LeighFisher

PHASE 2 REPORT

EVIDENCE OF REVENUE GENERATION AND ABSTRACTION FROM HISTORICAL OPEN-ACCESS ENTRY AND EXPANSION

Prepared for

Office of Rail and Road

Date

7 January 2016

REDACTED

Leigh|Fisher

TABLE OF CONTENTS

LeighFisher

Leigh|Fisher

1 EXECUTIVE SUMMARY

1.1 Purpose

When open access rail operators operate new services, the revenue they receive comprises newly generated revenue (from rail journeys that would not otherwise have taken place) and revenue abstracted from existing operators. Comprising a review of the historical performance of the current open access operators (OAOs), the purpose of this study is to identify the proportions of revenue that have historically been abstractive or generative.

1.2 Review of previous work

A number of previous studies have been undertaken on the demand and revenue generation and abstraction effects brought about from open access operator competition. Of these, work undertaken by AECOM for Grand Central, and an MVA report into East Coast applications were of most relevance to our study.

The AECOM report reviewed past demand and revenue forecasts undertaken for Grand Central and proposed improvements to the methodology. It was unable to rectify the problem of over-predicting Grand Central demand at established stations and under-predicting demand at new stations.

The "ECML Track Access Applications Assessment" report, produced by MVA Consulting in 2009, assessed a number of track access applications on the East Coast Mainline. The approach to modelling differential fares from this report has been used as a basis for the approach adopted by our study; it has been developed further to better take into account different journey distances and ticket types.

1.3 Approach

In order to understand the historical generative and abstractive effects of OAOs, we have developed an approach centred on the creation of PDFH demand and revenue models, which retrospectively forecast revenue for the ICEC franchise, Hull Trains and Grand Central in a "with open access" scenario, and ICEC revenue in a "without open access" scenario.

The steps involved in the study are as follows:

1. Development of a revenue model, which retrospectively forecasts the effect of the open access operators on market size and market shares. This includes modelling the effects of TOC-specific fares;
2. Incorporation into that revenue model a counterfactual scenario, in which the open access operators do not exist, thus allowing calculation of an abstraction-generation split;
3. Comparison of the modelled results against outturn revenue, identifying a gap between modelled and outturn revenue;
4. Investigation of explanatory factors for the gap; and
5. Creation of a set of direct demand models, to capture railheading effects that are not captured by the Generalised Journey Time (GJT) elasticity based modelling framework.

In developing the revenue model, two main issues emerged:

- Use of MOIRA 2 to model fares effects was found not to be appropriate; and
- Gaps and inconsistencies in historical LENNON data necessitated some manual adjustments to the data.

1.4 Investigation of explanatory factors

In order to explain the gap between modelled and outturn revenue, a number of potential demand drivers were explored:

- On board offer;
- Rolling stock quality;
- Lagged effects of timetable and fares changes;
- Station quality;
- National Rail Passenger Survey results as an indicator of service quality; and
- Marketing.

In most cases, we have found no evidence to create additional demand overlays for service quality-related factors. In particular, we could not find evidence that Open Access TOCs increase passenger satisfaction beyond their passengers paying lower fares, travelling on less crowded trains and (among the NRPS sample) being less delayed.

However, we have identified a longer lag in GJT effects following the introduction of open access services than that recommended by PDFH. This was implemented in the revenue model.

1.5 Direct demand modelling

We have undertaken direct demand modelling in order to capture the impact of improved station accessibility in areas previously poorly served by rail. The results show demand switching away from stations served by incumbent franchised train services linking directly with London, to open access stations (i.e. to those stations benefitting from the introduction of an open access London train service), and generation of new demand in these areas. For some open access stations, this helps to explain the gap between actual revenue and that forecast by the revenue model. Similarly in the case of some incumbent franchised served stations such as Newcastle, it helps to explain why the revenue model forecasts demand above actuals.

1.6 Generation and abstraction identified

Based on the revenue modelling work, and prior to the addition of any modelled railheading effects, we have identified a bounded estimate of generated revenue. In 2014/15, between 14% and 28% of revenue was generated – between 16p and 39p for every £1 of revenue abstracted. The estimate is bounded due to variation of our modelled “with open access” scenario from outturn revenue.

Subsequent to the addition of railheading effects, **between 21% and 31% of 2014/15 revenue** is estimated to have been generated – **between 27p and 45p generated for every £1 of revenue abstracted**. Over the last fifteen years, the estimate is slightly lower – between 22p and 46p – because generation takes longer to develop than abstraction, and so there is greater abstraction while the operations mature.

As growth between [redacted] and London (a key driver of modelled generation) has been lower in reality than that modelled, **we consider that the actual level of generation has been toward the lower end of this range**.

Currently, the NPA test uses a ratio of 0.3, i.e. open access services pass the test if they generate 30p of revenue for every £1 abstracted. The NPA ratio of 0.3 is within the bounds of the generation estimates identified by this study.

We have also made a number of important findings regarding the modelling of the effects of open access operators:

- Modelling the effects of TOC-specific fares brings the modelled shares of revenue to each operator much closer to actuals. As such, fare strategy is a key driver of the generation-abstraction split.
- For some flows, ‘direct demand’ models were useful in explaining the gap between outturn and modelled revenue. We believe that calibration of a catchment area’s residents’ propensity to travel is a key area for improvement in these models. This would be helped greatly by access to current TOCs’ customer databases, which would be invaluable in modelling the effects of proposed services. Availability of this data would greatly aid future research, and it may be worth exploring with TOCs (both franchised and open access) their willingness to share such data for future research, in order to help the rail industry’s understanding.
- The levels of generation depend on the degree to which the Open Access TOC serves locations and flows currently poorly-served by rail. About half of the East Coast line’s Open Access TOCs’ revenues come from these flows, and our estimate of generation is in this context. For an Open Access TOC for which ‘railheading’ was not important, the share of generation would be lower.

2 PURPOSE OF STUDY

2.1 Study Aims

In its assessment of track access applications, as well as considering infrastructure constraints, the ORR considers the extent to which a new open access service would generate new journeys, and the extent to which revenue is abstracted from existing services. The primary purpose of this study is to undertake analysis of historical market entry and expansion of open access operators (OAOs) to attempt to identify the revenue 'generated' by OAOs and the revenue 'abstracted' from franchised train operating companies (TOCs). The aim is that the evidence and findings of this study could be used to inform future access applications.

The proposed methodology for the study was shared with a number of key stakeholders, made up of representatives from current OAOs (Hull Trains and Grand Central), franchised TOCs (Virgin West Coast, Virgin Trains East Coast), and representatives from TOC holding companies (Stagecoach and First Group). The participants provided insights, potential areas of additional analysis and suggestions for changes to the methodology that have been taken on board in the methodology adopted in the study. These are discussed in greater detail in the methodology section of the report.

2.2 Background to Open Access Operators

Since privatisation a number of proposed OAOs have been determined not to be *"primarily abstractive of an incumbent's revenue without providing compensating economic benefits",* found not to cause capacity or performance issues, and the applicants were granted access rights to operate services on the UK rail network. Success has been mixed.

The first OAO, Hull Trains, has been operating since September 2000. It has expanded to running seven daily return services between London Kings Cross and Hull, and has access rights to December 2019. Hull Trains competes directly with Virgin Trains East Coast (VTEC) on its entire route, supplementing the one train per day between London and Hull provided by VTEC and its predecessors.

Wrexham & Shropshire commenced operations in 2008 with five daily services between Wrexham and London Marylebone, providing direct links to London that were not previously provided. Clauses in Virgin West Coast's Track Access Agreement at the time prevented direct competition at Wolverhampton, Birmingham and Coventry; the only competition with franchised TOCs was at Banbury. The service ceased in early 2011.

Grand Central began operating in 2007, and operates services on the East Coast Main Line between London Kings Cross and Sunderland (four daily services) and Bradford interchange (three daily services), with access rights until December 2026. Their services provide links to London from ten stations that had not previously had direct trains.

In August 2015, Great North Western Railway Company Limited (GNWR) were granted rights to run services between Queen's Park in London and Blackpool North (six daily services), with firm rights to call at Crewe, Preston and Poulton-le-Fylde, from the December 2017 timetable change date for a period of 10 years.

In addition to those granted rights, ORR has rejected a number of applications for track access rights because they have failed to pass the "not primarily abstractive" (NPA) test. In its present form, a threshold of generated to abstracted revenue of 0.3 to 1 is intended to recognise the potential benefits to rail users from competition. Practically, this means that where a potential operator is forecast to generate more than 30p of revenue for every £1 abstracted from other operators, it passes the NPA test. Previous rejected applications have included an application by Grand Central to run services between Blackpool North and London Euston in 2010, and an earlier application by Great North Western Railway Company Limited (GNWR).

2.3 The Five Stage Test

The ORR's 'Not Primarily Abstractive' test has been established in ORR's procedures for evaluating track access applications, and has been developed to *"promote competition for the benefit of users and to have regard to the funds available to the Secretary of State".* The test is made up of five stages which are set out below:

- **Stage 1:** using standard industry models (such as MOIRA and the passenger demand forecasting handbook) to make a broad estimate of the likely level of abstraction;
- **Stage 2:** review the estimate established in stage 1 with input from the applicant, potentially affected incumbent operators, funders and any other interested parties;
- **Stage 3:** using benchmarking and survey information from other comparable situations to refine estimates produced by stages 1 and 2; and
- **Stage 4:** assessing the likely impact of services one to two years after introduction to identify material impacts that would not occur immediately after introduction of the new services; and
- **Stage 5:** will consider other relevant factors against quantitative assessment produced under stages 1 – 4.

The types of flows served by OAOs fall into four categories, each of which raises different analytical issues in the application of the NPA test:

- flows with incumbent direct services to London;
- flows with no incumbent direct service to London;
- indirect / connecting passenger flows; and
- non-London flows.

With regard to the second category of flows, the primary industry tool for estimating the revenue and journeys impacts of changes to train service levels, MOIRA, works by applying Generalised Journey Time (GJT) elasticities to a base level of demand. Therefore, this incremental approach may not be appropriate for assessing demand and revenue impacts on flows with little existing demand or where the change in GJT is large; it thus may not accurately quantify the impact of OAOs providing train services not currently provided by a franchise. The Passenger Demand Forecasting Handbook (PDFH) recognises the limitations of MOIRA in using it to model such a situation (extract from section B10.5.1), explicitly referencing the introduction of an OAO:

[redacted]

This highlights the importance of the steps beyond Stage 1 of the five stages of NPA test, which this study aims to inform. As mentioned above, one of the aims of this study is to generate evidence for future access applications, and specifically Stage 3 of the NPA test, and to inform the implementation and approach to Stage 1 in future applications. Historical evidence and data available over the period covered by the operation of the OAOs has been used to inform what OAO revenue has been abstractive and generative.

3 REVIEW OF PREVIOUS WORK

We have reviewed several previous studies and documents which look at OAOs and on-rail competition:

- [redacted];
- On Rail Competition Analysis (2009) – ARUP;
- Second Witness Statement of Mark Jonathan Leving (2006) (From OAO court case);
- Assessment of Alternative Track Access Applications on the East Coast Mainline (2009) – MVA Consulting; and
- Assessment of Aspirations for Track Access on the West Coast Mainline (2011) – MVA Consulting.

Of these, the work by [redacted] and the MVA report into East Coast applications were of most relevance to our study.

3.1 [redacted]

(Document shared in confidence)

3.2 MVA

We reviewed the "*ECML Track Access Applications Assessment*" report, produced by MVA Consulting in 2009. In this report a number of different applications were assessed.

The methodology deployed by MVA was to build an assessment modelling suite, with modules for Station Choice, MOIRA, Fares, Capacity, and Air Competition for Anglo-Scottish applications. From this modelling suite, generated and abstracted demand was assessed and compared. There were many similarities between the approach adopted in the analysis presented in this report and our report. However, as a single static forecast which did not need to include macroeconomic factors or create counter-factuals, the MVA methodology was far simpler.

On discovering that MOIRA2 was not an appropriate tool for modelling the introduction of new operator specific fares, we were able to develop an approach based on the Fares Module used in the MVA report. Through recalibrating spread parameters, we have developed a more robust methodology, which we can suggest for future application assessments. The list of stations included in the MVA station choice model also formed a useful reference point.

4 METHODOLOGY

4.1 Summary

The approach undertaken has two main components that are summarised below:

1. **Identifying** the quantum of revenue abstraction and generation due to open access operators, by going back to a point in time prior to the entry of each of the OAOs, and essentially following Stage 1 of the Five Stage NPA test by creating a Passenger Demand Forecasting Handbook (PDFH) based forecast. This incorporates exogenous demand drivers, such as GDP and population growth, the demand effect of successive timetable changes, and the effect of changes to available fares, both on the overall size of the market, and on the shares of each operator. This forecast is compared against actual revenues, identifying a gap between actual revenues and what could be explained by exogenous driven growth, timetable effects and fares effects (Phase 1a). This identification stage primarily involved the development of a number of PDFH-based revenue models. This process is described in section 4.2 below, with further technical detail in appendix 1.

2. **Explaining** the gap between the forecast generated in the previous step, and further explanation of the breakdown between abstracted and generated demand, beyond the generalised journey time (GJT) and fares effects modelled in the prior stage. This included station accessibility ('railheading'), and other service quality effects, such as the on board environment of the rolling stock, station facilities, customer service, marketing campaigns and brand. The approach for identifying station accessibility (railheading) effects was to build a set of direct demand models for the areas newly served by the open access operators. This is discussed in section 4.4. Investigation of other demand drivers is described in section 4.3.

This section provides a high level description of the methodology employed. The modelling activities and analysis are technical in nature, and are described in greater detail in the appendices to this report.

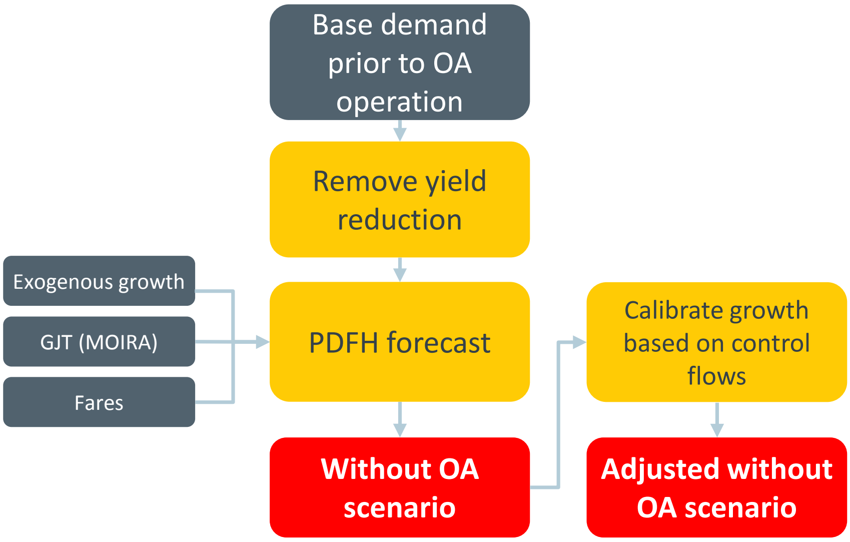
4.2 Identifying the quantum of generation and abstraction

Our approach to identifying revenue generation and abstraction is centred on the creation of a modelled counterfactual historical case, in which the open access operators do not exist. Theoretically, by comparing revenue for the ICEC franchise in this counterfactual case with outturn revenue, it should be possible to identify the proportion of revenue that has been abstracted from ICEC, and that which has been newly generated.

In practice, we use a PDFH revenue forecasting model to create the counterfactual, meaning it is also necessary to model the actual historical situation, to identify its deviation from actual outturn results.

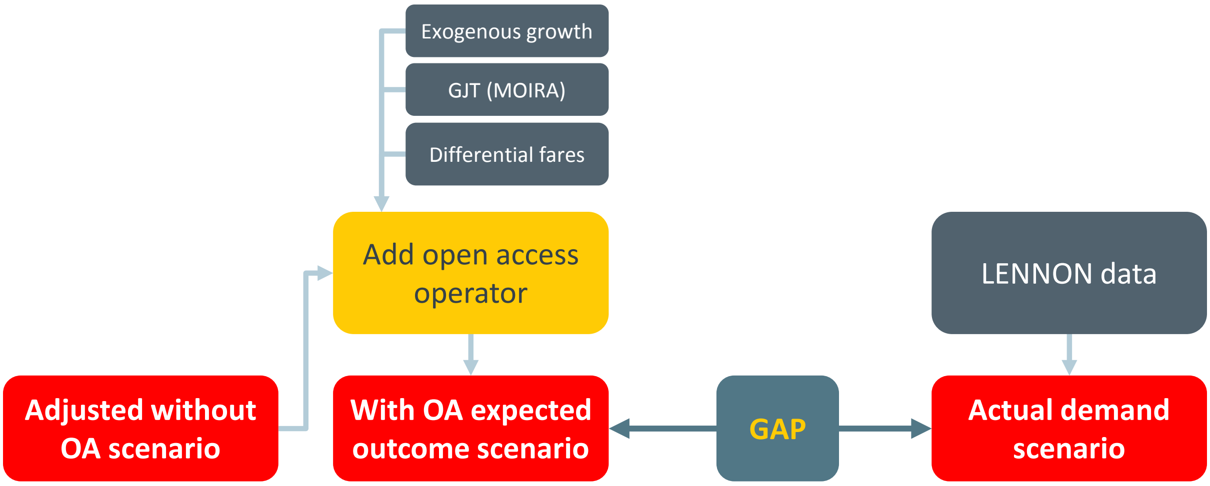
Consequently, this approach requires the development of three PDFH-based revenue forecasting models. Firstly, a model that forecasts revenue for the ICEC franchise in a hypothetical scenario in which the OAOs were not in operation for the flows affected by OAO competition, the "without OA" scenario, or "counter-factual" scenario. This process is shown in Figure 1.

Figure 1 Developing a counterfactual (without OA) scenario



Secondly, a model that forecasts revenue for the InterCity East Coast (ICEC) franchise, Hull Trains and Grand Central, from a base prior to the commencement of the latter two operations (rail year 1999/2000). This second model is the "with OA" scenario, and its development is illustrated in Figure 2

Figure 2 Developing a modelled with OA scenario



Finally, a model that forecasts revenue growth for a number of comparator flows unaffected by OAO competition on the East Coast and East Midlands franchises. This final model is developed in order to establish the extent to which outturn demand has deviated from PDFH over the study period.

The deviation of outturn demand from PDFH derived from the third model is applied as an overlay to demand in both the first and second models, in order that deviation of outturn demand from PDFH is reflected in the with- and without-OA scenarios. This is the "Calibrate growth based on control flows" step in Figure 1.

Even following the application of this overlay, we would expect that there will still be a gap between modelled and outturn demand, because exogenous factors, timetables and fares will not explain all of the demand effect of new direct services. This is addressed in section 4.3.

4.2.1 Modelled drivers of demand

The models described above include as demand drivers the effects of exogenous growth, timetable changes (GJT), and fares. The manner in which these are applied is described below. Further detail on the application of these drivers is provided in Appendix 1.

4.2.1.1 Exogenous growth

The models apply the impact of changes in:

- GDP/Capita;
- Employment;
- Population;
- Car Journey Time;
- Car Operating Costs;
- Car Ownership;
- Bus Journey Time; and
- Bus Costs.

These are the drivers prescribed in the PDFH framework for exogenous growth. PDFH 5.1 elasticities are applied.

4.2.1.2 Timetable

The generalised journey time (GJT) effect of timetable changes on demand is modelled using MOIRA. To avoid over-complexity that would result from modelling each and every timetable in the study period, the following significant timetable changes were selected for inclusion. These were selected on the basis of MOIRA's estimates of demand growth and market shares relative to a September 1999 base timetable on key East Coast Main Line flows.

- May 2000
- September 2000 – **HULL TRAINS COMMENCEMENT**
- June 2002
- September 2002
- December 2004
- June 2005
- December 2005
- December 2006
- May 2007– **GRAND CENTRAL SUNDERLAND COMMENCEMENT** (implemented from December 2007)
- December 2009
- May 2010 - **GRAND CENTRAL BRADFORD COMMENCEMENT**
- May 2011
- December 2011
- May 2013

4.2.1.3 Fares

Following comparison of ICEC fares on flows with competition, with those in the comparator flows model, we could not find conclusive evidence of lower yields on flows with competition. As such, an adjustment has not been made to ICEC yields in the without OA scenario. Fares effects are modelled using PDFH 4 conditional elasticities.

In the with OA scenario, it is necessary to model the effects of differential fares on market share. To do this, we have adopted the methodology set out in PDFH 5.1 for differential fares on open access operators. We have enhanced this approach by calibrating spread parameters for flows of varying distance bands and ticket types.

Due to data and modelling issues, there are limitations to the approach adopted for modelling fares. These are described in section 5.

4.3 Explaining the quantum of generation and abstraction

Following the development of with- and without-OA revenue models, it is possible to identify modelled generative and abstractive effects, based purely on the difference between the modelled with- and without-OA scenarios. However, the modelled factors (exogenous growth, GJT and fares) do not alone explain the changes in demand resulting from the introduction of new direct services.

As such, the second component of the study is concerned with identifying the "gap" between the modelled and actual outcomes. This process is outlined in Figure 2 above. We attempt to identify potential additional drivers of demand, such as the quality of rolling stock. Our approach to establishing the impacts of these drivers is described in this section 4.3, with further technical description provided in appendix 2.

A major driver of the gap is expected to be the demand increase caused by the provision of new direct services to London from stations which previously had no such services, resulting in changes to travel patterns not captured by the GJT elasticity effect. Modelling these effects requires development of Direct Demand Models. This involves a major modelling exercise, and as such is described separately in section 4.4, with further detailed description in appendix 3.

4.3.1 On board offer

There is limited evidence available in PDFH to quantify the demand effects of operators' on board offer. A notable change in on-board offer that occurred during the study period was the introduction of an improved first class offering on East Coast in 2011, which included provision of complimentary food and drinks.

