

The formal mathematical approach to constrained optimisation will lead to one shadow price in respect of each constraint potentially exceeded (if the constraint is not binding, the shadow price is zero). However, the interpretation of the shadow price becomes more complicated when we move away from pure economic considerations (cost minimisation or revenue maximisation).

If a constraint becomes binding, the classical optimisation approach will set the value of the related variable to the constraint, as opposed to “within” the constraint. However, from a heuristic point of view, a search process might propose a value of the shadow price which brings the variable to a value lower than the constraint. While this might be considered acceptable, the implication is that the shadow price is in fact too high – a lower price could be proposed that would still meet the constraint. Note that this has some repercussions for the conditions for terminating a search algorithm

With these general considerations in mind, we will discuss the issue of shadow prices in the transport modelling context. Because the so-called “gravity” model is well known, it probably makes sense to use this as an example, rather than the NAPALM model itself. However, the conclusions will be related to NAPALM in the final section.

4.2.2 The gravity model of distribution or destination choice

In its simplest form, the “gravity” model proposes that the number of trips “produced” in a particular zone i [P_i, say], will be allocated to destinations j in a manner proportional to the attractions in each destination [A_j, say] and to some negative function of the (generalised) cost C_{ij}. It is common to use the negative exponential function for this purpose, and we will do this, especially as this is consistent with the logit function for choice models.

With this in mind, we can write T_{ij}, the number of movements between i and j, according to the following formula:

$$T_{ij} = a_i P_i A_j \exp(-\lambda C_{ij})$$

Note that this implies that $a_i A_j \exp(-\lambda C_{ij})$ is the proportion of P_i which goes to j, and this allows us to re-write the model as one of destination choice, conditional on origin, in logit form:

$$p_{j|i} = \frac{A_j \exp(-\lambda C_{ij})}{\sum_j A_j \exp(-\lambda C_{ij})}$$

In the so-called singly constrained gravity model illustrated here, the A_j terms may be considered merely as indices of attractiveness: the model will not deliver the result that $\sum_i T_{ij} = A_j$. However, in some contexts, such as the journey to work, this equality is desirable. To achieve this, we require that the equation $\sum_i T_{ij} = A_j$ is imposed as a constraint. As is well known, this introduces a further multiplicative term b_j to the gravity formula, so that:

$$T_{ij} = a_i P_i b_j A_j \exp(-\lambda C_{ij})$$

Thus, the mathematical outcome is that there is a set of destination multipliers which can be calculated so as to ensure that the constraint is satisfied. As noted earlier, there will be one multiplier associated with each (independent) constraint.

The formula translates straightforwardly into the logit form, again by means of destination multipliers:

$$p_{j|i} = \frac{b_j A_j \exp(-\lambda C_{ij})}{\sum_j b_j A_j \exp(-\lambda C_{ij})}$$

It now becomes possible to “interpret” these multipliers in terms of generalised cost. If we write $\exp(-\lambda \beta_j) = b_j$, then the logit form becomes:

$$p_{j|i} = \frac{A_j \exp(-\lambda[\beta_j + C_{ij}])}{\sum_j A_j \exp(-\lambda[\beta_j + C_{ij}])}$$

and β_j is therefore the implied shadow cost of the destination constraint.

So far, so good. We now go on to the problems that arise within the transport modelling approach. These are of two, related, sources:

- multiple segments with different choice models;
- generalised cost components

Suppose that we have two separate "segments" competing for the same destinations: the segments have been distinguished **because** their choice mechanisms are different. In the simplest possible case, assume the following set-up:

$$T^1{}_{ij} = a^1{}_i P^1{}_i A_j \exp(-\lambda^1 C_{ij})\,;\; T^2{}_{ij} = a^2{}_i P^2{}_i A_j \exp(-\lambda^2 C_{ij})\,;\; \sum_i (T^1_{ij} + T^2_{ij}) = A_j$$

Note that we are assuming here that the generalised cost C_{ij} is the same for both segments: this is a restriction which may well not obtain in practice.

