Department for Transport

Response to the Peer Review of NAPALM

August 2011

The Department for Transport has actively considered the needs of blind and partially sighted people in accessing this document. The text will be made available in full on the Department's website in accordance with the W3C's Web Content Accessibility Guidelines. The text may be freely downloaded and translated by individuals or organisations for conversion into other accessible formats. If you have other needs in this regard please contact the Department.

Department for Transport
Great Minster House
76 Marsham Street
London SW1P 4DR
Telephone 0300 330 3000
Website www.dft.gov.uk

Contents

1. INTRODUCTION

1.1 In July 2010 the Department for Transport commissioned John Bates, a recognised expert in the field of transport modelling, to peer review the National Air Passenger Allocation Model (NAPALM). In October 2010, we received the final report[1] setting out the key findings of the review, including recommendations for the future development of the model. In this document, we have attempted to respond to the recommendations made in the final report.

1.2 Firstly we should record our opinion that the Peer Review has been a most valuable process that has provided many stimulating insights and been carried out with considerable thoroughness and rigour. The scope of the commission and its timescale were extremely challenging, especially given that there was the accumulation of ten years of technical development to take on board. We also appreciate that the Peer Reviewer went into considerable technical detail in examining datasets and modelling techniques beyond that recorded in the final document. This thoroughness adds considerably to our confidence in the Review's findings.

1.3 The report acknowledges the long history of the modelling work and the need to keep the review "within the spirit of the existing model". It therefore adopts a practical approach of:

a. investigating whether the model can be described as fit for purpose; and
b. suggesting somewhat more radical departures which could be followed in a less constrained timescale.[2]

1.4 In looking at the recommendations made by the Review it is important to keep in mind the distinction in (b) above. The Review contains several technical recommendations that are extremely helpful pointers for the long term development of the passenger allocation model. These represent significant departures from the historic form of the model, but do not in themselves affect the conclusions to (a), which found that "the model in its current form is broadly fit for purpose"[3] and therefore suitable for its current purposes in undertaking work programmed to be

[1] Peer Review of NAPALM, John Bates Services, October 2010
[2] Peer Review of NAPALM, John Bates Services, October 2010, p1.
[3] Ibid. p.44

completed in short timescales. This response will attempt to make this distinction as it affects specific items.

1.5 The Review has been organised into three main areas:

- The passenger airport choice (logit) model within NAPALM.
- The passenger to air traffic movements (ATM) modelling.
- The application of capacity constraints and shadow costs.

1.6 Some additional points are also included in the conclusion. We will address the technical issues raised following this structure, concentrating primarily, but not exclusively, on the matters included in the Peer Review's final section on Overall Conclusions and Recommendations. Given the balance of the Peer Review Report much of this response will necessarily deal with (1) the passenger airport choice model.

2. THE AIRPORT CHOICE MODEL

Presentation of logit choice models

The Review recommends (pages 25 and 44) 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."

2.1 The spreadsheet that supports the logit modelling documentation[4], Note10_014AnnexA v1.3.xls, includes adjusted rho squares for all overall models and coefficients, standard errors and p-values for individual variables. The p-value gives an indication of the confidence in each variable and the t-ratio can be calculated from the information provided as estimated coefficient/standard error. However, an updated version of the Note Annex A[5] including t-ratios has been provided.

Summary Response: This is accepted and the technical documentation has been updated.

Fares and airport choice

The Review notes (pages 25/26) that, although "some of the … differences in fare seem large enough to influence airport choice … the [model] estimation has not been able to identify this … despite commendable efforts …".

2.2 We agree with the Reviewer that data problems seem to prevent any possibility of resolving this in the short term. As a consequence fare is now not used as a choice variable in any of the models implemented in the new version of the model. Implementation has involved a significant

[4] The document referred to here and other technical documentation mentioned in this response are available from the Department on request. This includes Rules and Modelling, A User Guide to SPASM, Edition 2 DLL25, 2004, a large document that is the main technical guide to the aviation passenger allocation model. The Department's general transport appraisal guidance (WebTAG) includes a chapter on logit modelling, which underlies most transport demand modelling including that in NAPALM, at http://www.dft.gov.uk/webtag/documents/expert/unit3.10.3c.php.

[5] Note10_014AnnexA v1.4.xls

revisit of the recalibration/revalidation exercise. The re-estimation of the models without fares and the subsequent re-validation of the allocation actually showed very little change to the key airport choice parameters or the quality of the passenger allocation. This confirmed the judgment of the Peer Review about the weakness of the fare variable.