We used this case study as a basis to establish whether demand effects related to TOCs' on board offering could be established. Based on ticket sales data, it was not possible to establish a conclusive link between the improved on-board offer and increased demand. Further detail on this analysis can be found in section 9.1.

4.3.2 Rolling stock quality

We have reviewed the rolling stock used by the operators in the scope of this study. The operators have generally used a mixture of Mark 3 and 4 carriages, and Class 180 DMUs. We consider that these classes of rolling stock offer a broadly similar passenger environment.

In addition to this, there is insufficient evidence of historical train condition in terms of upkeep and cleanliness, and judgements on these factors would in any case be subjective. As such, we have concluded that it is not possible to quantify a rolling stock quality impact on demand. Further details are provided in section 9.2.

4.3.3 Lagged effects

The demand effects of changes to timetables may take some time to come to fruition. This can occur for a variety of reasons, such as lack of awareness of improvements, or a delay in people changing their habits.

In order to establish the effect of a lag in demand changes as a result of open access services, we examined the profile of demand growth on open access flows, as well as some other franchised flows that have benefitted from step changes in service frequency. Demand growth on these flows was compared to that on flows with no such large changes ([redacted] and [redacted]).

This analysis identified continued demand growth over a longer period than that defined by PDFH. On this basis, the lag profile in Table 1 was derived, and implemented in the modelling of introductions of open access services.

Table 1 Estimated lag profile for use in OAO forecasts

	Average Q1	Average Q2	Average Q3	Average Q4	Average Y1	Average Y2	Average Y3	Average Y4	Average Y5
Major new services	[redacted]								

Further details of the derivation of this lag profile are provided in section 9.3.

4.3.4 Station quality

Although we have been provided with some details of past station improvement schemes, this does not necessarily provide us with a full picture of historical station condition throughout the study period. We would hesitate to recommend that station quality impacts be considered an important component of Open Access Operator generation – station schemes could just as equally be ameliorating the impacts of Open Access Operations.

We thus do not include any station quality impacts, or an externality of these on other Train Companies, in our modelling work. Further details are in section 9.4.

4.3.5 NRPS as an indicator of service quality

Transport Focus supplied us with 32 waves of National Rail Passenger Survey (NRPS) data (since the survey began) for the three TOCs in the study. We produced 'logit' choice models to estimate the probability of users claiming to be fairly or very satisfied in their NRPS response (the headline measure of customer satisfaction).

When we control for the wave of the survey, the passenger's satisfaction with value for money, their satisfaction with the room to sit/stand and whether or not they reported that their journey was delayed, respondents travelling with either open access TOC (as opposed to the franchised TOC) are not significantly more satisfied.

This result was robust to a change in the sample years and removing satisfaction with value for money.

Our interpretation of this is that given levels of fare, crowding and punctuality - which have known demand impacts and accepted methodologies for their assessments - there is no evidence that open access TOCs generate further demand by way of customer service factors, and thus our model does not require an additional overlay. Further details of this analysis are provided in section 9.5.

4.3.6 Marketing/branding effectiveness

We have undertaken some research on historical marketing campaigns of open access operators. From the analysis undertaken it would appear that the majority of marketing schemes seem to be targeted primarily at generating additional demand by promoting rail use and travel destinations. However, this does not provide evidence that advertising would be purely generative.

The effect of marketing on rail demand is typically quantified by estimating rates of return on marketing spend. As part of this study we have not been provided with access to data on the historical marketing spend of ICEC or the open access operators. Even if such information were available, quantifying its impact would present significant challenges. As such, we have not modelled the effect of marketing campaigns. Further details of our research on historical marketing campaigns can be found in section 9.6.

4.3.7 Car parking

Availability of car parking at a station is a potential driver of passenger demand generation or abstraction. Examples of this might include:

- Generation of demand, where it was previously suppressed by a lack of available parking space at an incumbent operator's station car park, which now has available spaces as people switch to a station newly served by an OAO; or
- Abstraction, where a person who currently railheads to a nearby incumbent station, has the option of driving to an OAO-served station that is closer and has parking provision.

Conversely, a lack of spaces at an OAO-served station might prevent as much generation or abstraction as would otherwise occur.

To assess the effects of parking provision at stations, it is necessary to have some idea of the extent to which spaces are occupied. Car park occupancy data is not public domain and is not held by ORR. Operators attending the stakeholder workshop held at project inception were asked if they could provide car park occupancy data. However, no data has been provided. It has therefore not been possible to provide meaningful analysis of any effects of car parking provision in this study.

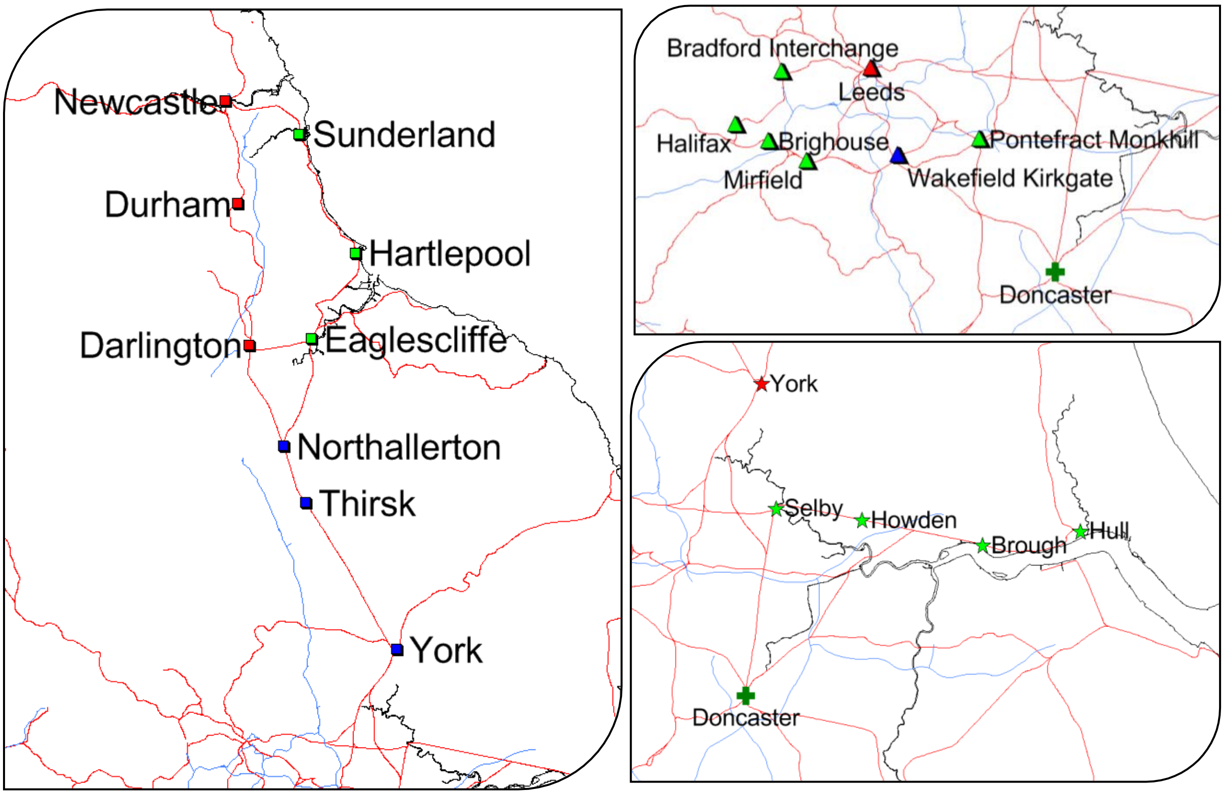
4.4 Direct Demand Modelling

Where open access operators provide new, direct services to stations previously poorly served, the result is similar to that of an entirely new station. In both cases PDFH recommends the same approach - the use of a 'gravity' or 'direct demand' model (PDFH 5.1 – B10.2). This is because in such cases, GJT and fares elasticities alone do not capture the full impact of the new service. In order to capture the impact of improved station accessibility, we created 'combined demand and accessibility choice' models based on a PDFH recommended approach (PDFH v4, section B6). This has been used to calculate how many trips are abstracted from other stations and how many new trips are generated due to the proximity and convenience of a new service offering for people in the vicinity of the open access station.

The approach is to create a series of population catchment models which determine trip rates from a particular zone to each station of interest, based on the generalised cost of making a journey via that station. The generalised cost used to determine an individual's choice of station includes the journey time and cost of accessing the station (including car operating costs), and the journey time, frequency and cost of the rail journey.

Three models have been created, two covering the Grand Central catchment areas (Sunderland and Bradford, separately) and a third for the Hull Trains catchment area. The area of each model is shown in the figure below. Each of the three models is calibrated to a 'pre-open access' market position, from which the 'post open access' impacts of a direct service and lower fares can be analysed. For the purpose of this analysis, only journeys to London are modelled.

Figure 3 Maps of Direct Demand Models



Following introduction of the new services, the models calculate trip rates to each station. This model also includes the effects of fare and GJT (which are already captured in the standard revenue model). As such, these effects are "reversed out" using standard PDFH elasticities prior to transferring a passenger miles overlay to the "with OA" revenue model.

Further technical detail concerning the development, calibration and outputs of the Direct Demand Models is contained in appendix 3.

5 LIMITATIONS

5.1 Summary

There are a number of limitations to the analysis carried out for this study, with the primary caveats relating to the modelling of fares effects and the availability of data.

The methodologies used for modelling fares are consistent with the size and scope of the study, and with PDFH and DfT guidelines and transport demand modelling. We have adopted an approach that we believe is the best possible given these constraints and available data, and we have made some improvement to the existing industry methodology for quantifying the market share effects of differential fares. However, there are limitations to our fares modelling, which are detailed below.

The LENNON data that was utilised extensively had a number of issues. These are also detailed below.

5.2 Treatment of fares modelling

We have strived to develop a fares modelling methodology that is robust as possible within the remit of this study. However, there are several key limitations to the treatment of fares in our modelling.

5.2.1 Market size

Our modelling of market size effects is based on PDFH 4.0 elasticities, in accordance with WebTAG guidelines. We have modelled the demand effect of yield changes for Full, Reduced and Season tickets in isolation, using conditional elasticities. This does not fully take into account ticket type switching effects. This approach is common, but a more rigorous application of prescribed PDFH methodology would fully incorporate ticket type switching. This would require the development of a fares model. Such a fares model would require development of a bespoke methodology to capture the effects of:

- changes to fares;
- changes to the validity of Off Peak tickets; and
- changes to the availability of Advance tickets.

There is currently no widely accepted methodology to capture all of these effects, and development of a fares model is beyond the scope of this study. We have tested the effect of using a single average yield (across all ticket types) and a single elasticity for each flow, applying the resulting demand multiplier to revenue for all ticket types. The adopted approach of modelling ticket types in isolation resulted in modelled effects that are closer to outturn. This is discussed in greater detail in Appendix 4 – Fares methodology sensitivity tests.

5.2.2 Market share

The methodology we have adopted to model the market share effects of operator-specific fares is prescribed in PDFH 5.1 (section B11.4), and has been used by ORR previously. We have improved upon this methodology by calibrating spread parameters for flows of different distance and for different ticket types. However, because the formula is explicitly designed to model effects on walk-up fares, we have not applied it to Advance fares. Instead, the market shares derived from the Off Peak fare differentials are applied to both Off Peak and Advance revenue. Additionally, the differentials for Standard class tickets have been used to model market share for both ticket classes.

5.2.3 Competitive response

We have performed analysis seeking to identify a difference in yields on flows where ICEC faces competition, relative to those where it does not. We have found no evidence of historical ICEC franchisees having a consistent strategy of reducing yields on flows with competition, relative to those without. This finding does not represent conclusive evidence that there has not been a competitive response, merely that we have not identified conclusive evidence of one. As such, we have not incorporated an increase in yields in the counterfactual (without OA scenario) to account for a competitive response in the with OA scenario.

Figure 4 ICEC Off Peak + Advance average yields (Grand Central Sunderland flows)

[redacted]

Figure 4 and Figure 5 illustrate the lack of evidence for a competitive response. Average yields on Reduced ticket types do not vary significantly on ICEC flows with competition from those without (Newcastle and Leeds), following the commencement of open access services. We have made similar findings for Full (Peak) tickets.

Figure 5 ICEC Off Peak + Advance average yields (Grand Central Bradford flows)

[redacted]

Based on the evidence available to us, we have not identified a competitive response. However, were a competitive response to have occurred, the effect of assuming it did not on the modelled generation-abstraction split depends on the fare elasticity:

- As stated above, we have used PDFH 4.0 in our modelling. For long distance London flows, PDFH 4.0 recommends an overall elasticity of [redacted]. This means that a decrease in fares as a result of competitive response would [redacted] the abstraction-generation split.
- However, PDFH 4.0 gives different conditional elasticities for different ticket types. For example, the conditional elasticity for Reduced fares on long distance London flows is between [redacted] and [redacted]. If there were a competitive pricing response occurring primarily through reductions to the price of Reduced tickets without affecting other ticket types, a ticket type switching model should be created to fully understand the effects.
- The latest version of PDFH (version 5.1) recommends an overall elasticity of [redacted]. This implies that that any competitive response would be generative. [redacted].

It should be noted that although it forms the latest guidance, the fares elasticities provided by PDFH 5.1 have not been included in DfT's WebTAG guidance. WebTAG[1] recommends the use of PDFH 4.0 elasticities, and these remain the most widely accepted parameters within the industry. To fully understand the effects of differential fares changes by ticket type a ticket type switching model should be created. This was outside of the scope of this study.

5.3 LENNON Data Issues

This study utilises LENNON revenue and passenger miles data for several purposes:

- As a base for forecasting
- For comparison of modelled revenue with outturn

[1] TAG unit M4, Chapter 8; https://www.gov.uk/government/publications/webtag-tag-unit-m4-forecasting-and-uncertainty-november-2014

- To derive average yields for fares modelling

Five key issues have been encountered with the LENNON data provided:

- No LENNON data is available prior to RY 2002;
- The transition between the CAPRI and LENNON systems in period 6 of 2004 has caused a discrepancy in the Passenger Miles data;
- There are anomalous results for both yield and revenue per passenger mile, for some flows and ticket types;
- There is some erroneous categorisation of ticket types which has occurred within the LENNON system; and
- It has not been possible to gain access to the full timeframe of periodic LENNON data required at this stage.

Mitigations have been found for most of the issues encountered, and as a result LeighFisher view the conclusions reached in this report as robust. Nevertheless, these caveats should be born in mind, especially when looking at results relating to the Hull Trains service introduction.

5.3.1 LENNON Data availability

LENNON data is unavailable prior to RY 2002, as this is the limit of [redacted] historical data storage. This is a particular issue for the Hull Trains service introduction, as the initial service introduction took place in September 2000. We contacted Hull Trains to ask if any they had the relevant data stored within any of their archives, but unfortunately they did not.

To mitigate this issue, data for RY2000 and RY2001 has been created through the extrapolation of LENNON data using MOIRA base data. This source of data was created from the CAPRI system originally, and it provides a sound basis for the creation of the "Actuals" for the model.

It should be noted however that we have only been able to extrapolate this data on a "National Rail" basis, due to the limitations of the MOIRA system. It has therefore been impossible to accurately assess Hull Trains' sales during the first six months of service.

5.3.2 CAPRI – LENNON Transition

A discrepancy has been observed in the Passenger Miles data, resulting from the introduction of the LENNON system in RY2004 P06, particularly for TOC-specific tickets. A similar issue has been observed across all TOCs. When questioned on the cause of the discrepancy, [redacted] have said:

"While we are not aware of any specific changes in how miles were calculated between the two systems, it's possible that the switch to LENNON did result in the higher numbers that are visible within the data."

To mitigate this issue, we have applied appropriate scalings to data for passenger miles prior to 2004 P06. These have been derived from a combination of MOIRA data, official ORR data for passenger miles, and comparison with Journeys data. This has allowed us to adjust data prior to this point to be consistent with later data.

Figure* 6 *EC Total Journeys and Passenger Miles data (2003 P06 = 1), Capri/LENNON unadjusted

[redacted]

5.3.3 Anomalous Yield Results

For numerous flows and ticket types, anomalous yields have been observed. This has particularly been an issue for stations that were without a direct service to London prior to Open Access service commencement. It is not unusual to find some issues with yield data when building a Revenue Model, and typically a useful tool to adjust for this issue is to use revenue per passenger Mile data instead. However, issues have also been observed in this data set, hampered by the inconsistent data prior to RY2004 P06.

As a result, it has been necessary to apply LeighFisher's subjective judgement on a case by case basis, in order to make manual adjustments to yield (ie revenue per passenger mile levels) in those instances where LENNON results give implausible movements in yield.

5.3.4 Erroneous Categorisation of ticket types within the LENNON system

It appears to be the case that some tickets have been miscategorised into the wrong product groups within the LENNON system. For example, a very high number of "East Coast only" full priced ticket sales were observed. As East Coast are the lead operator on almost all of the flows looked at, they would not normally be permitted to set an "East Coast only" walk-up fare.

To mitigate this issue, it has also been necessary to apply LeighFisher's subjective judgement on a case by case basis. For the example described above, it was deduced that these sales represented miscategorised East Coast Advanced Purchase tickets, and the data was adjusted accordingly.

5.3.5 Incomplete periodic dataset

At the time of release, it has not been possible to obtain the full periodic dataset necessary to investigate the lagged effects of the Grand Central Bradford service introduction. This is because of a delay in the restoration of data for RY2011 P05, P06, and P07.

6 RESULTS

6.1 Summary

The previous sections describe the approach undertaken to identify the differences in revenue between outturn and forecast revenue both with and without OAO competition, and the methodology to understand the drivers of this difference related to timetable and fares effects.

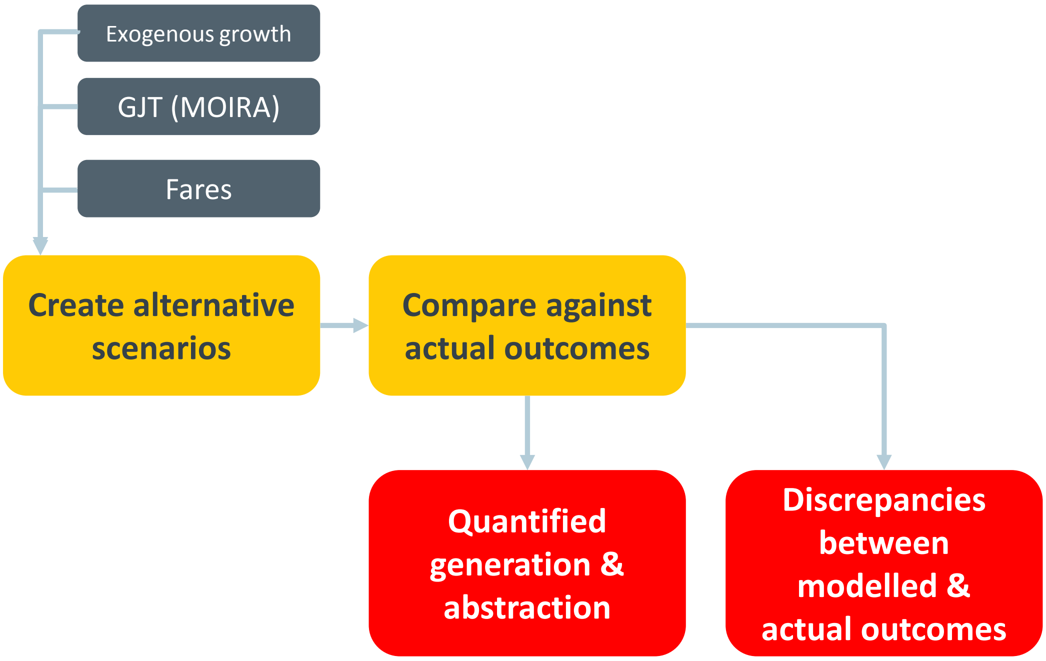
In this section we explain our results, identifying the quantum of revenue generation and abstraction. We then compare these to actual results, demonstrating the differences between our approach and alternatives, and the lessons that can be learned for future studies.

In turn, we discuss the following incremental approaches to quantifying generation and abstraction:

- Use of MOIRA;
- A Revenue Model;
- Inclusion of TOC-specific tickets; and
- Inclusion of direct demand ('railheading') modelling.

We then compare our results to the actual revenue observed on the East Coast Mainline. We use this to provide bounded estimates of the quantum of revenue generation and abstraction on the line.

Figure 7 Summary of modelled counterfactual v actuals



6.2 Use of MOIRA

In this section, we will describe the outputs of MOIRA for the Open Access Operations on the ECML.

Modelling the impact of a change in timetable is, by the recommendations of PDFH and WebTAG, best undertaken using MOIRA (sometimes referred to as 'MOIRA1'). PDFH contains instructions on calculating the impacts of improvements in service level, i.e. improvements in Generalised Journey Time (GJT). MOIRA is used for this purpose because:

- It contains data on the volumes of demand on each origin-destination (O-D) pair on the rail network, which are collated from LENNON;

- It is a model of ORCATS (the rail industry algorithm for allocating revenue between TOCs), and thus calculates the actual change in the distribution of revenue from walk-up tickets between operators following a service change;
- It calculates GJT with reference to the demand distributions and behavioural model contained in ORCATS. This rewards additional/accelerated/direct services at the times of the day that passengers want to travel; and
- It can identify calculate GJT (and thus demand impacts) for all the O-D pairs affected by the service change. This allows connecting flows to be taken into consideration very easily.

MOIRA, however, has well-known deficiencies (in part because there is no accepted behavioural model that would allow their resolution), such as:

- The demand distributions in ORCATS are at least twenty years old. It is known that there are more passengers travelling in the evening than MOIRA (or ORCATS) would reward, for instance;
- It does not take into account changes in station choice that may result from changes in train services at different stations – the GJT elasticity is applied to each O-D pair independently;
- MOIRA does not "know about" TOC-, route- or train-specific tickets, and so the base TOC earnings split is not generally correct. Differences in yield by TOC, route or train are not modelled;
- There is no explicit treatment of advance tickets – these are treated as "reduced" (i.e. "off-peak" or "super off-peak" tickets) even though they are likely to be sold in different volumes at different prices at different times of the day. Advance ticket-holders make the majority of passengers on many InterCity flows now; and
- There is no treatment of demand suppression caused by train crowding.