We wish to impose the constraint implied in the last equation on the choice models for the two segments. In line with the mathematical approach given earlier, this will imply one additional coefficient for each destination. In the absence of any other evidence, we could calculate a factor b_j which would be applied equally to both segments. However, if we do this, it should be clear that because the λ coefficients are different between the segments, the implied shadow costs β_j will be different. Alternatively, we could require that the shadow costs are the same, but then the factors will be different, with the result that, in proportional terms, the constraint will impact more on one segment than the other.

Other approaches are possible. We could, for example, pre-specify the impact on each segment, in relative terms, by saying that the implied multiplier b_j for segment 1 should be, say, double that for segment 2, because we thought that one segment was more sensitive to the constraint than the other.

Any of these approaches, and indeed an infinity of alternatives, can be made to achieve the constraint. The mathematical fact is that the destination constraint does not recognise the segmentation: it applies to the combined demand, and it has nothing to say about how the impact of the constraint should be allocated between the segments. This means that further assumptions are required on the part of the modeller. As has been shown, even apparently innocuous alternative assumptions – that the multiplier is the same, or that the shadow cost is the same – will lead to different outcomes, and independent judgment will be required to decide which set of assumptions is to be preferred. This needs to be unambiguously stated, since there is a danger that incorporating an arbitrary choice into the solution algorithm could lead to an undesirable outcome.

The situation is only made worse when we consider that the "generalised cost" measure used in transport models is a combination of different components (time, cost etc). Hence, it becomes possible to translate the shadow cost into any of these components. If the weights in the generalised cost function were the same for all segments, this would not be of any significance, but in the face of different specifications, it introduces a further aspect of potentially arbitrary judgment.

Since standard economic treatment of shadow prices naturally deals with them in money units, it might seem that this approach is appropriate in the transport case as well, but this cannot necessarily be assumed correct. Indeed, most transport models take the view that – in relation to inter-segment operations, as well as for forecasting purposes – time is a more appropriate numéraire. But the point is that some kind of judgment is required, and the outcome will be different depending on what approach is adopted. *A fortiori*, it should be clear that this has potentially major implications for the treatment of "shadow costs" in appraisal.

4.2.3 Implications for NAPALM

The airport allocation models relate to 22 different segments, which need to be combined to produce the demand at each airport (since the capacity restraint only applies to the total passengers). The coefficients on the various components of "generalised cost" vary between the segments as, to some extent, do the actual components included. Thus the problems illustrated by the gravity model example arise cogently. In addition, the conversion to ATMs introduces a further differentiation, since there are many more Larame graphs than the 22 segments.

In the case of the "general" airport choice models, there is essentially a unique definition of generalised cost in money terms, so that the shadow cost can be added directly: from this point of view, the only issue in relation to segments is in terms of the variation of the generalised cost parameter, which, comparably with the gravity model example, will transform the shadow costs into different multipliers. Further issues arise, however, for the "individual" airport choice models, since although the various components of generalised cost are all converted to monetary units, the weights on them vary. Hence, a decision has to be made about which component the shadow cost should apply to. This is exacerbated by the fact that the set of components varies between the segments. Currently only surface access and frequency variables are common to all models (with the exception of I to I, where there is of course no SA variable).

Some argument could be made along the lines of reflecting how the airports/airlines might actually respond to the capacity issues, and it would not be unreasonable to propose that an increment to the fare per passenger might be the most likely mechanism. In this respect it is unfortunate that few of the airport choice models have been able to include the fare as one of the determining variables (for reasons discussed elsewhere).

However, even if we accept the fare "mechanism", this would still leave room for further assumptions as to the **distribution** of the increment, eg in relation to overall fare, allowing for variation with say ticket type, long-haul/short-haul etc. While it could be argued that a uniform increment is "fair" on the basis of the actual costs being essentially equal regardless of the passenger circumstances, it is not clear that either the airlines or the airports would want to impose it in this way (unless we are talking about a trivial amount). Certainly an option of allocating in proportion to fare should be considered. Other options that occur would be a) to focus the cost on the likely source of capacity problems (eg peak times) or b) the inherent profitability of different routes, so as to ensure that the "suppressed" demand (whether lost to air or merely diverted to another airport) was the least valuable.