2.3 In general we would note that most often there simply was not enough variation in the fare data between airports from which passengers were potentially choosing. We would also note that:

a. All possible fare data, both CAA and IPS, was extensively analysed. Although ultimately rejected for inclusion in logit models, the data itself and the fare models developed still now provide a useful and enhanced resource for the modelling framework that was not previously available to DfT.[6]
b. Some of the Reviewer's doubts (page 11) about the quality of CAA fares data (p.11) are perhaps unduly harsh despite the footnoted qualification following a response from the CAA. Much of the modelled fare data was sourced from the International Passenger Survey (IPS). The IPS only collects fares in units of single UK£ fares and is carefully filtered, so ought to be regarded as a reliable fare source. We found that once the CAA fares are all converted to single UK£ fares there is a close correspondence with IPS.

2.4 In the long term we might still like to include fare as a choice parameter. Therefore we are interested in the Review's suggestion (page 45) that this might be achievable if the model was switched to incremental form. However, our view is that this might be the main advantage of switching to an incremental model form and there would be other difficulties with this course (see paragraphs 2.15 to 2.17 below).

Summary Response: This proposal was accepted and implemented with a detailed model revalidation exercise undertaken, excluding fares.

[6]E.g. modelled fares may still be used inside for generating composite generalised costs of air travel between o-d pairs for the purpose of demand suppression

Overall concern about surface access variable

The Review (pages 9 and 10) notes that NAPALM currently handles the surface access data "as if it were a mode choice model nested below the main (airport) choice model." It concludes that while "the general methodology is judged acceptable it could be challenged on many grounds." It notes in particular that "the lambda value[7] of -0.1 is high in comparison to the results of other airport mode choice estimated models …; …there is a case for allowing [lambda] to vary, at least between business and leisure segments; …. The VoTs [values of time] used are based on a number of assumptions which could be contested; [and, as] car is a dominant mode … combining car and rail without allowance for mode constants will tend to overestimate the impact of rail cost changes."

2.5 The Review's main reservation was that the surface access lambda (λ, taken as -0.1 for all modes) used to derive a composite (road + rail) generalised cost of airport surface access had not been estimated from UK airport data; it also potentially questions some (e.g. UK resident international scheduled and NFC leisure markets) model hierarchies.

2.6 In a nested choice model the most sensitive choice with the largest sensitivity parameter (λ) should be at the bottom of the nesting hierarchy and the least sensitive choice at the top. In practice NAPALM does not at present use a nested model. It is nonetheless important to consider the surface access sensitivity parameters which are used in the individual purpose airport choice models in deriving the common sensitivity used to create the composite surface access costs. If an airport choice surface access coefficient is greater than the value of -0.1 chosen for the composite cost (as happens in the leisure models), it implies that theoretically the allocation model would have been better specified with choice of surface access mode at the top of a nested hierarchy.

2.7 However the purpose of this parameter in almost all uses of NAPALM has been to produce a superior alternative to a simple weighted average of highway + railway district to airport surface access cost. Arithmetically averaged costs between the access modes are not appropriate, whereas using composite cost and a choice sensitivity parameter (-λ) means that

[7]Lambda is the "logit sensitivity parameter" in the equation defining the proportion of travellers choosing one alternative out of a range of possibilties, as set out in paragraphs 1.3.5 and 1.3.6 of the WebTAG chapter noted in the footnote to paragraph 1.2 above. Lambda is also variously described as 'choice parameter', 'mode share parameter', or, in this context, 'surface access coefficient." The term 'scaling factor', used to describe a (positive) factor that reflects the ratios of the lambdas for different response mechanisms as you move up a nested model structure, is also sometimes used to describe lambda itself. While the exponents in the logit equation are of course always negative, lambda itself is sometimes presented as positive (like negative price elasticities of demand) and sometimes as negative.

a mode choice with very high cost will contribute relatively little to the overall average access cost of the journey.

2.8 As a measure of separation between each district and each airport, we regard the current methodology and parameter as being adequate, but if the input surface access models were used to assess the effect of (say) new modes of access on airport choice (e.g. new rail links to airports) then this form of composite cost, and indeed potentially the model nesting structure, is less adequate.

2.9 In recent years no model runs have used the surface access costs for any purpose other than a weighted measure of separation between each ground origin district and each airport. However, if the model is to be used in future (for example) to assess the impact of a new mode such as High Speed Rail (HSR) on airport access, then some caution is required as the adopted sensitivity parameter (-λ) could overstate the HSR effect on airport accessibility. It is our view that this would have only a very small effect on the overall airport allocation, but that any more detailed assessment of HSR should be avoided until the changes recommended by the Peer Reviewer are implemented.