Nevertheless, MOIRA is a key benchmark for our results. It is straightforward to use and can quickly calculate the impact of changed timetables on all train companies' services.

The table below shows results from MOIRA for the Open Access TOCs in this study in the December 2014 timetable, based on revenue for the year ending March 2014, and revenue for the year ending March 2000 (i.e. before Hull Trains commenced).[2] Accordingly, they are in different price (and volume) bases.

Table 2 MOIRA modelled impact of Open Access operators in the December 2014 timetable

	2013/14 volumes and prices	1999/00 volumes and prices
With Open Access		
EC Revenue	[redacted]	[redacted]
GC Revenue	[redacted]	[redacted]
HT Revenue	[redacted]	[redacted]
Without Open Access		
EC Revenue	[redacted]	[redacted]
Impact of Open Access		
EC Revenue	[redacted]	[redacted]
Other TOC Revenue	[redacted]	[redacted]
Share of OA revenue generated	23.8%	19.8%
Generated per £1 abstracted	31.1p	24.7p

[2] The results for Year to March 2014 were estimated by deleting the OAO services from the timetable. The results for Year to March 2000 were estimated using the same timetable, but adjusted to take into account that the base revenue does not include OAO services. We have not adjusted the year to March 2000 outputs to take into account changes in prices or changes in demand that is not due to the OAO's GJT improvement.

The share of revenue generated is higher when modelled using 2013/14 volumes and prices. This is because flows where the open access operators are more generative rather than abstractive have grown faster than flows on which they are more abstractive – this could be because exogenous growth has been more favourable on the former types of flows or because of railheading, which is not modelled in MOIRA (but is accounted for in our direct demand modelling approach discussed later).

In our modelling work, we have used MOIRA's application of GJT elasticities to assess the change in market sizes and ORCATS shares that result from (ten) timetable changes.

6.3 A revenue forecasting model

MOIRA is an exercise in "comparative statics" – it should calculate *what demand/revenue would be* if *the new timetable operated today*. Aside from the deficiencies in MOIRA itself, there are a number of items that require a forecasting model:

- MOIRA does not incorporate the impact of *exogenous* drivers on revenue (i.e. the economic drivers of demand for rail travel). The abstraction estimate from MOIRA would only be valid if future growth in rail revenue was uniform across all flows (and across all ticket types). This means that a new train service that generates demand in future growth markets (e.g. a destination whose economic growth prospects are better than average, or the reduced ticket market which is usually forecast to grow faster than the season ticket market) would not be given enough credit, as is suggested in the above analysis showing revenue generation greater in 2013/14 volumes compared to 1999/00 volumes;
- MOIRA does not incorporate the impact of *endogenous* drivers on revenue (i.e. the factors that are within the control of the train operator). Rolling stock quality and marketing (amongst other things cited in PDFH and elsewhere) can be expected to lead to further growth in the rail travel market over and above exogenous factors. Alternatively, car park and train capacity may suppress future revenue growth. Changes in fares offering are also not incorporated in MOIRA results; and
- MOIRA does not incorporate any lagged effects. It may take time for demand growth to be fully realised. Incorporating lagged effects would tend to make a new service seem more abstractive over a given time period, because abstraction is (nearly) instantaneous but generation takes time.

For *forecasting* models, there is of course dissensus about future economic growth (and especially its location) and, indeed, of what initiatives train companies themselves will implement to grow their revenues. Our *backcasting* model has the advantage of being able to include real exogenous data, as well as to be able to calibrate other growth (i.e. growth in excess of that predicted by the exogenous factors and the elasticities in PDFH) based on outturn on other flows.

Revenue models do need to be appropriate for the rail routes served. Our model does not cover the whole of the National Rail network, only selective flows served by the three TOCs operating on the East Coast Mainline – a revenue model that covered every TOC and every flow would be difficult to manage. As a result, we rely on estimates from MOIRA to understand the externality of train service changes onto TOCs not explicitly modelled.[3]

Table 3 Impact of Open Access operators in 2014/15 in our Revenue Model, before differential fares

	2014/15 forecast volumes in 2013/14 prices
With Open Access	

[3] This could vary in either direction. Additional services between Crewe and London, say, might have a positive externality on Arriva Trains Wales by growing the market between North Wales and London. They might have a negative externality on London Midland by taking passengers between Crewe and Stafford that would otherwise use LM services.

EC Revenue	[redacted]
GC Revenue	[redacted]
HT Revenue	[redacted]
Without Open Access	
EC Revenue	[redacted]
Impact of Open Access	
EC Revenue	[redacted]
Other TOC Revenue	[redacted]
Share of OA revenue generated	21.4%
Generated per £1 abstracted	27.3p

The share of revenue generated in this model is lower than from using 2013/14 MOIRA alone. This is because the forecasting model has a 1999/2000 revenue base, with revenue forecast forward from this base year, and this 'vanilla' forecasting model does not entirely capture the growth on those segments served by the Open Access operators. This is illustrated on Figure 8 below:

Figure 8 Forecast error at selected stations, Revenue Model without differential fares or railheading

[redacted]

This deficiency would obviously not be applicable assessing Open Access operations *ex ante*. Using a revenue model suggests Open Access operators generating more revenue than using 1999/00 MOIRA, because Hull Trains and Grand Central have served markets which have grown faster (primarily the reduced ticket markets).

The other advantage of revenue forecasting models is that we can assess dynamic effects. Table 4 below shows the modelled impact of Open Access operators over the last fifteen years.

Table 4 Impact of Open Access operators in our Revenue Model, before differential fares

	Total volumes over fifteen years, in 2013/14 prices
With Open Access	
EC Revenue	[redacted]
GC Revenue	[redacted]
HT Revenue	[redacted]

	Total volumes over fifteen years, in 2013/14 prices
Without Open Access	
EC Revenue	[redacted]
Impact of Open Access	
EC Revenue	[redacted]
Other TOC Revenue[4]	[redacted]
Share of OA revenue generated	20.0%
Generated per £1 abstracted	25.0p

Over the full fifteen years covered by the Revenue Model, there is a slightly higher modelled level of abstraction than that estimated from the final single year in the model (as shown in Table 3) – this is because of changes in the timetable over time and because the generation process takes time to realise (the lag is implemented in the model), whereas changes in market shares are instantaneous.

6.4 Including the impact of TOC-specific tickets

Section 8.2 discusses our approach to modelling the impact of TOC-specific tickets. As shown in Table 5 below, this has a significant impact on the Open Access Operators' revenue lines and the share of revenue that has been abstracted.

Table 5 Impact of Open Access operators in 2014/15 in our Revenue Model, including differential fares

	2014/15 forecast volumes and 2013/14 prices	Total volumes and 2013/14 prices
With Open Access		
EC Revenue	[redacted]	[redacted]
GC Revenue	[redacted]	[redacted]
HT Revenue	[redacted]	[redacted]
Without Open Access		
EC Revenue	[redacted]	[redacted]
Impact of Open Access		
EC Revenue	[redacted]	[redacted]
Other TOC Revenue	[redacted]	[redacted]
Share of OA revenue generated	17.3%	12.4%
Generated per £1 abstracted	20.9p	14.1p

Adding differential fares does not have a large impact on total industry revenue (a reduction of approximately [redacted]) because the price elasticities of demand [redacted].

However, the modelled effect of differential fares is to increase the market shares of Grand Central and Hull Trains, and their revenue accordingly. No more revenue has been generated, so this means that the estimated level of abstraction is significantly higher.

4 To simplify the modelling, this value uses the impact on other TOCs in the December 2014 timetable. The impact would be slightly different when considering every timetable.

The impact of including TOC-specific tickets, as we shall see in section 6.5 below, is to increase the Open Access TOC's revenues and market shares closer to the observed LENNON totals. We think this is an important addition to modelling of Open Access operators, although we are conscious that we have the advantage of *ex post* data on actual yields on TOC-specific tickets which will not be available from an application.

6.5 Performance against actuals

Figure 9 Headline model performance (Modelled from LeighFisher, Actuals from MOIRA/LENNON)

[redacted]

As can be seen in Figure 9 above, our model performs reasonably well, tracking revenue until 2007/08, and then only [redacted] difference in 2014/15 year. Divergence from actuals is strongest in 2000/01 (where we can imagine that the Hatfield accident had a more severe impact on the East Coast Mainline than on our comparator flows) and since the downturn, where PDFH performs poorly but the overlay from our comparator flows does not close the whole gap.

However, we do not match the Open Access market shares as well. While Hull Trains revenue is reasonably well-replicated ([redacted] over-forecast at [redacted] is matched by revenue being under-forecast at [redacted]), Grand Central's revenue is under-predicted.

By definition, Open Access (and, indeed, franchised) operators' revenue is incorrect because of errors in our forecasts of market size and errors in our forecasts of market share. We shall discuss these in turn, as they present distinctive issues and solutions, and different issues on each of the flow types. We conclude by providing bounded estimates of the level of abstraction and generation by Open Access Operators on the East Coast Mainline.

6.5.1 Performance of market size

Figure 10 Forecast errors of total market size for four flow types

[redacted]

Table 6 Total market size for Revenue Model flows (except 'Other' flows)

Flow type	Flow	2014/15; 2013/14 prices		
		Actual	Modelled	Error
Flows with existing direct London trains	Doncaster	[redacted]	[redacted]	[redacted]
	Grantham	[redacted]	[redacted]	[redacted]
	Leeds	[redacted]	[redacted]	[redacted]
	Newcastle, Durham, Darlington	[redacted]	[redacted]	[redacted]
	Northallerton	[redacted]	[redacted]	[redacted]
	Retford	[redacted]	[redacted]	[redacted]
	Sheffield	[redacted]	[redacted]	[redacted]
	Wakefield	[redacted]	[redacted]	[redacted]
	York	[redacted]	[redacted]	[redacted]
Flows with new direct London trains	Bradford	[redacted]	[redacted]	[redacted]
	Brighouse, Mirfield	[redacted]	[redacted]	[redacted]
	Brough	[redacted]	[redacted]	[redacted]

Flow type	Flow	2014/15; 2013/14 prices		
		Actual	Modelled	Error
	Eaglescliffe	[redacted]	[redacted]	[redacted]
	Halifax	[redacted]	[redacted]	[redacted]
	Hartlepool	[redacted]	[redacted]	[redacted]
	Hull	[redacted]	[redacted]	[redacted]
	Pontefract	[redacted]	[redacted]	[redacted]
	Selby, Howden	[redacted]	[redacted]	[redacted]
	Sunderland	[redacted]	[redacted]	[redacted]
	Thirsk	[redacted]	[redacted]	[redacted]
Indirect/ connecting flows	Huddersfield	[redacted]	[redacted]	[redacted]
	North East Yorkshire	[redacted]	[redacted]	[redacted]
	North Lincolnshire	[redacted]	[redacted]	[redacted]
	Teesside	[redacted]	[redacted]	[redacted]
	West Riding	[redacted]	[redacted]	[redacted]
Non-London flows	Doncaster - Bradford	[redacted]	[redacted]	[redacted]
	Doncaster - Hull	[redacted]	[redacted]	[redacted]
	Grantham - Hull	[redacted]	[redacted]	[redacted]
	Hartlepool - York	[redacted]	[redacted]	[redacted]
	Northallerton - York	[redacted]	[redacted]	[redacted]
	Sunderland - York	[redacted]	[redacted]	[redacted]

It is clear from Figure 10 and Table 6 that our model over-forecasts revenue for flows with existing direct London services and under-forecasts revenue on flows that gain direct London trains.

Under-performance on some flows with direct services (e.g. Newcastle/Durham/Darlington) may well be *due* to the additional revenue earned on nearby flows gaining direct services (e.g. Sunderland/Hartlepool/Eaglescliffe); this would be due to changes in *railheading*. This would be one station abstracting the revenue of another. However, some of the under-performance may be due to regional economic factors not captured in our data, for instance, the strength of the Leeds economy compared to Doncaster or Hull not being captured by Yorkshire & The Humber's GVA.

Performance on indirect and non-London flows is more mixed, but the non-London flows are relatively small in size but very numerous. Non-London flows make up [redacted] (under-forecast[5] by [redacted]) of East Coast Main line revenue in 2014/15 but we would expect few to have experienced significant effects from open access (OAOs receive only [redacted] of the [redacted]); they are also affected by changes in other TOCs' services because LENNON reflects commissions turned over by ECML TOCs and by changes in the inclusion of PTE products in LENNON.

[5] This will be in part because our base is from MOIRA, not LENNON, so does not include historical commissions data (e.g. ICEC will earn no revenue in our model from Wakefield-Fitzwilliam tickets sold at Wakefield ticket offices). This will also be due to PDFH under-forecasting non-London demand throughout time period and especially since 2007/08.

6.5.2 Performance of market shares

Figure 11 Grand Central market shares on selected flows

[redacted]

Figure 12 Hull Trains market shares on selected flows

[redacted]

It can be seen from Figure 11 that we generally under-predict [redacted] market shares (the upper end of the red bars or lower end of green bars is the actual revenue share). However, the gap between orange circles and the open circles shows the improvement in model performance that comes from including differential fares.

Performance on [redacted] is less consistent, with our model generally over-predicting market shares – differential fares seem to have a smaller impact on [redacted] than modelled.

Figure 13 Modelled and actual market shares (compared to ICEC)

Flow type	Flow	2014/15 revenue shares	
		Actual	Modelled
Flows with existing direct London trains	Doncaster	[redacted]	[redacted]
	Grantham	[redacted]	[redacted]
	Leeds	[redacted]	[redacted]
	Newcastle, Durham, Darlington	[redacted]	[redacted]
	Northallerton	[redacted]	[redacted]
	Retford	[redacted]	[redacted]
	Sheffield	[redacted]	[redacted]
	Wakefield	[redacted]	[redacted]
	York	[redacted]	[redacted]
Flows with new direct London trains	Bradford	[redacted]	[redacted]
	Brighouse, Mirfield	[redacted]	[redacted]
	Brough	[redacted]	[redacted]
	Eaglescliffe	[redacted]	[redacted]
	Halifax	[redacted]	[redacted]
	Hartlepool	[redacted]	[redacted]
	Hull	[redacted]	[redacted]
	Pontefract	[redacted]	[redacted]
	Selby, Howden	[redacted]	[redacted]
	Sunderland	[redacted]	[redacted]
	Thirsk	[redacted]	[redacted]
Indirect/ connecting flows	Huddersfield	[redacted]	[redacted]
	North East Yorkshire	[redacted]	[redacted]
	North Lincolnshire	[redacted]	[redacted]
	Teesside	[redacted]	[redacted]
	West Riding	[redacted]	[redacted]
Non-London flows	Doncaster - Bradford	[redacted]	[redacted]
	Doncaster - Hull	[redacted]	[redacted]
	Grantham - Hull	[redacted]	[redacted]
	Northallerton - York	[redacted]	[redacted]
	Sunderland - York	[redacted]	[redacted]

6.5.3 Inferring levels of abstraction

Our modelling calculates level of abstraction – the loss in Inter-City East Coast revenue is abstraction, and onto this we add the impact on other TOCs, and OAO revenue that has not been abstracted is generated.

However, it also under-predicts open access operators' revenue. In total, our model forecasts OAO revenue at £[redacted] (in 2014/15, in 2013/14 prices) against an actual total of £[redacted][6]. It is reasonable

[6] We have not checked this against actual statements. It is unlikely to match, as our modelling excludes non-LENNON revenue (e.g. some PTE products, some car parking, catering) and some LENNON revenue we have not segmented (primarily because it is non-Geographical).

to suggest that there is uncertainty associated with the calculated share of abstraction versus generation.

We can take advantage of this decomposition between market size and market share to infer levels of abstraction. Errors in market share are, by definition, errors in the level of abstraction. We can thus calculate a quantity of *unmodelled* abstraction. We calculate this as:

$$(OAO\ actual\ revenue\ share - OAO\ modelled\ revenue\ share) \times actual\ total\ revenue$$

Across the thirty-six model flows, this totals [redacted] (30% of the gap) – a combination of [redacted] at stations (like [redacted]) where the OAO share was higher than modelled, and [redacted] at stations like [redacted] where the OAO share was lower than modelled (i.e. so less revenue than modelled had been abstracted).

The remaining [redacted] is an unknown mixture of abstraction and generation. This is illustrated in Table 7 below.

Table 7 Locations and Sources of OAO revenue

Flow type	Modelled Generated	Modelled Abstracted	Unmodelled Abstracted	Unmodelled, Abstraction and/or Generation	Total
Flows with franchised direct services	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Flows with new direct services	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Indirect/ connecting flows[7]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Non-London flows	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Flows with franchised direct services	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Flows with new direct services	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Indirect/ connecting flows	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Non-London flows	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total	13.8%	66.0%	6.1%	14.1%	

We know that at least 72.1% of revenue is abstracted, and at least 13.8% of revenue is generated. Between 13.8% and 27.9% of revenue, has been generated – between 16p and 39p for every £1 of revenue abstracted.

This is a static estimate – of 2014/15 alone. The level of generation is lower covering all of the last fifteen years, as in Table 5 – between 12% and 23% (13p and 30p).

[7] The "unmodelled" impacts are negative on these flows because the total market size is smaller than modelled.

6.6 Direct demand model – performance against actuals

Figure 14 Total demand, with Direct Demand Model

[redacted]

[redacted]

As can be seen from Figure 14 above, adding the direct demand model improves replication of actuals for Grand Central but deteriorates replication for Hull Trains and ICEC. The total revenue on the East Coast Mainline increases too, because the Direct Demand model gives a larger volume growth than GJT effects alone, particularly along the Hull and Sunderland routes, reflecting the significant improvements in accessibility to towns along the routes.

Figure 15 Impact of the Direct Demand Model by flow type

[redacted]

At most of the stations with new direct services, the direct demand model improves the fit of the model against actuals, although the change is relatively modest at some stations (e.g. [redacted]). It is clear that progress is made in understanding the changes in growth at some stations. However, at Leeds a significant volume of passengers switch to other stations, which makes the under-forecasting of the previous model worse. Equally, at [redacted], there is significant growth (because of a greater propensity to travel to London amongst a large population), making the over-forecasting of the previous model worse too.

Although there is a clear benefit to modelling railheading, there may well be scope for refinements both in the direct demand modelling, and the total demand model, to improve goodness of fit.

6.6.1 Results

The tables below show the results from the Revenue Model including direct demand. Table 8 shows the modelled levels of demand generation – 30% of modelled Open Access operator revenue is generated.

Table 9 shows a decomposition of outturn Open Access Operator revenue in 2014/15 into modelled and unmodelled, generated and abstracted components. The Direct Demand Modelling introduces interactions between flows, and so there is negative generation on "flows with franchised direct services" – this is because the total revenue on these flows is smaller because of Open Access entry, and the generated revenue is the difference between total revenue with and without Open Access. These passengers are then recorded as generated to "flows with new direct services"; the *total* ([redacted], 31%) quantum of generation is consistent with the model.

Table 8 Results from the Revenue Model with Direct Demand

	2014/15 forecast volumes and 2013/14 prices	Total volumes and 2013/14 prices
With Open Access		
EC Revenue	[redacted]	[redacted]
GC Revenue	[redacted]	[redacted]
HT Revenue	[redacted]	[redacted]
Without Open Access		
EC Revenue	[redacted]	[redacted]
Impact of Open Access		
EC Revenue	[redacted]	[redacted]
Other TOC Revenue	[redacted]	[redacted]
Share of OA revenue generated	29.9%	26.1%
Generated per £1 abstracted	42.6p	35.3p

Table 9 Breakdown of revenue sources from Open Access Operators – 2014/15, 2013/14 prices

Flow type	Modelled Generated	Modelled Abstracted	Unmodelled Abstracted	Unmodelled, Abstraction and/or Generation	Total
Flows with franchised direct services	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Flows with new direct services	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Indirect/ connecting flows	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Non-London flows	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Flows with franchised direct services	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Flows with new direct services	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Indirect/ connecting flows	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Non-London flows	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total	31.0%	72.8%	5.7%	-9.6%	

Incorrect market shares explain most of the gap between modelled and actual Grand Central revenue – our direct demand modelling has improved performance. However, Hull Trains performs worse, primarily because of under-performance compared to forecast at [redacted] ([redacted] less revenue than the model predicts there). We believe this could be because of a lower propensity to travel to London amongst residents of [redacted] than elsewhere in Yorkshire: the only data included in our model is the population and the distance from each station.

This means that "unmodelled abstraction and/or generation" is negative, rather than positive as it was in Table 7. This makes interpreting the results less straightforward. If the entire unmodelled difference in revenue is **not** abstracted, then 31% of open access operators' revenue has been generated (**45p in every £1**). If the difference was **not** generated, then 22% of open access operators' revenue has been generated (**27p in every £1**).

The equivalent numbers for the whole period (in 13/14 prices and without discounting) is between 18% and 31% (22p to 46p), and shown in Table 10 below.