In this context, we also need to discuss the difference between terminal and runway constraints. In the latter case, a uniform shadow cost **per ATM** is applied, but this is then differentially converted to a per passenger cost by using the Larame graphs (including the average load factor). The result is (“Rules & Modelling”, para I.28) that: “Differential shadow costs therefore apply according to different route and traffic types at the same airport when ATM constraints apply.” The practical consequence is that routes at the airport which use smaller planes will be allocated a higher per passenger shadow cost, thereby deterring such passengers to a greater extent. This will tend to discriminate against short-haul rather than long-haul routes.

It is assumed that the mechanism can be represented along the following lines (for convenience we ignore the domestic and ItoI segments, where the relation between passengers and ATMs is slightly different):

Using compatible notation, write D^k_{ij} as the total demand for segment k between UK origin i and international destination j, and $p^k_{A|ij}$ is the proportion allocated to airport A. Let $f_A{}^K{}_J(D_A{}^K{}_J)$ be the Larame curve appropriate to the relevant route at airport A, depending both on the segment K and the international destinations J (where K allows for some summation over segments k, and J correspondingly over destinations j). This converts the passengers at A to ATMs. The relevant number of passengers can be calculated as $D^K_{AJ} = \sum_{k \in K} \sum_{j \in J} \sum_i D^k_{ij} . p^k_{A|ij}$.

For each Larame curve, the number of ATMs at A is thus given by:

$$ATM^K_{AJ} = f^K_{AJ}\left(D^K_{AJ}\right)$$

and hence the total ATMs at airport A is obtained by summing over all the (relevant) Larame curve outputs (indices K, J). It is to this total quantity that the (runway) constraint applies.

Any shadow cost per ATM at airport A is assumed to apply equally to each ATM. Hence, for those ATMs governed by Larame curve $f_A{}^K{}_J$, the average passengers per ATM is given as D^K_{AJ} / ATM^K_{AJ} , and the shadow cost is distributed per passenger using this ratio.

In the light of all the foregoing, it should be clear that, while the model may be able to achieve the required constraints, the mechanism for doing so needs to be carefully considered, in order to avoid arbitrary assumptions (with consequently arbitrary outcomes in terms of passenger allocations).

Note that it is not being implied that the current procedure is inappropriate. The aim of this discussion is to emphasise how much room there is for judgment within the procedure, and to ensure that the details of the mechanism are open to scrutiny and debate.

4.3 The current position

The foregoing discussion makes it clear that there are a number of options to consider. In the first place, it is necessary to select which generalised cost component should be identified with the shadow cost.

It has been proposed that the most "neutral" component to select is the surface access variable, and this is the approach taken in the current implementation (except for ItoI). Despite some reservations about the form of the SA variable, I support this proposal. In its defence it may be said, firstly, that in the absence of a robust fare coefficient it is in fact the only variable which actually contains a true money component (as opposed to the variables which are converted to monetary units by means of the values of time etc.). Secondly, the airport allocation model estimation for the various demand segments comes up with a consistent pattern that the leisure SA coefficients are approximately 3 times the business coefficients, indicating a substantially higher money sensitivity for leisure, as would be expected. Hence, while recognising that other judgments could be made, this is certainly defensible.

With the (obvious) exception of the ItoI segment, the surface access variable is included in all the currently recommended models (for the Domestic segments, which have the "general" form, this is also true, though of less significance). Hence, within the current approach, the only question for implementation is the ItoI segment, where the currently preferred model has only IFT and Frequency, and the estimated coefficients are different by a factor of 2. Thus an arbitrary decision is required in this case, and the Frequency variable has been selected.

