An assessment of the feasibility of compiling a dataset of European Train Operating Companies

Final report

Office of Rail Regulation

London, October 2014

Content

1. Management Summary

The summary findings of this feasibility study are set out in three sections:

1. The interest in and appetite for participation from non-GB European train operators, and key issues and observations;

2. The interest in and appetite for participation from GB train operators, and key issues and observations;

3. An outline proposal for developing a credible pan-European dataset and comparative exercise.

Feedback from non-GB European train operators

- Findings from non-GB European train operators are based on interviews with seven national railways.

- Overall there is interest in participation in a comparative exercise from at least four or five non-GB European train operators. However, these are medium sized train operators, and further efforts should be made to encourage participation by the larger European railways (i.e. in France, Germany, Italy and Spain), as well as smaller, efficient train operators operating in a competitive environment.

- The primary issue for all non-GB European train operators is how much value and benefit can be provided through such an exercise and how this relates to the level of effort and resource required.

- There is a range of opinions on the type of pan-European dataset and comparative exercise that would be most beneficial to the participant railways.

- The main reason for participation given by the non-GB European train operators is internal business improvement through the development of a small dataset covering costs and revenues and relating these to outputs such as traffic volumes, railway performance and service quality.

- For some of the non-GB European train operators the positioning of unit costs and revenues is interesting only if the railways involved are of a similar size and structure; for others it is important to involve high performing, efficient railways operating in a competitive environment (e.g. Veolia, Netinera in Germany) as well as incumbent state railways, in order to have access to European best practices.

- A number of non-GB European train operators would like to use the outputs of such an exercise to inform and influence their dialogue with policy makers

and stakeholders, using robust evidence about the relationship between the existing market structure (e.g. degree of liberalisation, level of competition) and national rail performance.

- Any exercise would have to differentiate itself from other benchmark reports, such as the rail annual market monitoring report and the European railway performance index, which take a rather macro-economic view on the development of the railway sector, and are not focused on a detailed collection of cost and revenue data.

- Some non-GB European train operators understand the benefits of detailed good practice benchmarking, and in many cases this is already carried out to a limited extent through bi- or multi-lateral exchanges between operators, or operators recognise the benefits but have not prioritised such an activity to date.

- All the non-GB European train operators emphasise the importance of a high level of comparability and methodological robustness; data definition and comparability are seen as considerable but not insurmountable challenges which would need to be addressed by such an exercise.

- All the non-GB European train operators would require the use of anonymous results, and a high degree of confidentiality, as the data required for such an exercise is viewed as commercially sensitive.

- Non-GB European train operators also noted the benefit of using a third party to manage both the process and the confidential collation of the participants' data.

Feedback from GB train operators

- Findings from GB train operators (TOCs) are based on interviews with three TOCs and discussions with the Rail Delivery Group (RDG). Further discussions with the GB TOC owning groups, through the RDG are recommended.

- Each of the GB TOCs interviewed might support the development of a pan-European TOC database and a comparative exercise, although the strength of this support varies.

- The preferred approach would be to make national not TOC to TOC comparisons. These comparisons would be used to understand overall how GB TOCs are performing relative to operators in continental Europe, using robust evidence about the relationship between the existing market structure (e.g. degree of liberalisation, level of competition) and national rail performance.

- A national comparison could be beneficial if carried out robustly and accurately with aggregated data and complete anonymisation, but would require participation by almost all GB TOCs. And in countries with multiple operators, it would be important to have a sufficiently high proportion of those operators participating in the exercise.

- However, there are concerns over the purpose of the exercise - specifically ORR's intentions when it comes to using the results – and the possibility of the ranking of operators.

- TOC to TOC comparisons would not be welcomed by the GB TOCs as these would run into issues of commercial confidentiality, and TOCs would not be prepared to share data at this level of granularity.

- Owner groups with multiple TOCs already carry out internal benchmarking both within Great Britain and with operations in other countries. However for some GB TOCs with international operations the comparison with non-GB European operators is of interest.

- ORR needs to provide confidence to the TOCs that the benchmarking can be carried out robustly with accurate and meaningful comparisons e.g. infrastructure costs passed through as track access charges; ROSCO costs; etc.