Table 10 Breakdown of revenue sources from Open Access Operators – last fifteen years, undiscounted, 2013/14 prices

Flow type	Modelled Generated	Modelled Abstracted	Unmodelled Abstracted	Unmodelled, Abstraction and/or Generation	Total
Flows with franchised direct services	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Flows with new direct services	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Indirect/ connecting flows	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Non-London flows	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Flows with franchised direct services	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Flows with new direct services	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Indirect/ connecting flows	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Non-London flows	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total	31.5%	89.3%	-7.1%	-13.6%	

6.7 Sources of error

Sources of model error vary by flow. Table 11 below shows the size of this error on those flows where it exceeds £1m in absolute value (over the fifteen years, in 2013/14 prices):

Table 11 Modelling error for a selection of flows

Flow	2014/15 Error	Total Error (15 year)	Total Revenue (15 year)
Flow [redacted]	[redacted]	[redacted]	[redacted]
Flow [redacted]	[redacted]	[redacted]	[redacted]
Flow [redacted]	[redacted]	[redacted]	[redacted]
Flow [redacted]	[redacted]	[redacted]	[redacted]
Flow [redacted]	[redacted]	[redacted]	[redacted]
Flow [redacted]	[redacted]	[redacted]	[redacted]
Flow [redacted]	[redacted]	[redacted]	[redacted]
Flow [redacted]	[redacted]	[redacted]	[redacted]
Flow [redacted]	[redacted]	[redacted]	[redacted]
Flow [redacted]	[redacted]	[redacted]	[redacted]
Flow [redacted]	[redacted]	[redacted]	[redacted]
Flow [redacted]	[redacted]	[redacted]	[redacted]
Flow [redacted]	[redacted]	[redacted]	[redacted]
Flow [redacted]	[redacted]	[redacted]	[redacted]

It can be seen that the size of the error differs somewhat, and is of varying sign. Figure 16 shows the total (not just OAO) market size over time for the four largest 'errors'.

It can be seen that causes are likely to be varied. For instance, in Sunderland our model did not predict the demand decline that happened prior to the OAO service commencing, although it appears to estimate the growth fairly well. In [redacted], the model appears to correctly show demand stabilising after 2006/07 (when the market was, perhaps, 'mature') but incorrectly forecast the scale of growth resulting from the OAO service. In [redacted], the model forecasts demand well until 2005/06, but then there is growth in the following three years that is not modelled. In [redacted], there is model overpredicts for most of the period, perhaps reflecting local economic factors.

Of note, however, is the scales of the charts and the scale of the 'error'. [redacted] is far and away the largest contributor to the over-forecast of OAO revenue. It would appear from the historical data that our model over-forecasts the volume of generated demand in this market. As a result, **we consider that the actual share of generation in 2014/15 is to the lower end of the 0.27 - 0.45 range**.

Figure 16 Performance of model vs. actuals for the flows with the four largest 'errors'

[redacted]

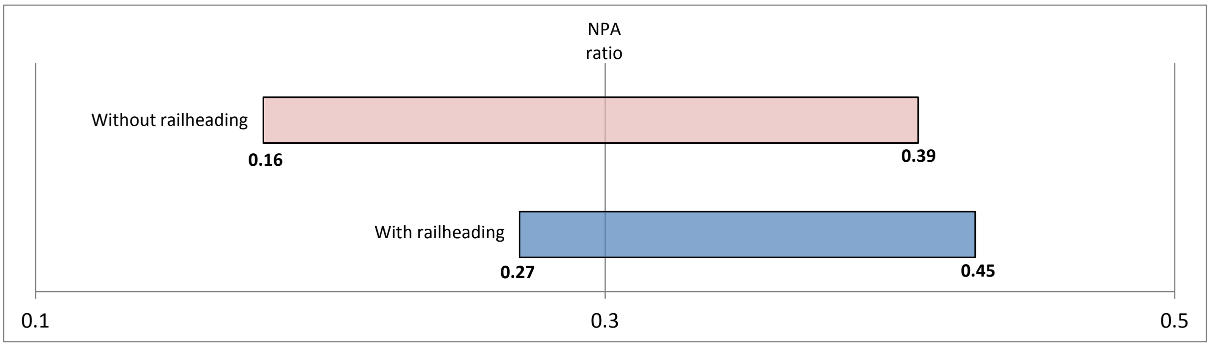
7 CONCLUSIONS & RECOMENDATIONS FOR FURTHER ANALYSIS

7.1 Conclusions

7.1.1 Abstraction – generation ratio

When railheading effects are included, we estimate that in 2014/15 between 22% and 31% of revenue had been generated – between 27p and 45p for every £1 of revenue abstracted. In the absence of railheading effects, our estimate was that 13.8% and 27.9% of revenue had been generated – between 16p and 39p for every £1 of revenue abstracted. These bounded estimates are shown in Figure 17.

Figure 17 Bounded estimates of generation-abstraction ratio in 2014/15 (centred on NPA ratio of 0.3)



The estimate is bounded due to a gap between modelled and actual revenue, the causes of which are not straightforward to identify. However, the actual value is likely to be towards the lower bound of the “with railheading” range.

7.1.2 Comparison to NPA test

The NPA ratio lies within both of our bounded estimates of generation, indicating that the existing open access operators have generated revenue close to the NPA ratio. Direct demand modelling allows the railheading effect to be captured, narrowing the bounded estimate. Where open access services would not serve stations with a previous lack of direct services to London, these bounded estimates may not be appropriate.

7.1.3 Limitations

The approach adopted to model fare market size is necessarily simplified, due to the large number of factors at play. It is possible that there are complex ticket-switching effects that are not captured by this study. Other aspects of fares modelling have also required simplifying assumptions. There were problems with data reliability, particularly miscategorisation of revenue by ticket type. We believe we have made sensible assumptions to mitigate these issues, but they should be borne in mind. We note also that our model performs reasonably well at forecasting the size of the total market until 2007/08.

7.1.4 Findings for future modelling

The direct demand modelling identifies generation/abstraction effects for a number of flows, correcting under/over-forecasting by the revenue model. However, findings in aggregate are less clear cut due to a small number of flows where the direct demand models predict demand that is at variance with actuals.

We believe that issues with direct demand models are due to difficulties calibrating a "propensity to travel" parameter for large catchment areas covering numerous stations. Variances in propensity to travel across the catchment area would be better reflected if existing rail customer location data were available as a model input.

Modelling of the effects of differential fares is crucial to identifying the abstractive-generative split of an open access service, producing modelled market shares that are closer to outturn market shares. A reasonable approach in modelling fares effects might be to assume track access applicants adopt a revenue-maximising fares strategy. However, this would require a high degree of confidence in the spread parameters used.

7.2 Recommendations for further analysis

There are limitations to the findings of this study due to lack of available data. Were this data available, the modelling could be refined.

- It has not been possible to model performance because average minutes lateness (AMLs) are not available from the start of the study period. PPM would be less useful for this purpose, but is also in any case unavailable.
- Crowding effects have not been captured. To do so would require detailed count data from 1999/2000 to the present day
- The suppressive/generative effects of car parking space availability have not been captured due to the non-availability of car park occupancy data.

Further examination of the PDFH forecast might help with understanding the reasons for deviation from outturn revenue. The PDFH model may be improved by using more disaggregated exogenous historical data. For example, [redacted] demand is currently modelled as driven by exogenous factors for the entire [redacted] region. There is no guarantee, however, that obtaining specific exogenous data [redacted] would necessarily improve fit.

LENNON data was acquired by product group, as this should in theory be enough to determine ticket type. In reality, we discovered revenue which we believe to have been misallocated. The extensive processing time involved in acquiring the data meant it was not possible within project timescales to request the data again. Thus, a more accurate breakdown of revenue by ticket type may be available. Product descriptions would, however, need to be manually assigned to product groups, and this would also be time consuming.

Further refinement of fares modelling might help in understanding why [redacted] is capturing higher market shares than modelled effects whereas [redacted] producing results closer to actuals.

Direct demand modelling may be better suited to smaller catchment areas. Where it is necessary to model large catchment areas, variance in propensity to travel across the catchment should be reflected. Existing rail users' locations (from TOC customer databases) would provide the best source of data to inform this.

Understanding of direct demand model calibration could be further enhanced by studying new services offered by an incumbent TOC, where customer data is available. A good example would be to model the effects of substantial service enhancements at Chester and Wilmslow from December 2008. This would pay dividends in improving estimates of abstraction and generation, as new Open Access operators planning to serve markets which are poorly served by rail at present and are likely to be particularly responsive to improved service levels would be assessed appropriately.

8 APPENDIX 1 – REVENUE MODEL DEVELOPMENT

8.1 Revenue model development

We have created a revenue forecasting model (Revenue Model) for three TOCs operating on the East Coast Mainline – the main franchised operator, InterCity East Coast, and the two open access operators, Hull Trains and Grand Central. The base year for this model is the year prior to commencement of operations of the first OAO, Hull Trains. The model forecasts expected demand and revenue using the PDFH v5.1 methodology. PDFH v4 is used for fares elasticities, per WebTAG guidance. The data sources for historical exogenous growth are shown in Table 12. These drivers of exogenous growth are those specified by the PDFH forecasting framework.

Table 12 Exogenous data sources

Exogenous Driver	Data Sources
GDP/Capita	CEBR Regional and City level historical data
Employment	CEBR Regional and City level historical data
Population	CEBR Regional and City level historical data
Car Journey Time	Extrapolation of data from Road Transport forecasts 2013, analysed on a flow by flow basis using Google Maps
Car Operating Costs	Analysis of multiple components of car operating costs, including: ▪ DECC historical fuel cost data ▪ Historical DfT fleet mix data (VEH0203) ▪ Historical DfT fuel efficiency data (ENV0103)
Car Ownership	Tempro V5.4 and v6.2 data
Bus Journey Time	Extrapolation of data from Road Transport forecasts 2013
Bus Costs	ONS RPI component data (series: DOCX)

The Revenue Model segmentation was created to give an appropriate level of disaggregation to model the three service introductions. Examples of four different types of flow were selected for each service introduction:

1. Flows with incumbent direct services to London;
2. Flows with no incumbent direct service to London;
3. Indirect/connecting passenger flows; and
4. Non-London flows.

The flow groups specified are given in Table 13. The role played by each of the flows for each OAO is listed in the table.

Table 13 Flow Groups, and Type of Flow (as defined above) for each operator

No.	Flow Group (to/from London unless otherwise stated)	Flow type	HT[8]	GC Sun[9]	GC Brad[10]
1	Grantham	Flow with incumbent direct London service	✔		
2	Retford	Flow with incumbent direct London service	✔		
3	Doncaster	Flow with incumbent direct London service	✔		✔
4	Selby, Howden	Flow with no/ltd. incumbent direct London service	✔		

[8] Hull Trains
[9] Grand Central Sunderland service
[10] Grand Central Bradford service

No.	Flow Group (to/from London unless otherwise stated)	Flow type	HT[8]	GC Sun[9]	GC Brad[10]
5	Brough	Flow with no/ltd. incumbent direct London service	✔		
6	Hull	Flow with no/ltd. incumbent direct London service	✔		
7	York	Flow with incumbent direct London service		✔	
8	Thirsk	Flow with no/ltd. incumbent direct London service		✔	
9	Northallerton	Flow with incumbent direct London service		✔	
10	Eaglescliffe	Flow with no/ltd. incumbent direct London service		✔	
11	Hartlepool	Flow with no/ltd. incumbent direct London service		✔	
12	Sunderland	Flow with no/ltd. incumbent direct London service		✔	
13	Pontefract	Flow with no/ltd. incumbent direct London service			✔
14	Wakefield	Flow with no/ltd. incumbent direct London service			✔
15	Brighouse, Mirfield	Flow with no/ltd. incumbent direct London service			✔
16	Halifax	Flow with no/ltd. incumbent direct London service			✔
17	Bradford	Flow with no/ltd. incumbent direct London service			✔
18	North Lincolnshire	Indirect/connecting flow	✔		
19	North East Yorkshire	Indirect/connecting flow	✔		
20	Sheffield	Flow with incumbent direct London service			
21	Teesside	Indirect/connecting flow		✔	
22	Newcastle	Flow with incumbent direct London service		✔	
23	West Riding	Indirect/connecting flow			✔
24	Huddersfield	Indirect/connecting flow			✔
25	Leeds	Flow with incumbent direct London service			✔
26	Doncaster-Hull	Non-London	✔		
27	Grantham - Hull	Non-London	✔		
28	Hartlepool - York	Non-London		✔	
29	Sunderland - York	Non-London		✔	
30	Northallerton - York	Non-London		✔	
31	Doncaster - Bradford	Non-London			✔
32	Bradford - H, B, M	Non-London			✔
33	Other - London	Indirect/connecting flow			
34	Other - Non-London - Core	Non-London			
35	Other - Non-London - Major	Non-London			
36	Other - Non-London - Other	Non-London			

8.1.1 The with OA expected demand scenario

The intention of this stage of the approach is to simulate Stage 1 of the ORR's five stage NPA test: *'using standard industry models (such as MOIRA and the passenger demand forecasting handbook'* to make a broad estimate of the likely level of abstraction', and to then compare this with actual OAO and East Coast revenues.

The "with OA" scenario is a revenue forecast in two dimensions. The model forecasts both market size and market share. Market size is driven by exogenous growth, timetable changes and changes in average fare. Market share is driven by timetable (from MOIRA) and differential fares effects.

Exogenous growth is consistent across the "with OA" and "without OA" scenarios. The source of historical exogenous factors is show in .

GJT effects are driven by MOIRA outputs. The study period is 16 years, and as such, with the twice yearly timetable changes there have been 32 timetable changes during this period. The majority of these timetable changes have not produced large changes in GJT, nor would it be practical to model every one of them. For this reason, we have selected those timetables that have resulted in material changes in train service provision and in levels of revenue and revenue apportionment, and included these in the model.

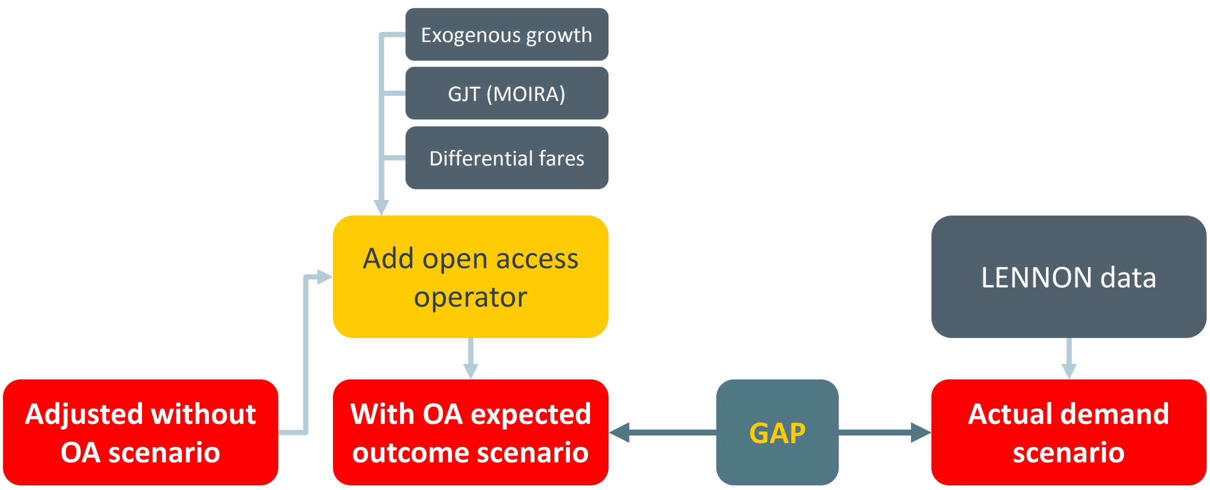
These timetables are as follows

- May 2000
- September 2000 – **HULL TRAINS COMMENCEMENT**
- June 2002
- September 2002
- December 2004
- June 2005
- December 2005
- December 2006
- May 2007– **GRAND CENTRAL SUNDERLAND COMMENCEMENT** (implemented from December 2007)
- December 2009
- May 2010 - **GRAND CENTRAL BRADFORD COMMENCEMENT**
- May 2011
- December 2011
- May 2013

These were selected on the basis of MOIRA's estimates of demand growth and market shares relative to a September 1999 base timetable on key East Coast Main Line flows.

The process used in the development of the "With OA" scenario is set out in Figure 18 below.

Figure 18 With OA Scenario



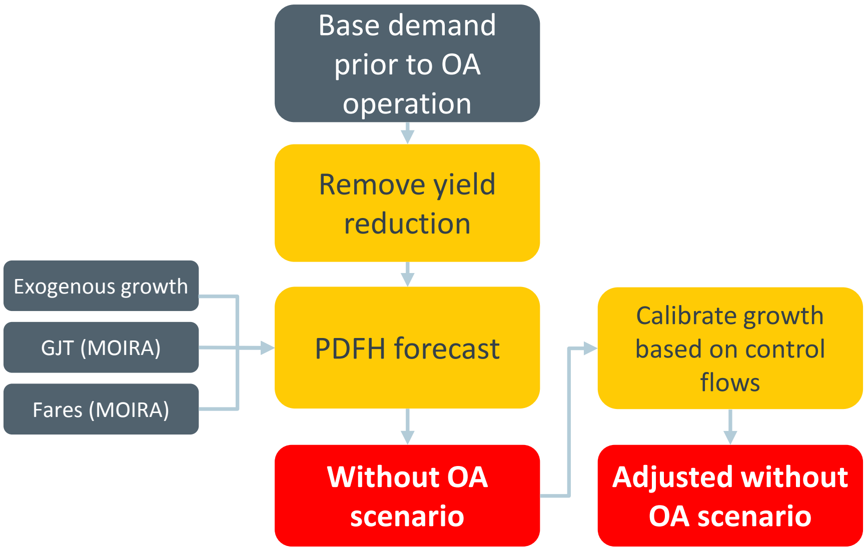
In each case, the MOIRA run uses on a consistent base demand matrix, ensuring that the effect modelled is solely based on the GJT change resulting from the timetable change. The outputs from each of MOIRA timetable changes listed above are processed to separate market growth (elasticity) effects from market share (ORCATS) effects.

To model the market size effects of differential fares on the market for travel for each flow, the average yield across operators is used. Differential fares modify the market shares output by MOIRA – the approach used to model this is described in 8.2.3.

8.1.2 The without OA expected demand scenario

Using the same model structure and flows as the "with open access (OA)" model, in the "without OA" version of the model, the open access operators are removed. All market share effects are removed, with 100% of revenue accruing to the ICEC franchise. Figure 19 shows the process involved in the modelling of this scenario assuming no competition from OAOs occurred.

Figure 19 Developing a counterfactual



To model the timetable GJT effects the same timetables are run in MOIRA, but with the open access services deleted.

Fares effects are modelled on the basis of ICEC historical yields. We have performed analysis seeking to identify a difference in yields on flows where ICEC faces competition, relative to those where it does not. We have found no evidence of ICEC franchisees having a consistent strategy of reducing yields on flows with competition, relative to those without. This finding does not represent conclusive evidence that there has not been a competitive response, merely that we have not identified conclusive evidence of one. As such, we have not incorporated an increase in yields in the counterfactual (without OA scenario) to account for a competitive response in the with OA scenario.

8.1.3 The comparator forecasting model

In order to identify the potential effects of fares competition on flows affected by OAO competition, a Comparator Flows Model has been developed that forecasts revenue over the same period as the OAOs have been in operation. This model takes account of the same factors as the Counter Factual Model (exogenous growth timetable effects, fares).

The flows the Comparator Model contains are (to/from London unless otherwise stated):

- Nottingham
- Derby
- Newark
- Chesterfield
- Leicester

- Leeds
- Grantham
- Newark - York
- Leicester - Sheffield
- Leicester - Nottingham
- Peterborough - Newcastle
- Peterborough - York
- Peterborough – Leeds

These flows are a combination of ICEC flows with no OA competition, and East Midlands Trains flows (also facing no competition). The number of major ICEC flows not facing competition from OAOs is limited. As such, we have also included flows from the adjacent intercity corridor, the Midland Main Line. We believe these are more appropriate for use as comparators than West Coast Main Line (WCML) flows, as the WCML experienced substantial disruption, then service improvements, as a result of the West Coast Route Modernisation programme.

We used MOIRA to calculate modelled timetable effects for each substantial timetable change.

The purpose of this model is twofold, firstly to identify the fares yield impacts in the absence of OAO competition on these flows. One impediment to the creation of a scenario in which the open access operator never existed is that it is difficult to know with certainty how the incumbent franchised operator's fares strategy would have developed in the absence of competition. The fares trajectories for the Comparator flows are used as to investigate whether fares should be changed in the "without OA" scenario. The outcome of this analysis is discussed in 8.1.2, above.

The second purpose is to attempt to account for the effect of potential PDFH forecasting error. This is likely to be particularly the case following the 2008 economic downturn, and this period forms the majority of the period covered in this study. A small study we recently undertook for DfT, and the Mott McDonald/University of Southampton study for PDFC, have shown under-forecasting of demand by the PDFH forecasting framework compared to outturn demand. In order to account for variation between outturn demand and our PDFH-based model counterfactual, we have developed a PDFH forecast adjustment based on the outturn forecast from the Comparator Flows Model and the forecasts it generates.

In producing PDFH forecasts and looking at the gap between observed and actuals, we noticed that on a year by year and a flow by flow basis, there were large variations. This could be explained by factors not considered in our forecasts, including:

- The impact of the Hatfield disaster;
- Introduction of Class 222s on the Midland Mainline;
- Redevelopment of King's Cross and St Pancras International stations;
- "Soft" changes in the customer service proposition, such as refurbishments, catering changes, and station improvements; and
- Revenue Protection initiatives.

With the exception of the Hatfield incident, our "control calibration overlay" has been applied only to calibrate for the large anomalies observed during and after the recession. It is well documented that PDFH has not performed well during this period.

8.2 Differential Fares

8.2.1 Investigating the use of MOIRA 2

In developing the methodology for this study, our original intention was to use MOIRA 2 to model the effects of differential fares in the modelled with OA scenario. Theoretically, using MOIRA 2 would have had a number of advantages. It would:

- take into account the relative infrequency of open access services (but also their lack of interchange penalty) by modelling demand train-by-train;
- obviate the need to develop a complex bespoke spreadsheet fares model, which would be a resource-heavy task; and
- provide information to ORR on the suitability of MOIRA 2 for this purpose.

In our exploration of the use of MOIRA 2, we encountered a number of issues that made its use impractical and/or not as useful for this task as was originally envisaged. These issues are as follows:

1. Fare level limitation

Although MOIRA 2 has the capability to model routed fares dynamically, this capability is grouped with time restrictions in the "Fares and Restrictions" function. This function is limited to modelling four fare levels, with the highest of these levels being an Anytime, Any Permitted ticket. In practice, two of the three remaining fare levels are already populated in the model for interurban flows (being Off Peak and Super Off Peak fares). This leaves only one remaining fare level for open access operator-specific tickets, which is insufficient.