2.10 At present the calculation of composite surface access cost takes place within the NAPALM C++ code. There would be considerable advantages in terms of flexibility of model specification if this functionality was externalised with composite costs being calculated as an off-model input. This would also allow the estimation of composite costs to be recalibrated to allow the use of mode specific constants as mentioned in the last of the bullets straddling the Peer Review pages 9 and 10.

Summary Response: In the near term and for the current programme of model runs, these issues are not expected to have any significant effect on the model results. However, we accept that there is a case for change and in the longer term aim to introduce a nested hierarchical domestic model and the facility to use mode specific constants. We will also increase flexibility and transparency by making the composite calculations outside the main NAPALM software. These improvements will in the longer term make the model a more robust tool with which to forecast the interaction between air transport and changes to surface modes. In the meantime, we will avoid using the model to perform more detailed assessments of HSR.

Reformulation of the surface access variable

The Review (page 26), in “turning to the longer term”, identifies reformulation of the surface access variable as a point worthy of investigation, “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.”

2.11 Bringing the surface access λ parameter in line with other studies of airport access will undoubtedly enhance the longer term model and would be desirable for compatibility with related models. However we note that:

a. we do not wish to confuse this annualised passenger to airport allocation and ATM forecasting model with the functionality of specific daily passenger to airport mode share models such as BAA's LASAM.[8] Such models need to represent daily peak periods and a considerably wider range of access choices such as taxi, car park, kiss and fly, charter coach etc. At present the primary function of NAPALM is as a provider of purpose split annualised surface access demand inputs to such models.
b. the logical conclusion supported by some of the estimation results is that ultimately some form of simultaneous mode and airport choice process might be preferable at least in some markets. However as the Review acknowledges this would be "a substantial task … not compatible with the current timetable". It should only be considered in the longer term. To our knowledge such model forms are ambitious and have only been implemented in the UK in academic studies of limited airport choice and with mixed results.[9]
c. Values of λ other than -0.1 have been calibrated and tested using recent airport surface access data. They are closer to 0 than -0.1 and this implies that composite costs would be lower than at present. The impact of this would not be very noticeable where the generalised costs between different modes are not similar and the composite cost would still be close to the least cost mode. The presence of more costly alternatives would have relatively little effect. However, if the generalised costs of alternative modes were similar, there would be a lower composite cost. It should be noted that our calibration of alternative λ values failed statistical confidence tests, so introduction of additional λ values would require further detailed work.
d. We did make test calculations of new airport access composite costs using the alternative λ values (despite the reservations on the quality of their calibration described above). The change they produced in the composite costs of airport access was modest compared to the changes in airport accessibility costs resulting from the change from NAAM to LDM as the source of airport access costs. The adoption of LDM composite costs when tested in NAPALM produced little change in the pattern of airport allocation compared to using the NAAM composite costs (as reported in a technical note[10]). On this evidence we think it quite reasonable to assume that the alternative λ values would have had even less effect in patterns of airport allocation..

2.12 A recommendation that we would like to implement as soon as convenient is to introduce a nested logit model into the domestic end-to-end market of the type discussed and illustrated in Section 2.8 of the Peer Review. Choice between airports has been found in the Review's

[8] LASAM: London Airports Surface Access Model.
[9] Ibid.
[10] Technical Note D0NFLSB/10/017 v1, June 16, 2010 LDM vs NAAM Comparison Notes

estimation to be more sensitive than the choice between surface and air travel on domestic journeys. We agree that having a single nest multi-nomial logit form for all choices, as at present, potentially overstates the sensitivity to surface mode changes such as HSR. There is significant change required to NAPALM C++ coding required to implement this change and there was, therefore, no scope to make these changes in time for the current round of model runs. This does not in our view invalidate including HSR in future year networks, but we agree that more detailed tests assessments of the impact of HSR should not be undertaken until the changes recommended have been implemented.

Summary Response: These issues do not preclude including HSR in future networks. However, in the longer term as part of the next phase of model development, we agree that there is a strong case for allowing the introduction of newly calibrated surface access λ parameters in the composite cost calculations and a hierarchical nested model for the domestic market. Until these changes have been implemented any NAPALM forecasts of the effect of HSR should be treated with caution and supplemented with evidence from external sources where appropriate.

Reconsideration of the frequency variable

The Review refers to (pages 13 to 15) "the conventional treatment of urban public transport", where it is assumed that people arrive at the stop or station at random and wait for the next departure, and that those assumptions "are much less appropriate in the case of long distance transport and air travel in particular". It comments that "the modellers' approach is generally in line with [a more realistic view of air travel] but … more influenced by the "half headway" [urban transport] rule than is really appropriate."