For the purpose of achieving the constraints (ie the outcome model demand), the impacts of alternative forms of implementation (i.e. using other variables such as frequency) will probably not be great (though alternative assumptions about how to allocate the costs across passengers could be more significant). But it should be noted that the uncertainty about the scale of the shadow cost could have greater consequences when we treat it as an actual component in the appraisal/CBA. Given that the ratio of direct frequency to SA coefficients is, for some segments, as high as 2.5, we might require rather more assurance that we were making the right choice of variable, or, at the very least, carry out appropriate sensitivity tests.

4.4 Conclusion on Capacity Restraint Procedure

i the "demand suppression" procedure is rightly based on the composite utility across all airports. However, more support is required for the controlling elasticities, and the functional form of the elasticity-based adjustment needs some attention. It appears necessary to check the code to confirm precisely what is being done at present, and to decide how any change should be implemented.

ii the detail of the mechanism for converting a uniform ATM shadow cost (in the case of runway constraints) to a per passenger cost by means of the Larame graphs needs more description. Presumably it is in some way allocated *pro rata* to the number of passengers allocated to each Larame graph, but no details have been seen. However, it will have the general effect of concentrating the allocation of the shadow price to routes operating smaller planes.

iii it is assumed that the estimated shadow cost per passenger applies equally to all passenger segments in monetary units of generalised cost. While this could be argued to be a neutral position, it is not necessarily equivalent to how airlines/airports would "price off" additional demand in practice. Sensitivity tests are recommended in terms of reasonable alternative assumptions about the **distribution** of shadow costs.

iv faced with the fact that different coefficients have been estimated for the various components of generalised cost (despite them being in principle all in monetary units), it has been decided to treat the shadow cost per passenger as if it is in "surface access" units. While this is considered reasonable, it is important to understand the potential sensitivity to making alternative assumptions. In addition, since the surface access variable is not included (for good reasons) in the ItoI choice model, sensitivity tests need to be carried out for alternative assumptions about the treatment of shadow costs for this segment

v in all cases where sensitivity tests are recommended, the impacts on CBA need to be examined as well as those on demand.

Finally, it should be noted that no review of the **algorithm** for achieving capacity restraint has been carried out. However, even from an informal reading it seems highly likely that significant improvements could be made here. In particular, there are remarks made (eg para H.41f in "Rules and Modelling") which indicate that the algorithm is too sensitive in those cases where the capacity constraint switches from being runway-based to terminal-based. At least in the medium-term, the potential for re-engineering the algorithm, making use of the advanced optimisation methods now available, should be investigated.

5 Overall Conclusions and Recommendations

There has been limited time to review what by any standards is a complex model, and therefore the conclusions need to be read in the light of this qualification. Nonetheless, I believe that I have been able to form a general assessment of the overall model, without necessarily appreciating every detail.

In this respect, I would like to make a strong plea for more documentation. It appears that in 2004 there was a concerted effort, and the resulting “Rules and Modelling” document was a major step forward. However, with the passage of 6 years, not all that is reported there is necessarily up to date, and in any case there were items that were rather cursorily described. The complexity of the model makes it essential that the documentation is kept up to date, and in an appropriate form. These remarks should also be taken to refer to the program code, which I suspect would benefit from a major overhaul before the details become impenetrable.

I have also taken more or less as fixed the demand segmentation, though it may be noted that for short-haul leisure travel, the *a priori* reasons for distinguishing Scheduled, NFC and charter are not compelling.

My overall assessment is that the model in its current form is broadly fit for purpose, but that some of its use for policy testing should be subject to caution. The development of the model has proceeded in a competent way, and while it does not always have full academic rigour, I do not see any major issues in that respect which would lead to fallacious outputs. I have also been impressed by the understanding of the consultants who are responsible for the model.

In making specific recommendations, I have had to bear in mind that not all of what I have termed short-term changes can necessarily be immediately implemented, because of the constraints of modifying the C++ code. I understand that my recommendation for removing fare from the airport choice models has indeed been carried out. The only other change which I think should definitely be implemented as soon as possible relates to the Domestic model, where a nested structure is required to reflect the choice between air and surface transport for this segment. The current model will both underestimate the amount of substitutability between airports and overestimate the substitutability between air and surface modes.