2. Treatment of Advance fares

In MOIRA 2, Advance ticket demand does not change dynamically in response to changes in other fares or ticket availability. Instead, a fixed proportion of demand is assumed to relate to Advance tickets. Over the course of the study period, the use of Advance tickets increased substantially. Using MOIRA 2 would mean that this effect would not be captured. It is also not possible to adjust the yield of Advance tickets. Advance tickets are allocated to trains based on available remaining capacity, following allocation of walk-up passengers to trains. This utilises the crowding functionality of MOIRA 2, the methodology for which is not widely accepted within the industry.

3. Availability of historical timetables and revenue journeys matrices

To use MOIRA 2 to model fares effects alone, it would be necessary to hold all other factors constant, i.e. to have an identical contemporaneous timetable and base demand matrix prior to and after the modelled introduction of each new set of fares. In order to model the effect of changes to the fares mix, the pre-existing demand levels for each ticket type in each modelled year would be required, along with an extensive set of historical timetables. We contacted DeltaRail, who advised that historical revenue-journeys matrices could be supplied at a charge. Historical timetables could be supplied without charge, but would require significant processing time to import to MOIRA 2, and would require manual adjustment to reconcile older service codes to those recognised by MOIRA 2.

To deal with the issues above, we developed a methodology that would address the deficiencies in functionality. This would work as follows:

- The four fare levels would be used as follows: Anytime Any Permitted, Reduced Any Permitted, Anytime operator specific, Reduced operator specific;
- The base revenue journeys matrix (base demand) would have all Advance ticket demand re-allocated to Reduced. There would be no distinction between Off Peak, Super Off Peak and Advance ticket types; and

- An average fare would be calculated for each of the Reduced ticket types. This would be a weighted average of the Off Peak, Super Off Peak and Advance ticket types.

This approach would have significant drawbacks:

- The necessity to merge all Reduced tickets into one fare level removes granularity and accuracy. For example, the time validity elasticity of the Super Off Peak demand base would not be reflected, nor would switching between Reduced ticket types;
- It would not be possible to model the effects of two open access operators with competing fares. This would be a particular problem at Doncaster; and
- Merging Off Peak, Super off Peak and Advance tickets from the base revenue journeys matrix would require repeated export of base matrices [redacted] for processing, and re-imports to MOIRA 2. The processing time required for each of these (both within and outside MOIRA 2) would be prohibitive.

In consultation with the ORR project team, we concluded that the compromises that it would be necessary to make to use MOIRA 2 meant that it did not appear to have any significant advantages over other potential approaches, which are outlined below, particularly given the extensive processing time and additional cost involved. In addition MOIRA 2 is still yet to be recognised in the rail industry as an accepted tool for modelling fares and crowding changes, with development and testing work ongoing to produce an industry accepted version.

8.2.2 Alternative approaches

We assessed two alternative possible approaches – the use of the differential fares functionality in MOIRA 1, and the use of the PDFH-recommended formula for the impact of open access fares. Each of these is described below. We have ruled out a third possible approach (set out in PDFH 5.1 B11.3.2) where a lead operator's fares are undercut by a minor operator, on the basis that its scope for application is limited to very specific competitive circumstances that are not especially relevant to this study.

MOIRA 1 differential fares

MOIRA contains functionality to model the effect of introducing operator-specific tickets. We made enquiries to DeltaRail and received confirmation that this functionality uses the equation in section B11.3.1 - Minor Operator Tickets of PDFH 5.1.

Advantages

- Straightforward to implement; and
- Allows disaggregation between full and reduced ticket types (as far as would be possible with MOIRA 2).

Disadvantages

- The formula this approach uses is designed for cases where the new fare is introduced by a "minor operator", typically operating slower services; and
- The formula does not include interchange penalties, thus not capturing an element of the relative attractiveness of the open access service, where it is operating to stations that the franchised operator does not serve directly.

Because this formula is designed to be used where differential fares are introduced by a "minor operator", we do not consider it appropriate for use in the context of open access operators serving locations that the franchised operator does not serve directly.

PDFH Impacts of Open Access Fares – (B11.4 in PDFH 5.1)

PDFH contains a recommended equation for calculating the impact of open access fares. This formula is sourced from MVA reports for the ORR assessing historical track access applications.

Advantages

- The equation has been developed specifically for the purpose of assessing the impact of open access operator-specific fares;
- The spread parameter of [redacted] was calculated based on [redacted] data at [redacted] in [redacted]. In terms of service quality (rolling stock quality and onboard offer), the offer at [redacted] from [redacted] at the time was comparable to that operated by the franchised operator. Differences in service quality to other locations will thus not be captured by this formula, but this is in line with the methodology for the study as a whole; and
- The equation includes MOIRA-derived market share as a variable, thus capturing GJT (including interchange penalty) through the ORCATS allocation.

Disadvantages

- Requires use of the calibrated spread parameter used by MVA based on Grand Central data to York. We do not have details of the derivation of this spread parameter.
- The fare differential is entered as an absolute monetary value. Thus, the spread parameter derived based on York data may not be applicable to shorter-distance flows.

Because this approach is designed specifically for the purpose of modelling open access operators' differential fares, we adopted it for this study. Details of its use are set out below.

8.2.3 Adopted differential fares modelling approach

The approach to modelling the market share effects of open access operators' differential fares is described in section B11.4 of PDFH. The formula used is as follows:

[redacted]

Where:

[redacted] = Market Share from MOIRA
[redacted] = New Market Share
[redacted] = walk up differential
[redacted] = spread parameter

PDFH provides a default spread parameter of [redacted], but recommends calibrating this in certain circumstances. It should be added that this spread parameter must assume a particular price base for demand to be homogeneous of degree zero in prices (i.e. if all prices, incomes, etc. were increased 10% then the minor operator market share would be unchanged) so as prices rise, λ should fall proportionally.

A worked example of this approach is shown below:

Table 14 Differential fares market share calculation - worked example part 1

	Incumbent franchised operator	Open access operator
MOIRA market share (i.e. ORCATS distribution of "Any Permitted" ticket revenue)	[redacted]	[redacted]
Fare	[redacted] (Any Permitted route fare)	[redacted] (specified operator only fare)

Market share following introduction of new [redacted] fare on open access operator	[redacted]	[redacted]

To determine the yield effect of this change in market share, it is necessary to identify the split between the ticket types sold. In the above example, the “Any Permitted” ticket revenue will continue to be allocated between the operators according to the [redacted] ORCATS split. The calculation of the yield effect is shown below.

Table 15 Differential fares market share calculation - worked example part 2

	Incumbent franchised operator	Open access operator	Total
MOIRA market share (i.e. ORCATS distribution of “Any Permitted” ticket revenue)	[redacted]	[redacted]	[redacted]
Market share following introduction of new £[redacted] fare on open access operator	[redacted]	[redacted]	[redacted]
“Any Permitted” ticket market share (based on incumbent operator share reduction)	[redacted]	[redacted]	[redacted]
New £[redacted] fare market share	[redacted]	[redacted]	[redacted]
Average yield	[redacted]	[redacted]	[redacted]

On flows where both Hull Trains and Grand Central have tickets available, we have calculated the market share as follows:

- For each open access operator, apply the formula separately;
- Identify which of the two operators has a higher resulting market share; and
- Use this market share as the total captured by the open access operators, split according to the relative sizes of the formula-derived market shares.

The reasoning behind this approach is that each operator has a limited (and in most cases, similar) number of services per day. A Hull Trains-specific ticket does not allow travel on Grand Central (or vice versa) so it stands to reason that the combined share of the two operators would not equal double the share of one.

As described in section 5.3.3, there were significant issues with the reliability of the LENNON data. In some cases, this has resulted in average yields varying significantly from year to year, where through analysis of actual fares changes we are aware that fares did not change to the same extent in reality. In addition to these issues arising from e.g. miscoding of ticket type, movements in the availability and usage of discounts from year-to-year can also impact the average yield.

The quality of the time series fare data available to us (average yields from LENNON) has meant that in some cases we have had to use our subjective judgment and expertise to make manual adjustments to remove fluctuations which we do not consider realistic. Because these issues were more problematic for First Class ticket yields, Standard ticket yield differentials have been used to model market shares for both ticket classes.

8.2.4 Calibration of spread parameter

Because the fare differential is entered into the formula as an absolute monetary value, if just one spread parameter is used, the formula would likely overstate OAO market shares on longer distance flows with higher fares (and thus higher absolute differentials), and understate OAO shares on shorter distance flows with lower fares. To prevent this from occurring, we have calibrated the spread parameter based on outturn market shares at a number of locations served by OAOs. The locations selected are similar in qualities to [redacted], in that:

- They are stations that [redacted]
- Other than frequency, which is captured in the formula through the MOIRA market share, the service quality in terms of e.g. rolling stock is similar across operators;

In most cases, the spread parameter has been calibrated using 2008/09 data. At [redacted], it has been calibrated using 2011/12 data, as competition at [redacted] did not begin until 2010. In all cases, the price base of the yield data used to calibrate the formula is the same (2013/14).

We have identified that the differentials vary substantially between Full and Reduced ticket types, and as such we have calibrated separate spread parameters for each of these ticket types.

Table 16 Spread parameters

Distance band	Mileage	Calibrated at	Full	Reduced
1	0-150	[redacted]	0.038	0.090
2	150-165	[redacted]	0.034	0.080
3	165-180	[redacted]	0.023	0.070
4	180-205	[redacted]	0.016	0.057
5	205+	[redacted]	0.014	0.038

Note that because the formula is designed for application to walk-up fares, it has been used based only on walk-up ticket data. The market shares have then been applied to Advance revenue, in the absence of an alternative means of quantifying this split.

8.2.5 Modelling of market size effect

Overall changes in market size as a result of changes to average fare are modelled using actual average yields. Fares are modelled using PDFH 4.0 conditional elasticities. This does not fully take into account ticket type switching effects.

Analysis of average yields for Advance tickets indicates that Advance yields have risen throughout the study period. Applying a conditional elasticity to these rising fares would result in decreases in demand for this ticket type. In reality, Advance revenue has increased, due to increased availability of Advance tickets (for which quotas are controlled by TOCs). As such, we have merged Off Peak and Advance demand into a single Reduced ticket type, and modelled Reduced fares impacts based on the average yield for both ticket types.

Season ticket fares are assumed to have risen in line with the maximum allowed increase. For all ticket types, market size effects are based on standard class yields, due to the data issues described in section 5.3.3.

9 APPENDIX 2 – INVESTIGATION OF POTENTIAL EXPLANATORY FACTORS

9.1 On board offer

In May 2011, East Coast launched a new First Class service, which provided complimentary meals and drinks to First Class ticket holders. We have investigated whether this new on board offering has had an impact on the First Class market for Open Access Operators (OAO).

We have concluded that the enhanced First Class service has not had a significant impact on OAOs.

9.1.1 Methodology

We picked out from the Revenue Model three sets of London flow groups:

- Flows served by Hull Trains but not Grand Central;
- Flows served by the Sunderland Grand Central service; and,
- Flows served by the Bradford Grand Central service but not by Hull Trains.

We then looked at the share of revenue enjoyed by the OAO on these routes, for both First Class and Standard Class. We compared this data for both a period immediately before the introduction of the enhanced First Class offering (RY2011 P08 – RY2012 P01) and a period immediately after the introduction (RY2012 P03 – RY2013 P02). By looking at market share changes for both Standard Class and First Class, it is possible to distinguish between general service changes (for example, timetable changes), and changes which specifically affect First Class.

9.1.2 Results

Table 17 Hull Trains Flows Market Shares

	HT market share before FC improvement	HT market share after FC improvement	Difference
First Class	[redacted]	[redacted]	[redacted]
Standard Class	[redacted]	[redacted]	[redacted]

Table 18 Grand Central (Sunderland) Flows Market Shares

	GC market share before FC improvement	GC market share after FC improvement	Difference
First Class	[redacted]	[redacted]	[redacted]
Standard Class	[redacted]	[redacted]	[redacted]

Table 19 Grand Central (Bradford) Flows Market Shares

	GC market share before FC improvement	GC market share after FC improvement	Difference
First Class	[redacted]	[redacted]	[redacted]
Standard Class	[redacted]	[redacted]	[redacted]

It is clear from this analysis that the [redacted] share of the First Class market has decreased at a much faster rate than its share of the Standard Class market. However, for [redacted] services, there is very little difference between market share changes by class of travel.

An explanation for this can be found by looking at changes in the Revenue earned per passenger mile between the two periods.

Table 20 % Change in Revenue per Passenger Mile – After First Class introduction compared to Before

	[redacted]	[redacted]	[redacted]
OA Operator First Class	[redacted]	[redacted]	[redacted]
OA Operator Standard Class	[redacted]	[redacted]	[redacted]
EC First Class	[redacted]	[redacted]	[redacted]
EC Standard Class	[redacted]	[redacted]	[redacted]

9.1.3 Conclusions

It is clear that as well as introducing an enhanced First Class service, East Coast First Class yields [redacted]. This appears to have been part of a concerted effort to [redacted] [redacted]. The [redacted] in First Class ticket sales and resultant [redacted] is a well-documented phenomenon.

The data for [redacted] flows seem to be distorted [redacted]. However it is apparent that GC and HT have adopted different strategies during this period. [redacted] yields have risen, allowing EC to capture a proportion of the First Class market. [redacted] yields however have fallen, [redacted].

From this analysis, we can infer that

- the enhanced East Coast First Class offering has not had a significant impact on OAOs; and,
- changes to the share of the First Class market enjoyed by [redacted] can be attributed to yield effects.

9.2 Rolling stock quality

Table 21 below shows an approximate timeline of rolling stock provision on the East Coast Main Line by operator.

Table 21 East Coast Main Line rolling stock timeline

<table>
<tr><th></th><th>2000</th><th>2004</th><th>c.2004/5</th><th>2005</th><th>2008</th><th>2010</th></tr>
<tr><td colspan="3">EC Inter-City 225 ‘as built’</td><td colspan="4">GNER Inter-City 225 ‘Mallard’</td></tr>
<tr><td colspan="5">GNER refurbished HST</td><td colspan="2">NXEC refurbished HST</td></tr>
<tr><td></td><td>Anglia class 170 (near new)</td><td colspan="2">Hull Trains Class 170 (with small kitchen)</td><td>Hull Trains 222 ‘Pioneer’</td><td>FGW class 180 (as built)</td><td>Hull Trains 180 refurbished</td></tr>
<tr><td colspan="4"></td><td colspan="2">Grand Central refurbished HST</td><td>Class 180 (near as built)</td></tr>
</table>

There have been some improvements in rolling stock interior quality on the East Coast Mainline – as, indeed, on the rest of the network – in the last fifteen years.

We have reviewed the fleets, and consider the current fleets are of broadly comparable quality. We therefore do not think that it is appropriate to include in our model a specific attraction of open access operators associated with rolling stock quality. The revenue model overlay, discussed in 8.1.3, applied to capture PDFH under forecasts of revenue will, in part, capture the increase in rail market size on the East Coast Main line that is coincidental with improved rolling stock quality on other train companies, especially Midland Mainline and East Midlands Trains.

It could be suggested, given the timing of the rolling stock improvements, that the entry of open access operators caused the East Coast franchised TOC to improve the quality of its rolling stock. We cannot identify this from the historical data.

9.3 Lagged effects

9.3.1 Rationale

The rationale for assuming lagged impacts for timetable changes (e.g. the appearance of a new direct Open Access service to London) is two-fold.

Firstly, it takes time for people to become aware of the new service and so to use it. This is used (in PDFH) as a justification for asymmetric response: existing users experience a service deterioration quickly and might cease to travel, whereas new users are not exposed to service level to the same degree. Equivalent arguments could be applied in a ‘railheading’ situation, where it takes time for users of one station to learn of an improved service at another. The time taken for such a response can probably be reduced by marketing effort.

Secondly, there is a change in the optimal spatial locations of economic activity (e.g. people’s houses, people’s offices). An hourly direct service at Chester (for instance) would encourage firms to locate in Chester; someone changing jobs and needing to be in London one day each week would be more likely to live in Chester. These changes are lagged *not* because of a lack of information, but because of transition costs: for example a firm might review its location only when the lease on its office is up for renewal.

9.3.2 Approach

We have periodic LENNON data through to the end of rail year 2009, and from mid-2011 until mid-2013. We have selected seventeen stations to consider the profiles of growth of demand to/from London (and stations via London – the same segmentation as used in our Revenue Model).

Table 22 Stations (to/from London) uses in lag analysis

Darlington	Stations (potentially) affected by the new Grand Central Sunderland service
Eaglescliffe	
Hartlepool	
Middlesbrough	
Sunderland	
Thornaby	
Chester	Stations potentially affected by 'Virgin High Frequency' timetable from January 2009
Crewe	
Wilmslow	
Bradford	Stations potentially affected by the new Grand Central West Riding service
Brighouse	
Halifax	
Leeds	
Mirfield	
Pontefract	
[redacted]	Stations unlikely to be affected by major new services during this time period
[redacted]	

We have used [redacted] & [redacted] as controls to capture background underlying growth. For those periods where we do not have periodic data, we have assumed the periodic volume is the annual volume divided by thirteen.

We have not adjusted demand for other changes that might be suggested by PDFH (e.g. local income, local population, price elasticities) because it is not so clear how well these effects match reality, especially following 2008 – it is well known that PDFH under-forecasts rail passenger demand in this period, where economic Output growth was disappointingly weak, and this can also be seen in our Revenue Model.

9.3.3 Analysis

The following two charts show how each station has performed since 2006 (the first line, '2006P13' is the year ending 2006P13 i.e. rail year 2006 which is forced to zero for each station; the last line, '2015P13', shows rail year 2015) compared to the total at [redacted] plus the total at [redacted]. Passenger miles in LENNON on included TOCs have been converted to trips by dividing by the mileage from London, so some 'excess' growth will be abstracted (e.g. adding direct London-Chester trains mean some passenger miles will move from Arriva Wales, for whom we do not have LENNON data, to Virgin West Coast for whom we do, meaning the mileage in LENNON for each trip will increase).

The charts illustrate the methodology for controlling for the flows: we remove the proportionate change in annual journeys that would have resulted if the flows had grown with the same profile as the total growth across [redacted] *and* [redacted]. Thus, the excess growth for [redacted] and [redacted] together sum to zero.

The total number of additional journeys is shown, rather than the proportion, as some stations (e.g. [redacted] and [redacted]) start from a very low base, whereas the additional journeys indicates the scale compared to other stations.

The charts show that for the most part, demand to OAO destinations has continued to grow at a faster rate than [redacted] + [redacted], indicating a lagged response to the improved service that is ongoing. The effect at [redacted] and [redacted] is even more pronounced.

Figure 20 'Excess' Growth on Sunderland and VHF Routes

[redacted]

Figure 21 'Excess' Growth on West Riding Route

[redacted]

The following two tables show the same data numerically, to demonstrate the proportionate ramp-up profile. We have selected a base year for each station, and removed any growth above/below the [redacted] and [redacted]'average' performance prior to that rail year.

The first table shows the total number of journeys (which will differ slightly from the chart because of the removal of growth prior to the base year). The second table removes underlying ([redacted]+[redacted] proportional growth since the base year – converting excess growth to a constant demand base. The third table converts this to a PDFH-type lag profile.

Table 23 Added trips (compared to [redacted]& [redacted]performance)

Base year	13 periods ending same period as base year										growth
& period	Rail year	06	2007	2008	2009	2010	2011	2012	2013	2014	2015 15 vs 06
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	+0.04x
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	+47.71x
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	+1.20x
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	+0.34x
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	+0.03x
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	+96.52x
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	+0.43x
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	+3.52x
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	+19.58x
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	+0.28x
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-0.06x
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	+21.12x
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	+11.62x
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	+0.34x
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	+4.59x
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	+0.45x
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	+5.47x

Table 24 Removing underlying growth (so same demand base)

Base year	13 periods ending same period as base year											growth
& period	Rail year	06	2007	2008	2009	2010	2011	2012	2013	2014	2015	vs base
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-0.2x
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	+34x
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	+0.6x
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-0.0x
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-0.2x
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	+70x
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	+0.0x
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	+2.3x
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	+14x
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-0.1x
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-0.3x
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	+15x
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	+8.2x
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	-0.0x
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	+3.1x
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	+0.1x
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	+3.7x

Table 25 Percentage of 2015 excess journeys achieved

13 periods ending this rail year, same pd as base year		2009	2010	2011	2012	2013	2014	2015
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	100.0%
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	100.0%
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	100.0%

[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	100.0%
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	100.0%
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	100.0%
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	100.0%
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	100.0%
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	100.0%
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	100.0%
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	100.0%
PDFH 5.1 – 'major new service'	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**100%**

Note. some stations (e.g. [redacted], [redacted]) are not included in this table as it is not clear that there is a trend that we have captured. Also, we have assumed demand in the rail years from 2010 inclusive is split equally among periods. This is likely to delay the growth profile slightly.

9.3.4 Findings

The following findings are apparent from the analysis of rolling annual data:

- **There was significant growth at Chester and Wilmslow**. Both these stations had had an irregular direct London service, especially during the off-peak, until the implementation of 'Virgin High Frequency' changes in January 2009; they now have hourly off-peak direct trains to London. Unsurprisingly, given none of the stations off the ECML which gained direct trains are served so frequently, these two stations both demonstrate more 'excess' growth than any of the ECML stations.
- **Underlying growth is uncertain**. The second chart shows a lot of 'missing' demand at [redacted], seemingly in excess of what could be explained by abstraction to [redacted] etc. Growth in the last ten years at [redacted] has been 26% against [redacted] 34% and [redacted] 43%. This provides a general note of caution, given idiosyncratic local factors can influence demand at individual stations which aggregate data (e.g. GVA in Yorkshire) cannot.
- **Ramp-up seems to be ongoing**. At only one of the modelled stations does faster growth (or in some cases, like [redacted], slower growth) than 'average' appear to have stopped. Chester and Wilmslow together added [redacted] trips to/from London in rail year 2015, compared to if they had performed like [redacted] and [redacted].