2.13 The modelling of the frequency parameter has been examined in some depth both in the Review and by the DfT itself during calibration. The argument that calculating standard public transport waiting times may be inappropriate as a frequency measure for air travel is broadly supported. However, it should also be emphasised that the current functional form with 'little a' parameter somewhat differentiates the frequency calculations from the urban transport rule and attempts to reflect the greater need of air passengers to tailor their arrivals to the timetable.

2.14 Simple frequencies by route as a measure of levels of service were tried in the Review and found to produce similar models and similarly sized parameters as the present frequency terms. So such a technique is attractive for future development, but there is no immediate case for change. The main advantage of using a route level of service variable is that it is simpler, easier to explain to non-experts and deals better with very low frequency and seasonal services. Given these benefits, we will

aim to incorporate this change when the logit models are next re-estimated.

2.15 A "wait time" based frequency term and a level of service based frequency term both produce similar models in terms of the sign and relative scales of other coefficients and in terms of the allocation of passengers. Therefore, although the Peer Reviewer's suggestions will improve the model, we have decided to retain the "wait time" frequency for the present model generation. Numerous tests have been conducted on the 'little a' parameter which controls the length of wait times. These confirmed the leisure value of 0.4 was robust but showed that some business markets such as long haul could have a changed parameter of 0.3 (from 0.2). This probably reflects increased security delays. However, airport allocation sensitivity tests showed that the changes to forecasts would be minimal and the current settings are retained for now.

Summary Response: The form of the frequency function has been extensively sensitivity tested during the Review and found to be adequate for current forecasts. There are persuasive arguments for introducing a simpler level of service variable which will not significantly change the forecasts but will be more easily understood by a wider audience and will deal better with very low frequency and seasonal services. This will be taken forward in the longer term, when the models are next re-estimated.

Estimating variables in logs

The Review suggests (pages 26 and 45) that, as one of several points "worthy of investigation [in] the longer term", it would be worth "considering whether some the variables might be better included in logarithmic form, as has been found in other studies."

2.16 We have never tested this, so it is difficult to comment. We agree with the Peer Reviewer that this is unlikely to have a major impact on the models. However we agree that it should be tested at the next opportunity when the logit models are re-estimated.

Summary Response: There is unlikely to be a significant improvement on the currently estimated models, but as it will not require us to incur significant costs, we agree that it is worth investigating in the longer term when the models are next re-estimated.

Nesting of direct and indirect routeing choices

The Review suggests (pages 27 and 45), as another issue for the longer term, "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)"

2.17 This is a helpful suggestion made after the completion of the estimation process. Introducing separate choice nests for direct and indirect routes prior to the airport choice nests may well potentially improve some of the most difficult to estimate models. No tests have been conducted in the Review or by DfT to confirm the potential of this approach. It would require significant re-coding of the NAPALM software.

Summary Response: The Review recommends this as a longer term test, and, given that implementation of a nested structure also requires significant changes to the C++ program coding within NAPALM, we support the view that this approach should be considered in the next estimation process. However, given the significant costs involved, we would only implement changes to the nesting if we were confident that the change would significantly improve the models' performance.

Incremental model structure

The Review suggests (pages 27 and 45), as further longer term issue, "considering the operation of the model on an incremental basis" and discusses this in more detail on pages 27 and 28.

2.18 The introduction of an incremental model whereby the estimated models are used to 'pivot' around observed airport choices was suggested by the Peer Reviewer during his independent testing where sample fitted Review models produced moderately successful passenger allocations. An incremental model form was initially raised primarily as a means of improving the quality of the passenger to airport allocation[11]. A useful by-product of the technique is that it can be used to assess the effect of changes in choice variables and might potentially be used to better reflect the impact of a future change in individual airport fares.

2.19 The technique requires observed data of ultimate origin and destination with choice of airport. Such origin-airport-destination data is generally available in the set of CAA data used by the DfT.