In the slightly longer term, I have a number of observations. In setting these out, I distinguish between the three main model components.

Airport choice

In documenting the estimated models, I recommend that both the coefficients **and** the estimated t-ratios should be given: the t-ratios are the simplest way of conveying the relative strength of the effect of the variable.

Although only two of the 19 “standard” models have been examined in any detail, the conclusions appear to be reasonably robust. The current choice data set is the best available, and the general approach to estimation is appropriate.

The conclusions for the model in the short term imply that

- there will be difficulties in testing fares policy as applied to specific airports or routes (of course, **global** fares policies can still impact (outside NAPALM) on overall air passenger demand);
- given some concerns about the specification of the surface access variable, the model’s prediction of the impact of specific **modal** changes in particular corridors on airport choice should be treated with caution.

Turning to the longer term, the following specific points are identified as worthy of investigation:

re-formulation of the surface access variable, in particular bringing the λ parameter more into line with other studies of airport access, allowing for variation between business and leisure purposes, and investigating how far the current composite variable reflects the balance between the highway and rail costs;

re-consideration of the most appropriate way to incorporate frequency into the model, and relaxing the general “waiting time” interpretation (this applies to the ItoI model as well);

considering whether some of the variables might be better included in logarithmic form, as has been found in other studies;

giving some consideration as to whether some “nesting” of alternatives might be appropriate (in particularly between direct and indirect airport choices for the scheduled market);

considering operating the model on an incremental basis

consideration of the most appropriate weighting to deal with estimation using “choice-based” samples.

Moving to an incremental basis would involve establishing a base year forecast in its own right. However, this brings with it a number of advantages, in particular the possibility to re-introduce (changes in) fare to the model. Since the surface access variable includes an explicit money component (the cost of accessing a particular airport), and it seems reasonable to consider that in choosing an airport, passengers could consider a direct tradeoff between fare and surface access cost, there is an *a priori* expectation that the coefficients on SA and fare should be the same.

Air Traffic Movements

More documentation of the Larame graphs would be welcome, and in particular a justification of the number of different graphs, and the way they are allocated to particular routes.

Given that the load factors are an essential element in the conversion between passengers and ATMs, some support for the values should be provided: if there is significant uncertainty about some of them, then sensitivity tests should be carried out.

Capacity Constraint

More support is required for the controlling elasticities affecting "demand suppression", and the functional form of the elasticity-based adjustment needs some attention. The code should be checked to confirm what is being done, and to decide how change should be implemented.

The detail of the mechanism for converting a uniform ATM shadow cost (in the case of runway constraints) to a per passenger cost by means of the Larame graphs needs more description. Presumably it is in some way allocated *pro rata* to the number of passengers allocated to each Larame graph, but no details have been seen.

The model assumes that the estimated shadow cost per passenger applies equally to all passenger segments in monetary units of generalised cost. While this could be argued to be a neutral position, it is not necessarily equivalent to how airlines/airports would "price off" additional demand in practice. Sensitivity tests are recommended in terms of reasonable alternative assumptions about the **distribution** of shadow costs.

Given that different coefficients have been estimated for the various components of generalised cost, it has been decided to treat the shadow cost per passenger as if it is in "surface access" units. This is considered reasonable, but it is important to understand the potential sensitivity to making alternative assumptions. In addition, sensitivity tests need to be carried out for alternative assumptions about the treatment of shadow costs for the ItoI choice model.

In all cases where sensitivity tests are recommended, the impacts on CBA need to be examined as well as those on demand.

Finally, while no review of the **algorithm** for achieving capacity restraint has been carried out, it seems highly likely that significant improvements could be made here. At least in the medium-term, the potential for re-engineering the algorithm, making use of the advanced optimisation methods now available, should be investigated.