The analysis is complicated somewhat by the steady improvement in service levels (Grand Central moving from three daily trains to five on the Sunderland route and three to four on the West Riding route – these may not qualify as 'major new services' however) so excess growth may continue to occur. This problem would have been even worse at Hull Trains stations, where there were significant enhancements to both speed and frequency in the first years of operations. Grand Central also suffered with poor performance in the early months of operating on both of its routes, which may have delayed the response further.

The most straightforward service enhancements have been at Chester and Wilmslow. The former shows faster 'ramp-up' than PDFH; ramp-up at the latter is somewhat slower. As our findings over other stations suggested lagged response is slower than the PDFH 'Major New Stations' recommendation, we have chosen to implement the demand ramp-up at Wilmslow in our model.

9.3.5 Quarterly trends

We have considered taking the same approach to identify quarterly trends. Unfortunately, year-on-year growth is much more variable when we consider only three periods than considering thirteen. The chart below shows year-on-year growth between Leeds and London (East Coast's largest single flow), compared again to the [redacted] & [redacted] base, for rolling three period and rolling thirteen period totals.

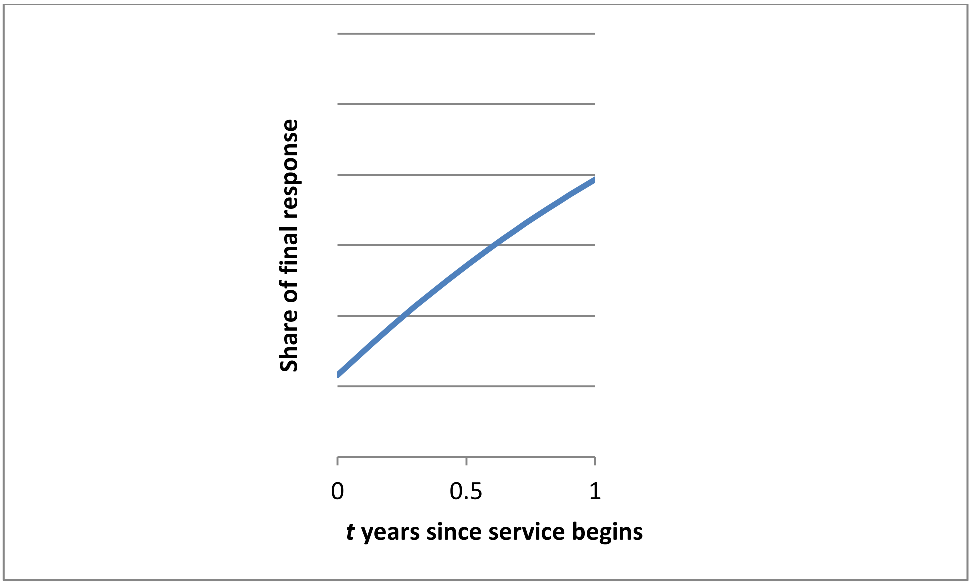
Figure 22 Leeds to London three v thirteen period growth

[redacted]

We control for "normal" growth (three periods on the same three periods the previous year), so seasonal trends *common across all stations* (e.g. the timing of Easter) should not appear. But nevertheless, there are clearly significant station-specific fluctuations, which may reflect engineering work or special events for example. In our view, it would not be realistic to expect to identify lagged responses at a station-level from quarterly data.

In order to derive a ramp profile for the forecasts used in this study we fitted a third order polynomial curve to calculate the ordinates of the lagged response curve for Wilmslow – one of two stations which should not be contaminated by further service changes. The lagged curve is shown below (with standard errors shown below the estimates) and the graph below. Note that the response we identified for the first year is the area *below* the curve (i.e. between t=0 and t=1).

[redacted]



We use this only to derive impacts for the first three quarters (which are included in PDFH) – these are clearly subject to significant uncertainty (as they are not estimated, but simply a consequence of the *assumed* functional form of the lagged effect). The quarterly impacts are calculated as the area under the curve. The impacts are as follows:

Table 26 Estimated lag profile for use in OAO forecasts

	Average Q1	Average Q2	Average Q3	Average Q4	Average Y1	Average Y2	Average Y3	Average Y4	Average Y5
Major new services	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

9.3.6 Lags in ticket choice

In addition to timetable-related lags, a potential further channel for lagged impacts is in TOC choice. This would apply to the first change only – it might take time for passengers to use different websites to buy their tickets or to make a different selection on the ticket machine because they are unfamiliar with the level of service offered by the new TOC.

One would expect this time lag to be less significant now, because of the importance of (TOC-specific) Advance tickets means that passengers are likely to be offered several train-price combinations (involving different train companies) and 'mix-and-match' franchised and open access tickets. Further, it is difficult to use actual data to identify this, however, because of the key confounding factor of the availability/price of Advance tickets. The availability/price of Advance tickets could easily move passengers to and from interavailable products even if there is no change in the train service.

We have periodic data for Rail Years 2008 and 2009, which covers the first sixteen months of Grand Central's Sunderland route (it started at the beginning of 2008 period 10), and below show the periodic market shares at four stations affected by the new service.

Figure 23 Market Shares of tickets by "route" at four Sunderland route stations

[redacted]

Table 27 'Grand Central only' share of Passenger Miles by period, station to/from London

	2008 P09	2008 P10	2008 P11	2008 P12	2008 P13	2009 P03	2009 P06	2009 P09	2009 P10
Eaglescliffe GC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Hartlepool GC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Sunderland GC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
York GC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

There are fluctuations each period, so conclusions are necessarily tentative; in addition, the Grand Central only share in Period 10 is low, probably reflecting poor performance in the early months. However, at the two stations which hadn't previously been served (Eaglescliffe and Hartlepool) the long term market share is reached very quickly; at the other two stations there is speedy ramp-up with the long term market share reached within the first quarter. [redacted] appears to be an exception, although may have been affected by changing availability/prices of Advance tickets.

9.3.7 Recommendations

We have reviewed the speed of demand responses using periodic LENNON data, and identified lagged responses that appear to be somewhat slower than recommended by PDFH. We have applied these in our modelling. We would hesitate in recommending these for other studies, as they are based on a single station only (with another two stations as controls), but using the single station with a significant improvement in London service that occurred at one point in time only and sufficient demand that growth is smooth.

We cannot recommend introducing lags in the up-take of operator-specific tickets, however – adjustment seems very rapid based on our findings, although it is difficult to use actual data (as we have done) given the potential confounding factors of advance ticket availability and price specific to these flows.

Introducing lags in response would make Open Access Operations seem more abstractive, because it takes longer to generate demand (and revenue) to the railway system then it does to abstract it – indeed, where passenger use any permitted tickets then the abstraction is instantaneous, and given ticket retailing is impartial, the changes in market shares of TOC-specific tickets is rapid.

However, the longer the Open Access Operation is assumed to continue, the less significant the 'ramp-up' period becomes. Generation at Wilmslow was about [redacted][11] in the first five years compared to a counterfactual in which it were instantaneous; ramp-up in abstraction would have been very close to [redacted]. Considering ten years, generation would be [redacted], and so on. Although Track Access Applications are time-limited, existing operations do not appear to be so. It would need to be considered carefully whether it is appropriate to include time lags in appraising applications, because their impact can be reduced by modelling the operation as continuing for a longer period.

9.4 Station quality

As with rolling stock quality, the revenue model overlay, discussed in 8.1.3, applied to capture PDFH under forecasts of revenue in our Revenue Model will, in part, capture the increase in rail market size on the East Coast Main line that is coincidental with improved station quality on other routes, especially Midland Mainline and East Midlands Trains.

Grand Central supplied us with information on five station improvements which it had part-funded.

Table 28 Summary of station schemes on Grand Central route

Station	Scheme (summary)	Funding	Date
Eaglescliffe	Additional 68 car parking spaces. Other enhancements.	[redacted]	Summer 2013 – Summer 2015
Wakefield Kirkgate	Restoration of station building. Provision of retail outlet. Creation of office accommodation. Improvements to subway waiting facilities. Additional 50 car park spaces.	[redacted]	Sep 2013 – July 2015
Mirfield	Installation of CIS on Platforms 1 & 2. Community-based mural on underpass.	[redacted] from Grand Central Other funding partners: West Yorkshire Combined Authority, Northern Rail	Nov 2014 – Spring 2015
Northallerton	Additional 90 car parking spaces.	[redacted]	Aug 2012 – Spring 2013
Brighouse	Provision of platform planters.	[redacted]	Summer 2014

These station enhancements could generate rail demand, particularly on other train companies (at none of these stations is Grand Central the primary operator). However, the impact of these schemes is likely to be relatively modest, and difficult to quantify (especially for the car parking schemes, where we do not have the necessary data to identify how much, if it all, demand had been suppressed prior to the additional spaces being added).

[11] Assuming no demand growth in the interim. Because there is exogenous demand growth over the five years, this percentage should be slightly nearer one.

We have MOIRA data that could incorporate the impact of the additional car parking spaces at Northallerton, with the results shown in Figure 24 below.

Figure 24 TransPennine Express Journeys (C) at five stations. Allocations based on Dec 14 Wednesday timetable. Source: MOIRA

[redacted]

It can be seen that there was [redacted] growth at Northallerton once the car park scheme was complete. However, it is not clear that the result of this would be the expanded car park. The timing of the demand growth is not entirely consistent with this – there is no visible demand change in the first six months after the scheme was complete. Some demand may have been abstracted from other stations – there is lower growth at [redacted] and [redacted]. It is also not obvious how much of any impact could be attributed to Grand Central, as opposed to TransPennine Express itself.

We would hesitate in recommending that station quality impacts be considered an important component of Open Access Operator generation – station schemes could just as equally be ameliorating the impacts of Open Access Operations.

This is particularly true for car parking schemes, where additional spaces funded by Open Access Operators can be relieving a constraint on demand on their own operations, and the impact on other TOCs is simply relaxing a constraint caused by Open Access Operations – if 90 car parking spaces at Northallerton are occupied by Grand Central passengers, then the impact on TransPennine of adding the spaces is simply to remove a demand constraint that would not exist were it not for Grand Central.

We thus do not include any station quality impacts, or an externality of these on other Train Companies, in our modelling work.

9.5 NRPS as an indicator of service quality

9.5.1 Introduction

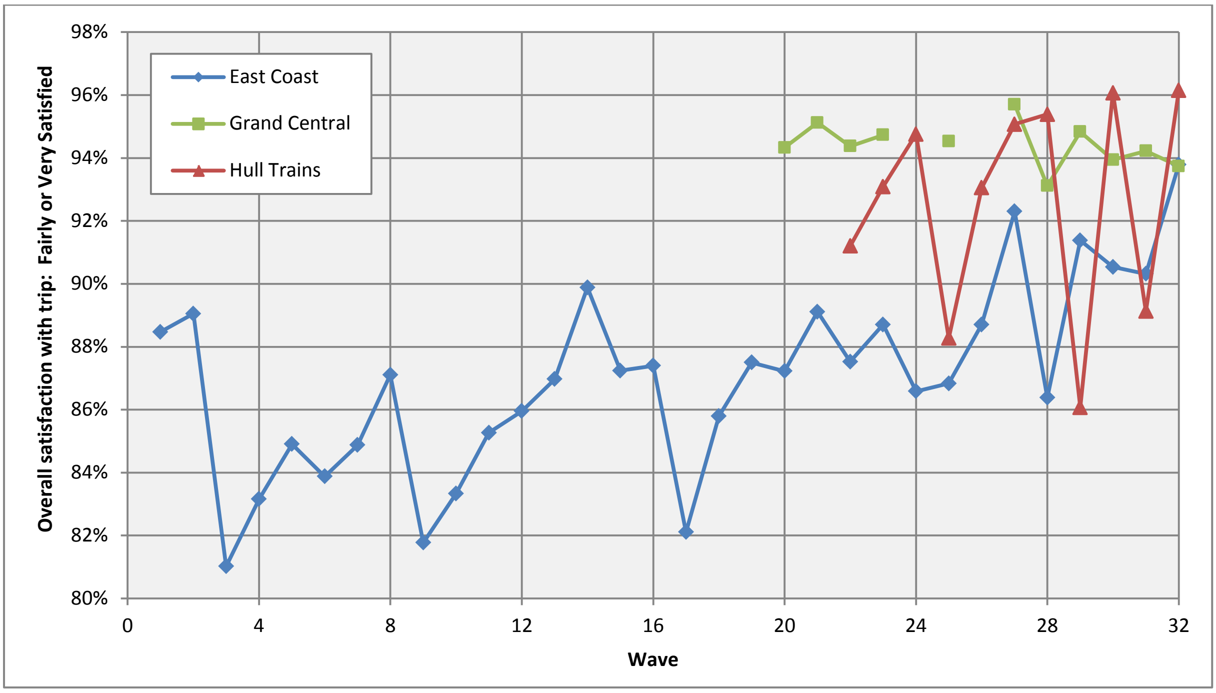
We initially planned to include a number of supplementary demand drivers in our model to reflect the other demand drivers included in PDFH, e.g. CCTV. This was discussed at the stakeholder workshop, where we requested from the OAOs the possibility of them supplying data on the levels of coverage of these attributes (e.g. the dates of introduction of CCTV and the proportion of trains covered).

Coverage of these factors is necessarily limited. At the stakeholder workshop, it was suggested and accepted that satisfaction levels from the National Rail Passenger Survey (NRPS) would be a more complete and appropriate measure, incorporating all 'soft factors', including those not included in PDFH, and explaining demand growth better.

Use of NRPS satisfaction levels is complicated by the lack of previous research linking passenger satisfaction to demand levels (Accent produced a report, *Customer Satisfaction and Customer Demand,* for PDFC in October 2010 which suggested in most sectors there was no identified link between satisfaction and demand), and the potential for overlap between measures of passenger satisfaction and the demand drivers already quantified in PDFH (e.g. fares and crowding).

Transport Focus helped us by supplying a full set of NRPS survey responses for East Coast, Grand Central and Hull Trains since the first wave (Autumn 1999) until the most recent (wave 32 in Spring 2015) – Grand Central and Hull Trains participate at their own expense but have not done so for every wave. Figure 25 below shows the headline overall satisfaction for the TOCs.

Figure 25 Overall satisfaction for the three TOCs. Source: LF analysis of NRPS data supplied by Transport Focus



It is easy to see that satisfaction with the open access operators has been higher than with East Coast and its predecessors for most of the period with data coverage.

9.5.2 Analysis

We know from previous Transport Focus work (e.g. Rail passengers' priorities for improvements, October 2014) that certain key drivers of satisfaction are:

- Enough room to sit or stand;
- Value for money; and
- Experience of delays.

These broadly map to factors recognised as key demand drivers in PDFH:

- Crowding (chapter B6);
- Fares (chapter B3); and
- Punctuality/Reliability (chapter B5).

While NRPS does collect data on whether not a passenger's journey was delayed, it only collects data on satisfaction with crowding and with value for money, not on load factors[12] or the price paid for the journey.

An example cross-tabulation is shown in Table 29 below, which shows overall satisfaction conditional on the passenger's views of the room available on the train, during the Spring 2015 NRPS wave. It should be noted that the reported NRPS satisfaction with *sufficient room for all the passengers to sit/stand* cannot be recovered from this table, as some passengers did not answer the question asking their overall satisfaction with the trip.

Table 29 Overall satisfaction given 'sufficient room', Spring 2015 NRPS wave

Operator	Share of passengers very or fairly satisfied overall, given *Rating of train...Sufficient room for all the passengers to sit/stand.* Shown as % (sample size)		
	All respondents (including no answer)	Very good/fairly good	Neither good nor poor/ fairly poor/very poor
East Coast	93.8% (1078)	96.3% (833)	80.3% (188)
Grand Central	93.7% (577)	96.3% (449)	66.8% (52)
Hull Trains	96.1% (559)	98.1% (484)	79.3% (56)
All of the above	93.9% (2159)	96.4% (1766)	79.8% (296)

It can be seen that the difference in overall satisfaction between operators changes when we control for whether or not there was sufficient room to sit or stand (and, indeed, the difference between operators is no longer statistically significant).

We undertook a multivariate analysis of NRPS data to understand how all of the factors affected passenger satisfaction, to try and understand the source of variation in customer satisfaction. This is an extension of more straightforward cross-tabulation, potentially controlling for many variables.

We used 'logit' choice analysis (common in transport modelling), to understand the causes of customer satisfaction. Our modelling is undertaken in BIOGEME, a common software package for choice analysis. The results are shown in below. The probability of meeting the binary dependent variable (e.g. overall satisfaction) is $\frac{1}{1+e^{-\text{variables}}}$, as can be seen from model 1 (in the table below) which replicates the NRPS results (e.g. the probability of overall satisfaction for a Hull Trains passenger is $\frac{1}{1+e^{-2.68-0.53}}$ =96.1%). In the models with constants for each wave (the constants are not reported) the reference wave is Spring 2015, but the NRPS results cannot be replicated using these tables.

Table 30 Outputs from models (coefficients;* p*-values in brackets). Source: LF Analysis of NRPS data supplied by Transport Focus

Model	1	2	3	4	5	6	7	8	9
Dependent variable (Overall Satisfaction…)	very/ fairly	very/ fairly	very/ fairly	very/ fairly	very/ fairly	very/ fairly	very/ fairly	very	very/ fairly
Constant	2.68 (0.00)	2.65 (0.00)	−0.08 (0.66)	0.382 (0.03)	−0.142 (0.50)	−0.074 (0.71)	−0.181 (0.11)	−2.85 (0.00)	−0.407 (0.00)
Operator = GC	−0.006 (0.98)	0.664 (0.00)	−0.174 (0.02)	0.069 (0.32)	−0.307 (0.00)	−0.188 (0.01)	−0.170 (0.01)	−0.140 (0.00)	−0.128 (0.04)
Operator = HT	0.530 (0.04)	0.404 (0.00)	−0.082 (0.21)	−0.089 (0.17)	−0.104 (0.32)	−0.074 (0.23)	0.074 (0.27)	0.064 (0.00)	−0.048 (0.43)
Room to sit/stand – very good/fairly good	–	–	1.42 (0.00)	1.65 (0.00)	1.42 (0.00)	1.42 (0.00)	1.42 (0.00)	1.19 (0.00)	1.42 (0.00)
Journey not delayed	–	–	1.63 (0.00)	1.64 (0.00)	1.66 (0.00)	1.62 (0.00)	1.57 (0.00)	1.20 (0.00)	1.63 (0.00)

12 The NRPS dataset seems to include enough data to match passengers to trains on routes with relatively infrequent services, and with a complete database of train loads (e.g. from loadweigh systems or passenger counts for each train) it would be possible to calculate the impact on satisfaction of changes in load factors and/or crowding multipliers.

Very/fairly satisfied with VFM	–	–	1.08 (0.00)	–	1.25 (0.00)	1.08 (0.00)	1.11 (0.00)	0.743 (0.00)	1.09 (0.00)
Very satisfied with VFM	–	–	0.67 (0.00)	–	0.633 (0.00)	0.722 (0.00)	0.814 (0.00)	1.20 (0.00)	0.661 (0.00)
Constant for each wave	–	✔	✔	✔	✔	✔	✔	✔	–
Weights	TF	TF, wrt waves	TF, wrt waves	TF, wrt waves	TF, wrt waves	TF, inc. between	No weights	TF, wrt waves	TF, wrt waves
Waves included	32	20-32	20-32	20-32	27-32	20-32	20-32	20-32	20-32
Rho-squared	0.660	0.520	0.621	0.597	0.657	0.620	0.660	0.191	0.619
n	2,054	26,804	26,804	26,804	13,306	26,804	26,804	26,804	26,804

n.b. in models 2 and subsequently we include *only* the portion of the NRPS sample who answer the questions about overall satisfaction, whether there was enough room to sit/stand, whether their journey was delayed *and* whether they were satisfied with value for money. If we did not make this restriction then there would be a larger sample shown for models two and four.

The first two columns show, as Figure 25 made clear, that there is a statistically significant difference in overall satisfaction between TOCs without controls – model 1 shows the results for Spring 2015 (wave 32) only, and model 2 shows the results for the NRPS period since the Open Access TOCs began to be surveyed (in wave 20, Spring 2009).

In our preferred model 3, however, we control for whether *the journey was delayed*, whether there was *enough room to sit or stand* and the passenger's level of satisfaction *with value for money*. We interpret this as controlling for the levels of crowding and the fare paid. It can be seen that each of these factors have a significant impact on the level of overall satisfaction. It also changes the TOC-specific impacts. Hull Trains passengers are not significantly more likely to be satisfied than East Coast passengers, and Grand Central passengers are *less* likely to be satisfied than East Coast passengers.

Models 4 through 9 test the robustness of these findings. Model 4 does not control for satisfaction with value for money, as it might be argued that TOC-specific 'soft factors' mean that, for a given level of fare, open access passengers would be more satisfied. In this specification, Grand Central passengers are no longer less satisfied than East Coast passengers, but neither OAO's passengers are more satisfied.

Model 5 considers only the last three years of NRPS surveys; models 6 and 7 use different sets of weights (in the base specification we use Transport Focus' own weights but adjusted so that the sum of weights matches the number of observations *within each wave*[13], in model 6 we use Transport Focus' own weights adjusted to match the number of observations across waves 20-32). Model 8 tests the determinants of whether passengers are *very* satisfied overall, and model 9 removes the wave-specific constants. None of these tests have impacts on our qualitative findings.

9.5.3 Conclusion

The NRPS surveys provide a rich resource which we have used to undertake a piece of analysis on passenger satisfaction on East Coast passengers survey data. The headline NRPS 'overall satisfaction' measure shows that the Open Access Operators have more satisfied passengers. We have identified that, **when we control for satisfaction with value for money, crowding and delays, Open Access Operators' passengers are no more satisfied than East Coast's.**

Our model already controls for the demand effects of changes in yield, which should correspond to changed perceptions of value for money. We have not been provided data on levels of crowding or delays.