[11] Bates, p.23

2.20 However there are a number of reasons why this approach is rejected in both the short and medium term:

a. The principal motive for such an incremental model form in relation to the alternative 'absolute' model is to improve the standard of the model performance and in particular its validation against observed data. However, as the model validation exercise has demonstrated, the absolute model form has produced a high standard of model validation.[12] The current model validation is superior to that reported in UK Air Passenger Demand and CO_2 forecasts, 2009, and no criticism was made of the model performance reported in that document. The calibration at the individual route level in NAPALM does allow for the introduction of generalised cost adjusters which can significantly improve model fit at the micro level in the final rounds of calibration. The Peer Reviewer only became aware of these adjusters at the end of the process after seeing model validation documentation and has commented that this additional calibration tool somewhat weakens the case for an incremental model.[13]
b. An incremental model would need to pivot on observed data at the individual route level. However, this poses major problems as the input network in the form of the route structure changes dynamically as the model forecasts year on year. Specifically it is not possible to pivot changes in passenger numbers on routes which have either been cancelled or started up by the NAPALM route viability algorithms as the model rolls forward year on year.
c. NAPALM retains functionality to model new airports. It is difficult to see how new airports can be represented with an incremental model form.
d. Not all airports are in fact covered by detailed CAA ultimate origin-airport-destination observations. Demand using Newquay, Plymouth, Blackpool and Coventry has never been captured in CAA surveys and has to be synthesised in the base inputs. CAA demand observations at Southampton and Norwich are too old to be directly usable and also require synthesis.

Summary Response: Implementing this recommendation would require us to incur significant costs and is associated with a number of problems. Therefore, as it is not clear that adopting the recommendation will significantly improve the already high performance of the models in comparison with independent base year data, we currently do not plan to move to an incremental model structure.

[12] See the technical notes on Validation, Note10_018 Napalm Validation Update_v4.doc
[13] See p.24.

“Choice-based” samples

As a final issue for the longer term on the Airport Choice Model the Review suggests (pages 27 and 45) “consideration of the most appropriate weighting to deal with estimation using “choice-based” samples.”

2.21 There is an argument that the datasets sampled in the Review contained many observations where passengers in practice had little choice and that a different model might be formulated from a more limited sample focused on marginal choices between airports (e.g. the scheduled US model).

2.22 In the discussion on choice data (Section 2.4) the Review laments the lack of real choice for most passengers in the US dataset because of the relatively small number of airports offering significant services. However, separate logit models have to be estimated for short haul, US and other long haul international markets. We considered it worthwhile that one of the two main sample datasets for the Peer Review should be a market other than short haul and also a market where hub/transfer traffic was significant. [The sample short haul market was UK resident NFC leisure, which was proposed because it promised a high degree of airport choice but was simpler with little transfer traffic – this market was analysed, more briefly than the long haul data, in Section 2.6 of the Review]. In practice the US market had the highest available choice of all the non short haul markets. The 10% proportion of transfer traffic was representative of the overall national proportion. Nonetheless we do agree that there is merit in concentrating on data where airport choice is more finely balanced. Achieving a sufficient sample size could require additional survey work.

2.23 It is unlikely that customised new surveys for such choice data will be possible in the near future. Nonetheless there is scope to use the wealth of CAA and IPS revealed choice data held by the DfT to change the selection of datasets to focus on demand from districts with marginal choices in future estimations. This would considerably shrink the size of the estimation datasets which could prevent the estimation of robust models. But it is nonetheless considered a worthwhile methodology to test in future re-estimations.

Summary Response: The choice of datasets sampled does reflect the structure and geographical definition of the current models. However, in the longer term, when we next re-estimate the models, we will attempt to examine observations of more marginal choices if enough data can be assembled. A more radical alternative would be stated preference surveys, but this would be a significantly more expensive course and unlikely to offer sufficiently high value for money.

Additional points not included in the Peer Review's Overall Conclusions

Alternative specific constants (ASCs)

The Review notes (page 16) that "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." It makes no recommendation.

2.24 Theoretically adding a constant into the logit models will improve baseline performance, but as the Review states, if too large such constants run the risk of dominating the forecasting. We have run versions of the model with and without ASCs and agree their effect on model performance is small. Therefore we conclude that there is no strong reason to introduce constants into the logit formulation.

Summary: This was tested during estimation and found to have only a minor effect.

Investigation of residuals

In its examination of the scheduled UK business market to the US the Review comments (page 17) that "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 …. It is recommended that this is further investigated".

2.25 We accept that such a search for patterns of residual is sound advice and would expect to incorporate such investigations as standard procedure during the next model re-estimation exercise. However, we also note that the process of calibrating particular routes through the

addition of generalised cost “fare adjusters” also serves to highlight and permanently record outliers.

Summary: This does not affect current forecasts but good practice for future model re-estimation.

3. ATM MODELLING

Larame graph documentation

In its discussion of Larame graphs the Review concludes (page 34) that "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."