[13] It is possible to attach weights to observations in BIOGEME, but it is necessary that the sum of the weights equals the number of observations (which would not necessarily be the case in a sub-sample of NRPS data)

We do not think that we need to include a specific overlay in our model – given satisfaction is explained by factors with known demand effects, it is not necessary to try to estimate a separate demand impact for passenger satisfaction itself.

9.5.4 Reliability and performance

To assess delays, PPM is too aggregate a measure – for instance, if most of the delays on East Coast are in Scotland, then Leeds-London passengers will have a different "PPM" – and in any case, Open Access NPS respondents are much *less* likely to report being delayed, despite PPM being, if anything, *higher* for East Coast.

This may be because Open Access Operator passengers are delayed because of cancellations (both have slightly higher CaSL (Cancellations & Significant Lateness) levels than East Coast) and then use the incumbent TOC and so are not captured as Open Access passengers in NRPS.

Figure 26 PPM (MAA) for ECML TOCs. Source: ORR

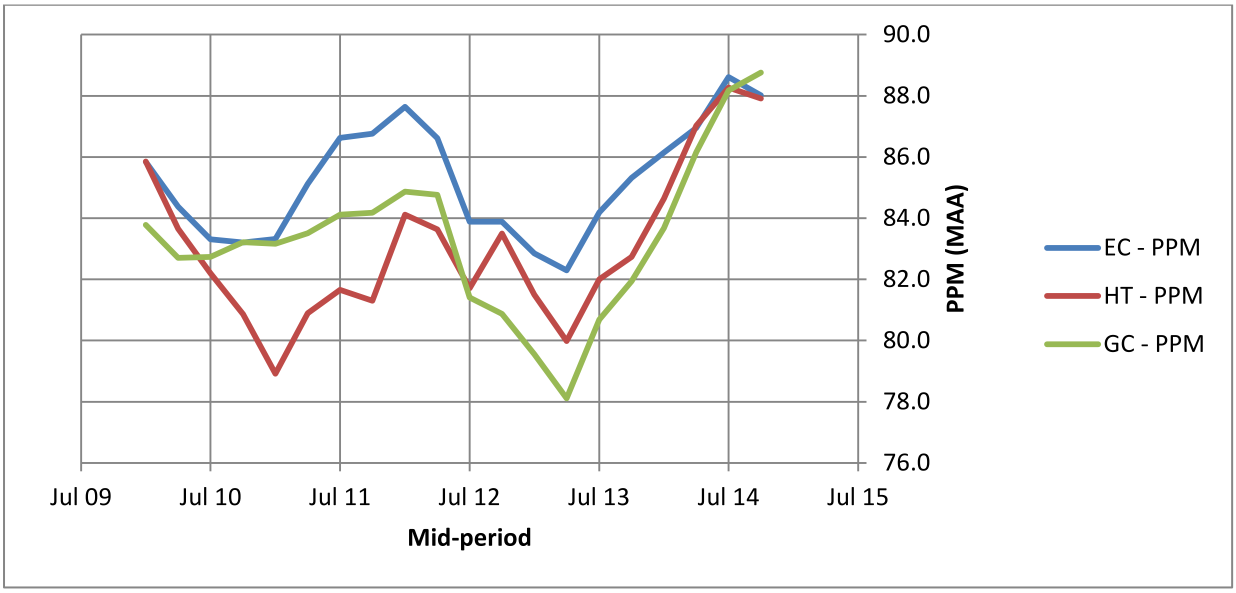
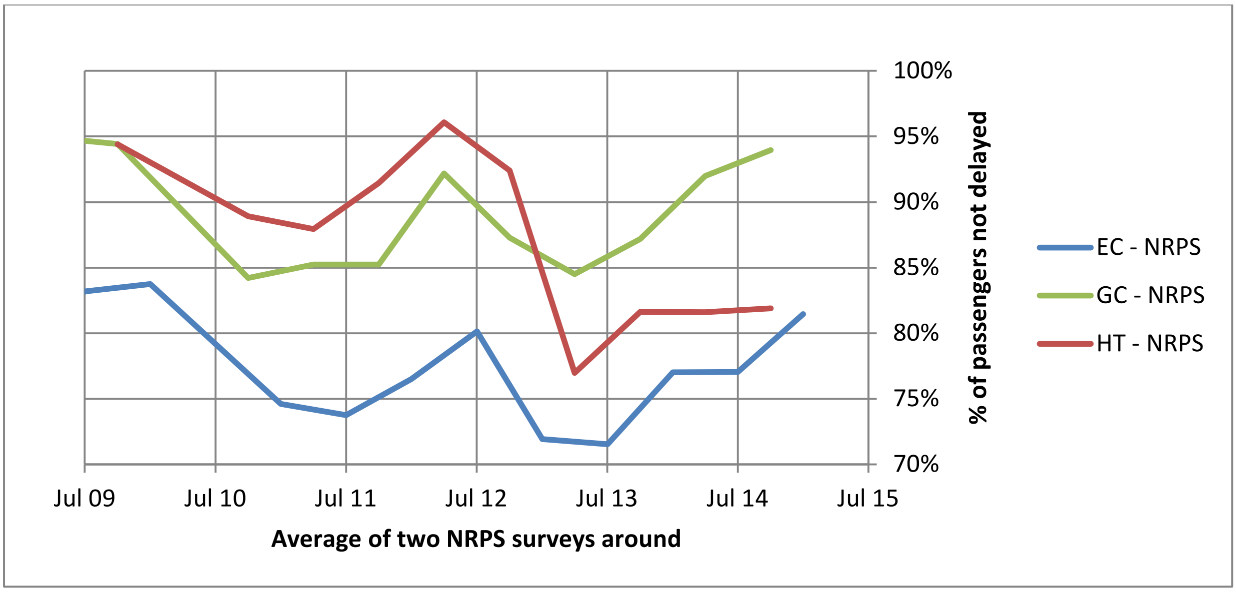


Figure 27 Share of passengers reporting as not being delayed. Source: LF analysis of NRPS data



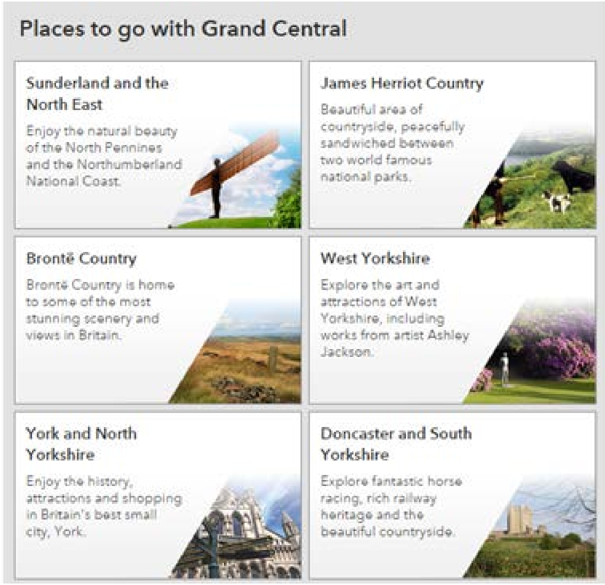
Leigh|Fisher

9.6 Marketing/branding effectiveness

We have investigated the marketing strategies of both Grand Central and Hull Trains. To quantify the effects of their strategy, it would be necessary to have access to both companies historical spend and Return on Investment (ROI) data. As this is highly commercially sensitive information, it is not possible to undertake this quantification. We also do not have access to historical marketing material. However, we have undertaken desktop research into the current marketing at these companies, and searched for evidence of “abstractive marketing”. This can be defined as an attempt to abstract revenue through marketing campaigns involving direct comparison with a competitor.

Our investigations show that the majority of OAO marketing activity appears to be aimed primarily at stations without a direct East Coast service to London and targets new rail demand.

Figure 28 Generative marketing material of OAOs



A key theme is destination marketing targeting leisure travel and tourist activity. Both Hull Trains and Grand Central provide extensive destination guides on their websites.

We have found no evidence of campaigns explicitly targeting the East Coast existing customer base. A key theme for both companies is the promotion of their low fares which could encourage downtrading to the OAO as well as encouraging new rail travel.

Figure 29 OAO marketing of low fares

In this study, we have not established the scale of marketing promotional spend by the OAOs, nor the impact on passenger demand of such marketing activity. It is likely that the marketing activity has played some role in increasing awareness of and generating demand for open access services. We have no evidence of the extent to which such marketing and promotional activity has been generative or abstractive.

10 APPENDIX 3 – DIRECT DEMAND MODELLING

10.1 Approach

Where open access operators provide new, direct services to stations previously poorly served the result is similar to that of an entirely new station. In such cases, GJT and fares elasticities alone do not capture the full impact of the new service. In order to capture the impact of improved station accessibility, we created ‘combined demand and accessibility choice’ based on a PDFH recommended approach (PDFH v4, section B6). This has been used to calculate how many trips are abstracted from other stations and how many new trips are generated due to the proximity and convenience of a new service offering for people in the vicinity of the open access station.

The approach is to create a series of population catchment models which determine trip rates from a particular zone to each station of interest, based on the generalised cost of making a journey via that station. The generalised cost used to determine an individual’s choice of station includes the journey time and cost of accessing the station, and the journey time, frequency and cost of the rail journey.

Three models have been created, two covering the Grand Central catchment area (Sunderland and Bradford, separately) and Hull Trains. The area of each model is shown in Figure 30 below. Each of the three models is calibrated to a ‘pre-open access’ market position, from which the ‘post open access’ impacts of a direct service and lower fares can be analysed. For the purposes of this analysis only journeys to London are modelled.

Figure 30 Direct Demand Model Maps

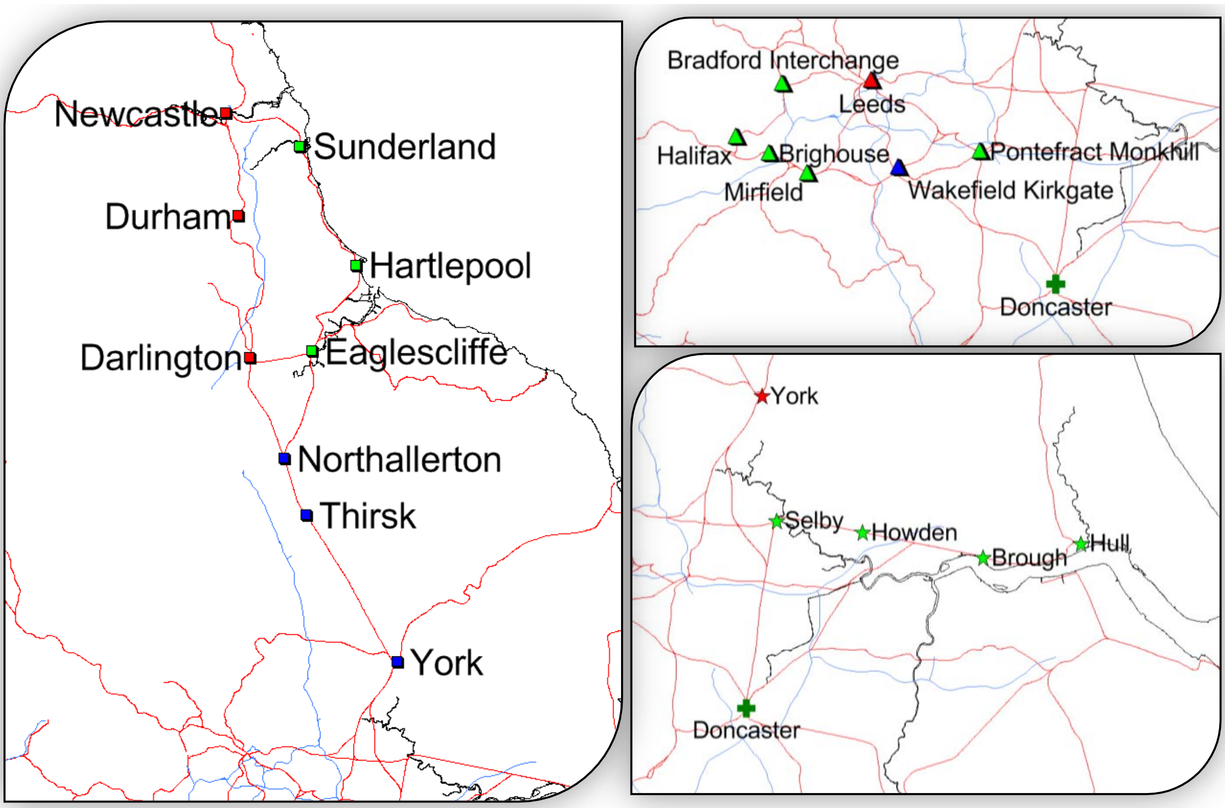
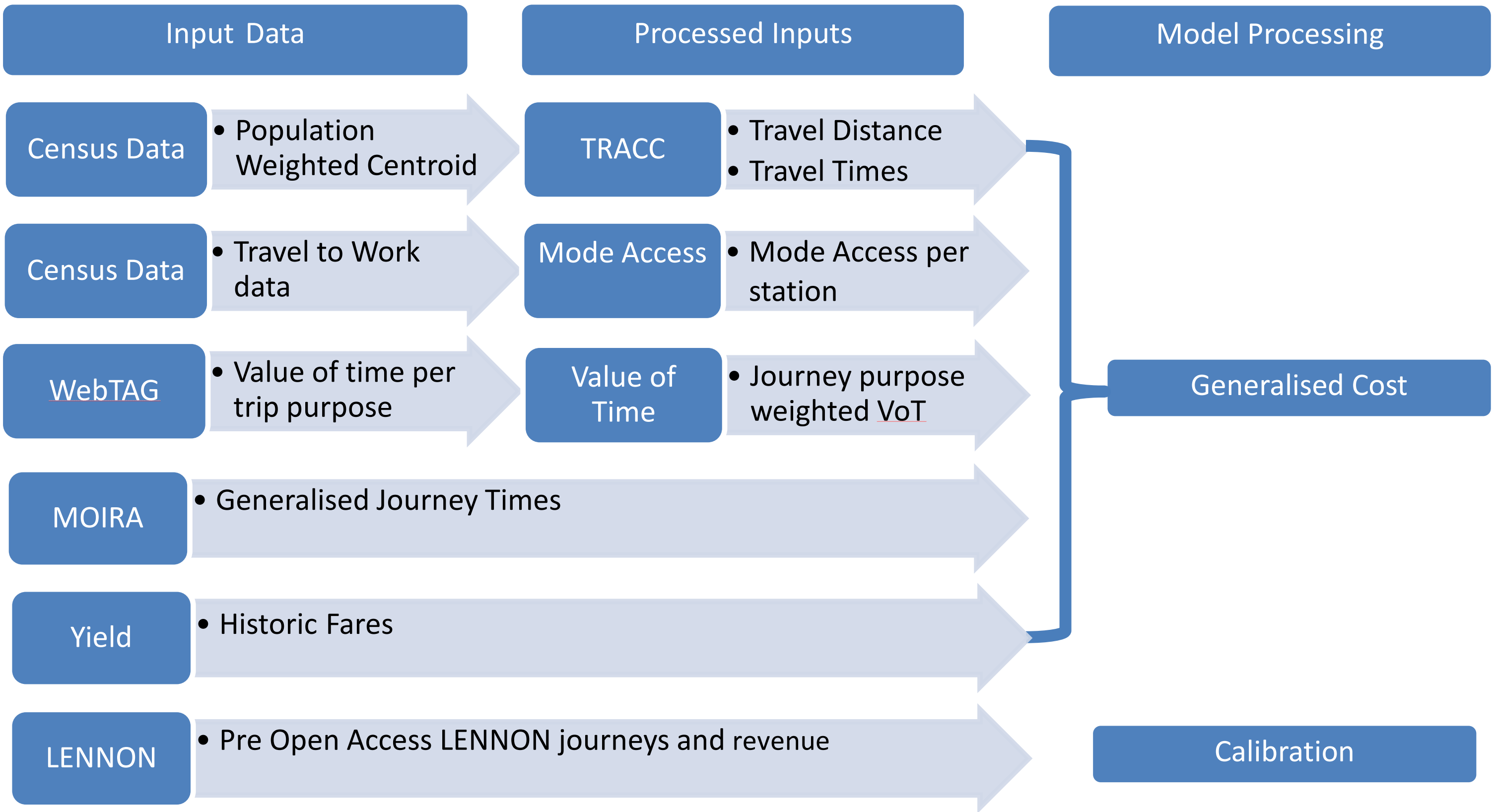


Figure 31 Direct Demand Model Flow



10.2 Processed Inputs

10.2.1 Catchment Area Population

The catchment areas of the origin stations were split into zones based on Middle Layer Super Output Area (MSOA) boundaries as defined by the ONS in the 2011 census. The location of the population weighted centroid of each zone was used as the point for calculating journey times and costs to new and existing stations.

We defined the potential population catchment area for each station using 2005 NRTS data. We used GIS mapping software to map the MSOAs from which passengers use each station. We could view the catchment areas for all stations in a model in one map as shown in the example below for the Sunderland model. We then selected additional zones to include where it looked sensible even if they weren't included in the NRTS data. It should be noted that the NRTS station origin data did not determine demand at each station, only how large the catchment area should be.

Figure 32 GIS Station Catchment Map of Sunderland

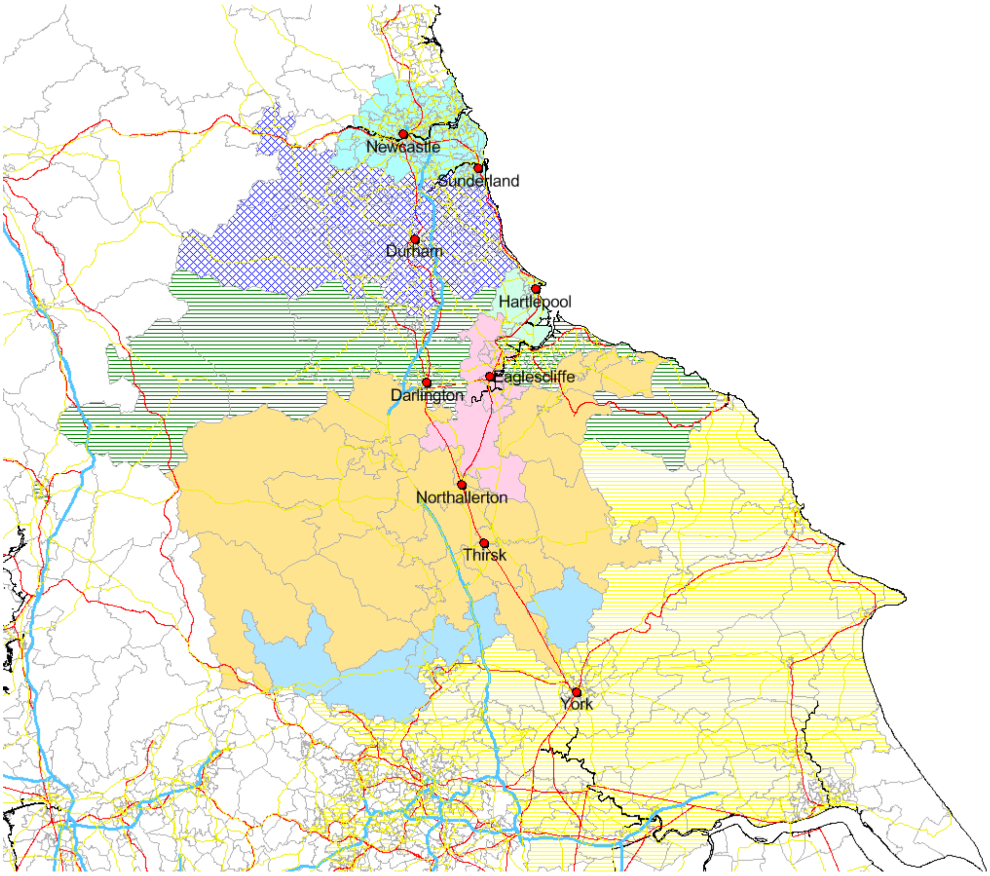


Table 31 Direct Demand Model Zones and Populations

Model	Number of zones	Population in catchment area
Hull	622	4,765,401
Sunderland	505	3,869,000
Bradford	574	4,335,939

10.2.2 Station Access Costs

The average station access costs were calculated from each zone using the average mode share and the generalised cost of each mode for accessing the station.

We used the population weighted centroid of each zone to calculate car driving and public transport access times to each station. This was done using TRACC software, whereby the coordinates of each origin zone centroid and destination station is input to obtain travel times and distances.

Using these travel distances we could calculate the generalised cost for each zone to station pair for a range of different access modes. The generalised cost of the station access leg of the journey includes the cost, and the travel time, which is converted into a cost using the value of time from WebTAG A 1.3.1. The WebTAG recommended values are for different journey purposes. We calculated a weighted average for each modelled area, using NRTS journey purpose data.

The journey time assumptions for bus, taxi, car driver and car passenger modes come from the TRACC analysis outputs. For bus fares, we analysed the maximum distance and fare of bus routes in the North East of England and created a representative value of £0.186 fare per bus kilometre travelled, with a minimum fare of £0.80. For long bus journeys above 3 hours the bus journey time cannot be calculated in TRACC. In these cases we approximate using the distance and an average minutes per kilometre calculated from the rest of the data set of bus journey times.

For car and van drivers, we apply a vehicle operating cost of 12 pence per kilometre, based on the2014 WebTAG recommended value including fuel and non-fuel costs. All prices in the models are converted to 2010 prices to be consistent with the 2010 WebTAG value of time. Car parking costs are also applied. We analysed a range of parking opportunities within 2km of each station, in order to calculate an average daily car parking charge (using www.nationalrail.co.uk and www.parkopedia.co.uk).