3.1 The Peer Review analysed and described all 'Larame' passenger to aircraft size inputs. A total of 244 graphs were examined. There were no technical concerns about the process. However, a request was made for more supporting technical documentation on the definition of the Larame graphs, particularly of their supporting statistical data. This request is acknowledged and supporting analysis will be provided in the next round of technical documentation.

3.2 In addition to the types of differentiation listed in the Review, inherited from the CAA/NATS SPAM model and its documentation, we feel it is worth noting that the recent expansion of Larame graphs has been driven by an appreciation of the value of differentiation by airline in providing a more accurate translation between route level passenger demand and aircraft sizes and types likely to be used to meet that demand. Such airline differentiation has been drawn from analysis of detailed CAA route level statistics.

Summary Response: A research paper describing the underlying statistical analysis of the Larame graphs will be included in the updated technical documentation that is due to be produced later in 2011.

Load factors

The review comments (page 3) that "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."

3.3 Selection of load factors is based on analysis of route level CAA statistics. The detailed model validation exercise has included a report of

modelled versus actuals of passenger loads on each route by airline type.[14] This provides clear and objective confirmation of the accuracy of the load factor selection process.

3.4 We also accept an implicit criticism that the programming mechanics were somewhat clumsy in that different aircraft size graphs were often defined where the only change was load factor. This does not affect model performance, but is clearly a potential programming enhancement that can be investigated at the next stage of software development.

Summary Response: Concerns about base year load factor selection have been addressed by detailed reporting of aircraft loads in the model validation.

[14] The technical note on Validation is Note10_018 Napalm Validation Update_v4.doc

4. CAPACITY CONSTRAINTS

Elasticities controlling demand suppression and their application

The Review comments (page 42) that " 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."

4.1 The Review has confirmed the soundness of the principle of calculating the composite utility across all airport and routeing choices of travelling between particular origins and destinations. However it also calls for more justification of the controlling elasticities used and further investigation of the programming code that applies these elasticities. There is also some questioning (page 37) of using short haul elasticities for domestic travel.

4.2 The controlling elasticities apply to generalised costs. These generalised costs are expressed in units of "surface access £s", which the Review considers "reasonable". At present we have deemed it is most suitable to change the approach so that we now apply current fare elasticities calibrated from the econometrics to the NAPALM passenger purpose markets. It is accepted that generalised cost and fare elasticities are not exactly the same, but at present we only have clear evidence of fare elasticities.

4.3 Since the Review, new and different fare elasticities have been applied to the domestic market and have been adopted for demand suppression. The business elasticity (-0.3) now differs from the international value, but the leisure value remains similar (-0.7). The domestic suppression elasticity does not necessarily have to include switching to surface modes as this effect is already captured in the domestic multi-nomial logit model.

4.4 It is worth noting that sensitivity tests have been undertaken and reported on varying the elasticity values.[15] These tests concluded that the effect of suppression elasticities is relatively small and declines when more capacity is provided or the level of input demand forecasts drops. However, the test report[16] also concluded that the λ parameter used to calculate composite utility of travel for an o-d pair could be an equally important factor. As evidence of revealed demand suppression is not easily obtained, we have not yet been able to calibrate a new parameter, however we accept that further investigation in this area would be worthwhile. Some tests on differing values for the domestic elasticity might also be worthwhile, although more robust justification would be required for switching away from the econometric elasticities. In terms of overall passenger forecasts the effects are likely to be modest.

4.5 The C++ program code which applies the suppression elasticities has been extracted and analysed for international and I to I markets.. We note that an improved methodology that requires the calibration of a θ parameter for applying a suppression elasticity is suggested in the Review (p.36). The Review does not expect its adoption to significantly change model results; however, we may attempt its introduction at the next software development stage because it represents "standard" methodology.

Summary Response: Controlling elasticities have been changed from the historic values to fare elasticities derived from the latest round of econometric modelling. A note on the effect of the values was produced[17]. The relevant program code has been identified and its investigation is continuing. Any changes to results from the possible alternative methodology are not expected to be significant.

[15] Note D0NFLSB/10/007 v1, March 23, 2010, The Impact of Demand Suppression Elasticities
[16] Ibid
[17] Technical Note D0NFLSB/10/007 The Impact of Demand Suppression Elasticities (March 2010)

More documentation on the conversion of ATM to passenger shadow costs

The review comments (page 43) that "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."

4.6 An ATM shadow cost is allocated at the route level to each ATM using a constrained runway. A single airport-wide shadow cost is converted to a per passenger shadow cost by dividing that overall runway cost by the modelled number of passengers on each route calculated from the aircraft size ('Larame') graphs as the model run progresses. This process is alluded to throughout the narrative in Section H of the earlier technical documentation Rules and Modelling, A Users Guide to SPASM, April 2004. This mechanism has not been made sufficiently clear and we will describe the process more thoroughly in the next round of technical documentation.

Summary Response: We are planning to include a full description of this aspect of the modelling in the updated technical documentation due later this year.

The equal application of shadow costs to all passenger segments

The Review notes (pages 43 and 46) that "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."

4.7 The contention in the Review is that rather than applying a uniform shadow cost to all ATMs or all passengers (if a terminal shadow cost), airports would in practice discriminate by, for example, differential slot pricing by time of day, or airlines by destination or ticket type. One option suggested is to allocate shadow costs in proportion to fares. The Review does not suggest that the "neutral" position of uniform ATM or terminal shadow costs is necessarily wrong. But it does assert that it is worth

gaining an understanding of the potential impact of alternative shadow cost regimes.

4.8 The Review has acknowledged that the more significant runway (ATM) shadow costs do in effect introduce price discrimination as they convert to a cost per passenger by dividing the runway shadow cost by the average aircraft loading for the route. So each ATM shadow cost has a different effect at the route level. Passengers on routes served by smaller aircraft (e.g. many business oriented routes) will face a higher charge per head and the shadow cost process will discriminate against short haul routes and in favour of long haul routes served by larger aircraft.

4.9 The issue is tied in with the question of the extent to which the fares faced by passengers reflect price discrimination between passenger segments on the part of the airline. If airlines perfectly price discriminate between passenger segments it is possible that the only passengers who would experience an increase in the fares they face as a result of a capacity constraint would be the most marginal passengers (with lowest willingness to pay) that yield the airline the least profit. In practice we find the effect of shadow costs in the passenger forecasting is to impact most heavily on the most marginal leisure passengers who are the most "footloose" or liable to be suppressed from travelling altogether.

4.10 We also recognise that if price discrimination in the application of shadow costs could be precisely modelled, it would have implications for the appraisal because it suggests that some consumer surplus may be captured by producers.

4.11 We accept that when a terminal shadow cost is applied, there is no effective price discrimination. With terminal shadow costs all airport users face the same premium and there is no discrimination by route. In practice, there is a presumption within the model that at the constrained airports, particularly those in the South East, runways will be the binding capacity constraint. There are a number of mechanisms currently enforced in the capacity constraint algorithm which ensure that where both the runway(s) and terminal are at or near capacity, the runway will be taken as the binding constraint unless the terminal is very significantly more overloaded.[18] We also go to some lengths in the capacity constraint procedure to prevent switching between ATM and terminal shadow costs as this leads to large swings in the levels of shadow cost faced by passengers between iterations or successive years and therefore problems with convergence.

[18] See Rules and Modelling, A User Guide to SPASM, 2004, paragraphs I.53-I.57.

4.12 Ultimately, we are not aware of any robust evidence on which to base alternative assumptions about how airlines (or airports) might seek to respond to excess demand in practice. We recognise, however, that in this and in other respects, the model is based on a stylised representation of reality. For example, the scope of the model is annual and does not include peak period modelling or, alternatively, there is no ticket type modelling (business/leisure purpose split does not proxy well for first/business and economy class passengers). A further complication is that the application of shadow costs is deeply embedded in the software and would require significant programming effort and cost to implement. Therefore for the current generation of model runs we will retain the current methods, placing emphasis on the importance of the model's route level aircraft size modelling capabilities. However, in publishing the results of the model we will recognise that this is an area in which judgement must necessarily be applied. In the longer term we will investigate whether we can find a cost effective method of sensitivity testing other ways of applying shadow costs.

Summary Response: Implementing this recommendation would require significant changes to the model code, at significant cost to the Department. Further, we believe the current 'neutral' assumption provides a reasonable indication of how airlines or airports might respond to excess demand and are not yet convinced that there are any alternative mechanisms for applying shadow costs that would be as robust. In the medium term we will discuss this with other aviation experts. If we identify alternative mechanisms for allocating shadow costs to passengers, we will investigate whether we can find a cost effective method for testing these in the model.

Flexibility in the monetary units represented by shadow costs

The Review notes (pages 43 and 46) that "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 I to I [interliner] choice model, sensitivity tests need to be carried out for alternative assumptions about the treatment of shadow costs for this segment."

4.13 The Review considers that the expression of shadow costs in units of "surface access £s" to be "reasonable", but raises the question of whether it will always be reasonable especially as some passengers such as I to I interliners make no surface access journeys in the model

and the sensitivities should be understood. We do not consider that there is an urgent need for change as no more acceptable pricing units are obviously available. However, we accept that it will be helpful if in due course for the program software has more flexibility to deal with different pricing units.