Table 32 Station Parking Fees

Sunderland Model	Parking fee	Hull Model	Parking fee	Bradford Model	Parking fee
Darlington	£5.77	Hull	£7.67	Bradford Interchange	£5.31
Durham	£8.58	Howden	£0	Brighouse	£3.51
Eaglescliffe	£2	Selby	£4.04	Mirfield	£0
Hartlepool	£0.46	Retford	£5	Pontefract Monkhill	£0
Middlesbrough	£6.45	Leeds	£15.73	Leeds	£15.73
Newcastle	£8	Brough	£4.6	Sheffield	£14.79
Northallerton	£4.6	Doncaster	£8.72	Doncaster	£8.72
York	£10.55	York	£10.55	Halifax	£5.16
Sunderland	£8.44			Wakefield Kirkgate	£1.08
Thirsk	£3.7				
Scarborough	£6.01				

We used www.yourtaximeter.com to find a range of fares and calculated an average fare per kilometre of £1.54.

Walking and cycling modes only have a cost in terms of journey time, calculated using the following assumptions of average speed:

- Walking, 4.8 km per hour, based on the walking speed assumption used in the TRACC software;
- Cycling, 21.5 km per hour, based on a reported average speed of 23 kmph for men and 20 kmph for women.

We used Census Travel To Work data to calculate a mode share for station access from each zone. The data had to be cleaned to remove implausible travel behaviour, for example when respondents stated they walk to work to a destination many kilometres away. For zone to station pairs with an insufficient census sample size we assumed the average mode share for that local authority.

Access time factors were applied to each mode as shown in the table below, taken from PDFH v5.1 Table B10.2. These are applied to the station access modes to reflect the relative convenience for passengers of each mode.

Table 33 PDFH Access Time Weightings

Mode	Access Time Factor
On Foot (walking)	[redacted]
Bicycle	[redacted]
Bus	[redacted]
Taxi	[redacted]
Car- Park and Ride	[redacted]

Mode	Access Time Factor
Car- Kiss and Ride	[redacted]
General	[redacted]

A maximum station access distance cap of 60km is assumed. A high station access cost of £10,000 is applied to zone and station pairs with a driving distance greater than 60km, as long as there is another station closer to the zone. This cap is applied to prevent the logit model assigning passengers to drive to stations which are a particularly long distance from their home as this would not be realistic of passengers' behaviour.

10.2.3 Rail Journey Costs

The generalised cost for the rail part of the journey consists of the Generalised Journey Time and fare from each station to London. GJTs are taken from MOIRA using timetables before and after the Open Access services commenced as shown in the table below. Where the level of open access service improved over the years, the Post Open Access scenario uses the GJTs for a year following the introduction of the full operation timetable. As the routes analysed are non-commuter distances and flows, we used the MOIRA GJTs corresponding to reduced ticket types.

All journey times are converted to cost in the model using the 2010 Value of Time.

Table 34 Source years for GJT, fares and base demand

Model	Pre-Open Access Year	Post-Open Access Year
Hull	2000	2007
Sunderland	2006	2013
Bradford	2009	2013

Historical fares were obtained from http://www.farehistory.info/. The fare assumption for the pre-open access scenario was an Any Permitted off peak return fare. The assumption for post-open access was a blend of the Any Permitted and Open Access only fare, weighted by the proportion using each operator (incumbent and open access) on each particular flow based on 2012 LENNON data. Fare assumptions were based on 2008 for Sunderland, 2009 for Hull and 2010 for Bradford, depending on the availability of data. All were uplifted to 2010 prices for consistency with the Value of Time.

10.2.4 Base Demand Data

For the Sunderland and Bradford models, we have calibrated the model using LENNON data for the pre and post open access demand and revenue on flows to London. We have used MOIRA historical journeys and revenue in the Hull model, as the LENNON data does not predate Hull Trains beginning operations.

10.3 Model Parameters and Processes

PDFH recommended parameters are used to determine the share of demand at each station from each zone, based on the distribution of the population, access times and cost from each zone to the stations, and time and fare of the rail journey.

Where:

LeighFisher

[redacted]

The generalised cost [redacted] destination station J is [redacted] car operating costs, bus fare, taxi fare and journey times for station access and overall travel time.

[redacted] is the ‘Utility of not using rail’. This value is used to calculate the proportion of the population who travel by train, which decreases as rail flow distances increases.

[redacted]

F_j is the attractiveness of a destination and is a function of population, and drive times raised to an egress elasticity. PDFH 4.1 (B6.10) gives a list of F_j values that can be used, depending on the destination of travel. PDFH states the F_j value for London, given the relativity of size and attractiveness of the city is 1,996,947.

Table 35 PDFH Recommended Attractiveness of London

J (Destination)	F_j (Attractiveness)
London	1,996,947

As passengers that live in the zone ‘a’, nearest a station ‘i’, can potentially use other origin stations ‘k’ to travel by rail to the destination station ‘j’, the population of that zone ‘P_{ai}’ is weighted by the proportion of the rail journeys that would use the alternative station. This is the first term of the traffic volume above. This is then weighted by the proportion of the population who would travel by rail and by the attractiveness of the destination station (F_j).

The volume of journeys from a zone using each origin station is proportional to [redacted], such that there will be an increase of trips from a given population to a station with increased services when there is fewer competing stations nearby. There will also be more trips when the generalised cost for the total journey is lower. This gives logical order that those living closer to the station with increased service levels will participate in more rail trips than those living further away from a rail station.

Table 36 PDFH Recommended Parameters

Parameter	Explanation	Value

[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]

10.3.1 Calibration

The models are calibrated to the pre open access base demand.

The table above shows the PDFH recommended values for the main parameters in the model. However, θ is changed in order to calibrate in terms of the share of the population making rail trips, or the 'propensity to travel'. This parameter is calibrated such that the total modelled journeys from all stations in the model equals the total LENNON journeys from all of the stations to London.

Table 37 Modelled Theta Values for Calibration

Model	θ
Hull	0.469
Sunderland	0.453
Bradford	0.444

It is then necessary to ensure that the model allocates share of demand to each station in a way that closely resembles the actual base demand data. In some cases it is necessary to add an additional station penalty (positive or negative) to the generalised costs for choosing a particular station, in order to achieve this station-level calibration. For example, in the case of Eaglescliffe station in the Grand Central Sunderland model, based on the inputs of generalised costs of accessing Eaglesliffe from the various population zones and the cost of the rail journey to London, the model allocates much more demand to this station than was reflected in reality before the commencement of Grand Central services. Therefore it is necessary to add a station penalty of [redacted] to costs of journeys via Eaglecliffe in order calibrate to close to the actual base demand share.

10.3.2 Impact of new services

After the model has been calibrated, we apply the Post Open Access GJTs and fares and assess the resulting change in demand from each station. This effect includes the impact of improved journey times and lower fares from open access stations relative to other stations within the modelled area. However it does not include the impact of exogenous growth, other timetable and service changes, or other factors. Therefore the modelled post open access demand at each station cannot be expected to match the Lennon actuals exactly. However it is compared against actual data for sense checking and model validation purposes.

In the case of incumbent east coast stations, the direct demand model gives a reduction in demand as passengers switch from travelling from areas previously not served by direct services to travel from a nearby open access served station. These reductions were not seen in the actual lennon demand data as there was other exogenous growth and timetable and competition effects. However the direct demand model isolates this impact of new open access services on demand generation and station switching, with all else held equal.

It is necessary to separate the impact of improved accessibility to previously unserved areas on station switching and new generated demand so that this effect can then be applied in the revenue model, over and above the demand growth forecast by exogenous growth and other factors. The elasticity effects of the change in fare and the change in GJT alone are calculated for each station served by a new Open Access service. This element of the change in demand is then excluded from the overall increase in demand. For incumbent stations which do not experience a change in GJT or fare in the model, any change in demand forecast by the Direct Demand model is inherently arising from this station accessibility effect only.

10.4 Results

The tables below show the spread of demand from each local authority in the main catchment area to the stations, before and after they are served by open access, as forecast in the model.

10.4.1 Sunderland Direct Demand Model

The tables below shows the modelled allocation of base demand between stations from the population in the north east local authority areas previously unserved by direct services to London.

Figure 33 Sunderland Model Local Authorities Map

Table 38 Hartlepool Local Authority Demand Between Stations

Hartlepool	Hartlepool	Eaglescliffe	Newcastle	Durham	Darlington
Before Open Access	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
After Open Access	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table 39 Sunderland Local Authority Demand Between Stations

Sunderland	Sunderland	Hartlepool	Newcastle	Durham	Darlington
Before Open Access	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
After Open Access	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table 40 Stockton-on-Tees Local Authority Demand Between Stations

Stockton-On-Tees	Eaglescliffe	Hartlepool	Northallerton	Darlington	Durham
Before Open Access	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
After Open Access	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Prior to Grand Central operating, demand from Hartlepool station was relatively low with passengers railheading to Darlington and Durham, with share from these stations falling post open access. Some population areas within the Hartlepool local authority are equally close to Eaglescliffe station and so this station receives an increased share of demand after Grand Central services commence.

From the Sunderland Local Authority area, the proportion of demand using Sunderland station increases from [redacted] to [redacted], whilst the proportion of demand railheading to Newcastle and Durham falls significantly.

Before Open Access, only [redacted] of the demand from the area around Eaglescliffe station (Stockton-on-Tees) would travel from Eaglecliffe, instead railheading to an East Coast station. With the additional direct services, [redacted] of passengers choose to travel from Eaglescliffe and Darlington share declines from [redacted] to [redacted].

Bradford Direct Demand Model

Figure 34 Bradford Model Local Authorities Map

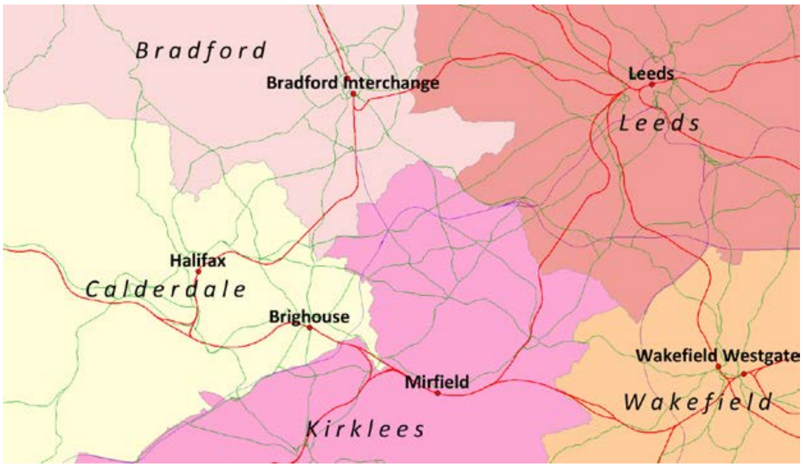


Table 41 Calderdale Demand Between Stations

Calderdale	Leeds	Huddersfield	Brighouse	Halifax	Wakefield
Before Open Access	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
After Open Access	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

The table below shows the demand impacts for Calderdale local authority, which hosts both Halifax and Brighouse stations. Whilst the model predicts relatively small growth in the proportion of demand from Brighouse, most of the demand within Calderdale is closer to Halifax, which sees a 12 percentage point increase. The proportion of demand travelling from inside Calderdale to Leeds to start a journey decreases from [redacted] to [redacted]. Within this model a large amount of demand is abstracted from Leeds.

Hull Direct Demand Model

Table 42 Kingston Upon Hull Demand Between Stations

Kingston Upon Hull	Hull	Brough	Howden	York

Before Open Access	[redacted]	[redacted]	[redacted]	[redacted]
After Open Access	[redacted]	[redacted]	[redacted]	[redacted]
Change in journeys	[redacted]	[redacted]	[redacted]	[redacted]

Table 43 East Riding of Yorkshire Demand Between Stations

East Riding of Yorkshire	Hull	Brough	Howden	York	Doncaster
Before Open Access	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
After Open Access	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Change in journeys	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Prior to the operation of Hull Trains, passengers from the East Riding of Yorkshire area would have been most likely to travel to London by railheading to York. With the introduction of Hull Trains services the proportion of demand using Hull, Brough and Howden increases and demand from York and Doncaster decreases. Brough demand increases in absolute terms from both local authorities but as a share of total demand only from East Riding of Yorkshire.

Leigh|Fisher

11 APPENDIX 4 – FARES METHODOLOGY SENSITIVITY TESTS

11.1 Introduction

As described in section 5.2.1, the modelling and analysis undertaken in this report utilises PDFH 5.1 elasticities, with the exception of the effect of fares modelling. The market size effects of changes to fares are modelled using PDFH 4.0 elasticities. This is in accordance with WebTAG guidance.

PDFH 4.0 recommends an overall fares elasticity of [redacted] for long distance flows to and from London. Theoretically, this should mean [redacted] [redacted]. PDFH 5.1 [redacted] recommends an overall fares elasticity of [redacted] for such flows. This would indicate that overall reductions to fares would [redacted].

A sample set of fares elasticities is shown in table Table 44.

Table 44: Rest of Country to/from London fares elasticities (101–200 miles shown for PDFH 4.0)

	PDFH 5.1	PDFH 4.0
Overall	[redacted]	[redacted]
Full	[redacted]	[redacted]
Reduced	[redacted]	[redacted]
Seasons	[redacted]	[redacted]

*[redacted]

We have undertaken a set of sensitivity tests on the treatment of fares changes within our model.

11.1.1 Sensitivity test 1 – PDFH 5.1 elasticities

This sensitivity test replaces the PDFH 4.0 Full, Reduced and Seasons fares elasticities used in the base methodology with PDFH 5.1 fares elasticities. This is the primary sensitivity test, and is reported on in the most detail below.

11.1.2 Sensitivity tests 2, 3 and 4 – fares modelling methodology

As discussed in 5.2.1, a full ticket type switching model has not been developed for this study. In the absence of such a model, there are different ways in which the elasticities could be applied. In the base methodology (and in sensitivity test 1), we have applied the elasticities to the yield change for each individual ticket type separately. In sensitivity tests 2, 3 and 4, alternative approaches to the application of these elasticities are tested. The purpose of these tests is to establish whether any of these approaches might produce results that differ significantly from that in the base methodology.

Sensitivity test 2: PDFH 4.0 ticket type-specific conditional elasticities applied to average yield across both Full and Reduced ticket types

Sensitivity test 3: PDFH 4.0 overall elasticity applied to average yield across both Full and Reduced ticket types

Sensitivity test 4: PDFH 5.1 overall elasticity applied to average yield across both Full and Reduced ticket types

Note that in sensitivities 2, 3 and 4, the effects of changes to season ticket fares continue to be modelled separately, as we consider this to be largely a discrete market.

11.2 Results

The results of the sensitivity tests, along with the original modelled results, and outturn revenue, are shown in Table 45.

Table 45: Summary of Sensitivity test results

				PDFH4.0 ticket type (TT) specific conditional elasticities applied to TT-specific yield changes	PDFH5.1 overall elasticity applied to TT-specific yield changes	PDFH4.0 TT-specific conditional elasticities applied to overall average yield change by flow	PDFH4.0 overall elasticity applied to overall average yield change by flow	PDFH5.1 overall elasticity applied to overall average yield change by flow
£millions, 2013/14 prices			**Actual**	**Base Method**	**1**	**2**	**3**	**4**
2015 EC Revenue			[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
2015 HT Revenue			[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
2015 GC Revenue			[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
15-year OAO Revenue (£m)			[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Of which…	Modelled Generated			[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Modelled Abstracted			[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Unmodelled Abstracted			[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Unmodelled, Abs or Gen			[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Generated per £1 abstracted	15-year	Lower		22p	26p	31p	29p	27p
		Upper		46p	50p	53p	52p	50p
	2014/15	Lower		27p	26p	32p	31p	29p
		Upper		45p	44p	44p	46p	45p

Figure 35: Total revenue across all TOCs in study

[redacted]

Because the change in real fares in the 'do minimum' scenario is fairly modest, and because the change in real fares caused by open access entry is also fairly modest, there is little difference in the outputs between the four sensitivity tests and the 'base' modelling.

At forecasting total revenue received by the ECML TOCs, the base method appears to perform slightly better between 2007 and 2009, and slightly worse than Sensitivity test 1 from 2012.

However, the improvement in total forecasting performance is not really focused on the worst-performing flows: for instance, sensitivity test 1 overforecasts "Other to/from London" revenue by £28m in 2014/15, whereas in the "base" revenue on this flow was overforecast by only £18m. There do not appear to be strong reasons for preferring any of the sensitivities over the base case.

Figure 36: TOC % market shares – actual, sensitivity test 1 and base methodology

[redacted]

Figure 37: Excess of actual revenue over modelled for a selection of flows, for sensitivity test 1 and base methodology

[redacted]

Figure 37 above shows that, for some London flows, sensitivity test 1 improves the accuracy of the model forecasts (the difference between actual and modelled revenue), but on other London flows, the performance of sensitivity 1 in accurately modelling revenue is worse. Figure 38 below shows the equivalent chart including all the sensitivities – the results are broadly similar.

Figure 38: Actual minus modelled revenue for four sensitivities and the base

[redacted]

Table 46: Summary of impact modelled as sensitivity 1

	2014/15 forecast volumes and 2013/14 prices	Total volumes and 2013/14 prices
With Open Access		
EC Revenue	[redacted]	[redacted]
GC Revenue	[redacted]	[redacted]
HT Revenue	[redacted]	[redacted]
Without Open Access		
EC Revenue	[redacted]	[redacted]
Impact of Open Access		
EC Revenue	[redacted]	[redacted]
Other TOC Revenue	[redacted]	[redacted]
Share of OA revenue generated	29.3%	27.8%
Generated per £1 abstracted	41.5p	38.6p

Table 47: Sensitivity 1, sources of actual 2014/15 OA revenue (2013/14 prices)

Flow type	Modelled Generated	Modelled Abstracted	Unmodelled Abstracted	Unmodelled, Abstraction and/or Generation	Total
Flows with franchised direct services	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Flows with new direct services	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Indirect/ connecting flows	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Non-London flows	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Flows with franchised direct services	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Flows with new direct services	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Indirect/ connecting flows	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Non-London flows	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total	30.5%	73.5%	5.5%	-9.5%	

This table is, as before, a little difficult to interpret. This is because (for the purposes of classification) we treat new revenue as generated. For 'flows with franchised direct services', there is negative generation as there is less revenue on these flows. This negative generation is transferred to 'flows with new direct services'. The cell showing total generation is correct.

Table 48: Sensitivity 1 - sources of OAO revenue for fifteen years (undiscounted sum in 2013/14 prices)

Flow type	Modelled Generated	Modelled Abstracted	Unmodelled Abstracted	Unmodelled, Abstraction and/or Generation	Total
Flows with franchised direct services	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Flows with new direct services	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Flow type	Modelled Generated	Modelled Abstracted	Unmodelled Abstracted	Unmodelled, Abstraction and/or Generation	Total
Indirect/ connecting flows	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Non-London flows	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Flows with franchised direct services	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Flows with new direct services	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Indirect/ connecting flows	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Non-London flows	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total	33.4%	86.6%	-7.3%	-12.7%	

For 2014/15, between 20.9% and 30.5% of Open Access Operator Revenue is generated- between 26.5p and 43.8p. For the whole fifteen years, the equivalent figures are 20.7% to 33.4% - between 26.0p and 50.1p.

Again, as the worst-performing flow in the model is [redacted] (representing more than half of the absolute errors in 2014/15), we think the model is over-forecasting generation rather than abstraction, and so the estimate is likely to be towards the lower end of this bounded estimate.

Table 49: Sensitivity 1 – modelling errors

Flow	2014/15 Error	Total Error (15 year)	Total Revenue (15 year)
Doncaster	[redacted]	[redacted]	[redacted]
Retford	[redacted]	[redacted]	[redacted]
York	[redacted]	[redacted]	[redacted]
Bradford	[redacted]	[redacted]	[redacted]
Brighouse, Mirfield	[redacted]	[redacted]	[redacted]
Eaglescliffe	[redacted]	[redacted]	[redacted]
Halifax	[redacted]	[redacted]	[redacted]
Hartlepool	[redacted]	[redacted]	[redacted]
Hull	[redacted]	[redacted]	[redacted]
Selby, Howden	[redacted]	[redacted]	[redacted]
Sunderland	[redacted]	[redacted]	[redacted]
Thirsk	[redacted]	[redacted]	[redacted]
Other - Non-London - Other	[redacted]	[redacted]	[redacted]
Sum	[redacted]	[redacted]	[redacted]

11.3 Conclusions

11.3.1 PDFH 5.1 fares elasticities

At an overall level, using PDFH 5.1 fares elasticities moves modelled revenue numbers closer to actuals. However, this is because:

- The sensitivity significantly increases forecast revenue on the "Other – London" flow by [redacted] (in 2014/15). Revenue on this flow is already being over-forecast in the base case ([redacted] error increases to [redacted] error in 2014/15).
- The sensitivity increases revenue on Non-London flows ([redacted]), marginally improving our significant under-forecast on these flows (still under-forecast by [redacted]). PDFC research suggests that this under-forecasting is probably due to localised macroeconomic, demographic and modal competition factors, rather than being related to fares elasticities.

On a flow-by-flow basis for London based flows where competition is introduced, sensitivity test 1 model performance is inconsistent – better on some flows, and worse on others.

In our analysis we have not found consistent evidence of a competitive response from franchised operators in the form of lower fares. Many studies assume that such a competitive response occurs. With a fares elasticity of (<-1), revenue increases when fares decrease: this might suggest that incumbent operators are acting sub-optimally by not decreasing fares, even in the absence of competition. However, operators' fares strategies could be driven by other factors. For example, capacity constraints might mean decreases to fares do not result in increased passenger numbers.

For these reasons, we do not believe there is compelling evidence to adopt PDFH 5.1 fares elasticities in place of the approach recommended by WebTAG guidance.

11.3.2 Alternative approaches to application of fares elasticities

Total revenue from sensitivities 2, 3 and 4 are shown in Figure 35 and Figure 38. Performance of the model in these sensitivity tests does not differ significantly from the base methodology, and none of the sensitivity tests are unambiguous improvements. We do not see compelling evidence to adopt one of these methods of application of fares elasticities, over that in the base methodology.