Summary Response: We currently have no plans to reflect this recommendation in changes to the model. However, we will review the case for introducing price unit flexibility at the same time as discriminatory shadow cost pricing in the longer term.

Review the algorithm for achieving capacity restraint

The Review notes (page 39) that the model structure "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. … 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. … there is a danger that incorporating an arbitrary choice into the solution algorithm could lead to an undesirable outcome." And later (page 43) "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 [indications] 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.

4.14 This algorithm is at the heart of the NAPALM program and consumes most computer processing time. We agree with the proposition that over time it has been subject to many piece-meal modifications and might benefit from root and branch "re-engineering" and optimisation in the medium term. Optimisation would be complicated by NAPALM having two possible and quite different constraints: runway and terminal. But such re-engineering might also free up processing capacity for more sophistication in other areas such as those discussed under the previous two headings and it might also assist in the increasing need for the model to run continuously into more distant future forecast years.

Summary Response: This suggestion is welcome, and we agree that there could be significant benefits to our reviewing the algorithm at the heart of NAPALM. Therefore, while implementing this recommendation is likely to be costly and time consuming, we would expect to prioritise this as part of longer term model development work, subject to being able to secure sufficient resources.

5. OTHER CONCLUDING POINTS

"a strong plea for more documentation"

The review makes (page 22) "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."

5.1 The last major technical documentation effort was Rules and Modelling: A Users Guide to SPASM, 2nd Edition produced in April 2004. Since then there has been regular documentation of various technical issues as they arise in a lengthy series of technical notes. A general reader's guide to the NAPALM program has also been provided in the 2007 and 2009 editions of UK Air Passenger Demand and CO_2 Forecasts. It is accepted that the technical documentation of not just NAPALM but also its associated models in the Air Passenger Demand & CO_2 Forecasting Appraisal System needs to be refreshed. We therefore propose to include provision for a full update of the technical documentation in the next (and imminent) contract specification.

Summary Response: The Department has commissioned updated documentation to be prepared later in 2011.

Scheduled / NFC / Charter Distinction

The Review comments (page 44) "that for short-haul leisure travel, the a priori reasons for distinguishing Scheduled, NFC and charter are not compelling."

5.2 The Peer Review does not extend to the National Air Passenger Demand Model (NAPDM) where the distinction has been dropped and replaced by

NFC and charter proportion sub-models immediately prior to the translation of the econometric demand forecasts into NAPALM. It should also be noted that the distinction has been withdrawn from the domestic end-to-end model which has been amalgamated into a single segment model. To some extent the "full service"/NFC distinction is a legacy of the model's development in 2000-2001 when NFC growth was booming and driving down fares at an unprecedentedly rapid rates. Subsequently because of the difficulties of making a precise distinction between two types of scheduled carrier and for consistency with NAPDM, we share the longer term aim of removing this distinction. However there are a number of reasons why the distinction should be retained in NAPALM international travel inputs and outputs for the current generation of the model:

a. Charter airlines clearly do operate to different business and operational models and this is reflected in all CAA reporting and elsewhere in the industry. The case for abolishing the distinction between charter and other airline types is far weaker than between the scheduled segments.
b. NFCs do not operate on a hub and spoke business model and are excluded from choosing indirect routeings via hubs in NAPALM. There has been some evidence of informal or DIY interlining (most notably 2004 to 2006) at Gatwick and Stansted. However the effect is small scale and restricting the choice of indirect routes via hubs to the "full service" scheduled airlines on balance produces a better choice model performance.
c. The passenger to ATM modelling currently achieves significantly greater accuracy through being able to apply separate aircraft size ("Larame") graphs to NFCs which have unique passenger demand to aircraft size relationships. This enhances the downstream fleet mix and CO_2 models.
d. The fleet turnover modelling also benefits from the distinction between "full service" and NFC airlines because of the unique patterns of fleet operation and aircraft purchasing policies of NFCs. It also lends greater accuracy to the fleet mix modelling and enhances downstream CO_2 modelling. This is relevant as the next programmed use of the model on MAC Curve modelling makes extensive use of the ATM, fleet mix and CO_2 modelling.

The scheduled, NFC, charter distinction is deeply embedded in nearly all the models in the framework and its removal within the current programme would be costly.

Summary Response: We share a long term aim to remove the scheduled/NFC/Charter distinction in NAPALM, and plan to review the case for doing this when we next come to re-estimate the models. However for the current generation of model runs, this segmentation is retained.