- GB TOCs are clear that the resources available for such an exercise would be limited and that all efforts would need to be made to remove duplication and minimise resource requirements.

- GB TOCs have not indicated a willingness to contribute financially to this work.

- It is recognised by the GB TOCs that the data needs to be pulled together confidentially, so the participation of a capable third party is an essential element of the exercise.

- The participation of the Rail Delivery Group (RDG) is considered beneficial as a forum to co-ordinate the perspectives and inputs of the various GB TOCs.

Proposal and Conclusions

- Any comparative or benchmarking exercise needs a very clear objective and scope. We recommend that ORR sets out clear objectives and confirms its preferred scope or scopes, and preferred participants, based on this feasibility study.

- We suggest that ORR builds confidence in the objectives and benefits of the exercise by engaging with the GB TOCs through the RDG on the objectives, purpose and approach of the exercise.

- Any comparative exercise needs to provide genuine, tangible value in relation to the amount of effort put in.

- From our discussions to date with TOCs in Great Britain and continental Europe, there are three distinct approaches to the creation of a pan-European database that could be envisaged:

  1. National comparisons to identify and understand differences in the overall costs and performance of train operations in each country, and the reasons for and drivers of these differences.

  2. Cost and performance comparisons between individual train operators in different countries.

  3. Identification and sharing of good practices between TOCs with drill down studies to look at specific topics.

- A detailed TOC to TOC comparative exercise would be preferable from the perspective of sharing good practices and identifying improvement opportunities. However, based on feedback from both European and GB TOCs, there is a lack of willingness to participate in a TOC to TOC exercise, whilst there is some appetite for a comparison of aggregate national datasets.

- A new pan-European database should build on existing datasets (e.g. management accounts provided to DfT) and minimise effort and inputs on the part of the TOCs.

- ORR needs to ensure that there is a sufficiently robust methodology with a high level of comparability and methodological robustness. The desired level of confidence requires a substantial contribution by peers to define data items and collect them accordingly and therefore there needs to be realism on the level of resources required to achieve this.

- The exercise would need to be put together and the data managed by a third party for purposes of confidentiality.

- An extension of such an exercise to drill down into specific topics would need to be discussed and agreed between the railways once the platform has been established. The railways interviewed expressed a clear view that the participant railways should be in control of shaping the programme.

- There is a potential contradiction between the railways' interest in exploring root causes and producing actionable insights whilst at the same time keep-

ing a tight level of confidentiality. This would suggest that it would be easier to compare between peers where the level of competition is low (i.e. between state railways) rather than between operator groups which deliver a number of operations in different countries.

2. Background

On the basis of recommendations made by Sir Roy McNulty's Value for Money study ORR is now considering extending its activities to monitor the costs and performance of the GB railway system, and is seeking to understand the optimal way to do this, including developing comparative analysis of GB Train Operating Companies (TOCs) with other railways across Europe.

In order to investigate costs and revenues further, two comparative studies have been commissioned already. The first one looked into the development of the cost and performance of GB franchises over time, and the second one was carried out as an international benchmarking study, including a set of six continental European comparators. It compared a variety of cost, revenue and output parameters with a special focus on train staff and rolling stock costs.

ORR is now considering the establishment of an ongoing platform to collect the necessary information and insights that are needed to guide future passenger rail services in Great Britain. As a next step the ORR commissioned this feasibility study to look at the possibilities of compiling a panel dataset of European TOCs[1].

The medium to long term objective of the ORR is to establish a detailed, comparable panel data set, including a number of years of historic data, covering GB as well as a sufficiently large number of continental European train operators, including data on costs, cost drivers and operating characteristics.

For any study to be successful there will have to be sufficient alignment between the objectives of ORR and the TOCs both in Great Britain and elsewhere in Europe.

The key questions in relation to the feasibility of a European panel dataset are:

• What are the key issues and **priorities** of European (both GB and non-GB) railway operators that should be considered in an ongoing data comparison and benchmarking project?

• What are the possible **datasets** that could be developed and how can the constituent data items be collected and made consistent and comparable?

• Under what **circumstances** and **conditions** are European railway operators interested and prepared to join such a platform?
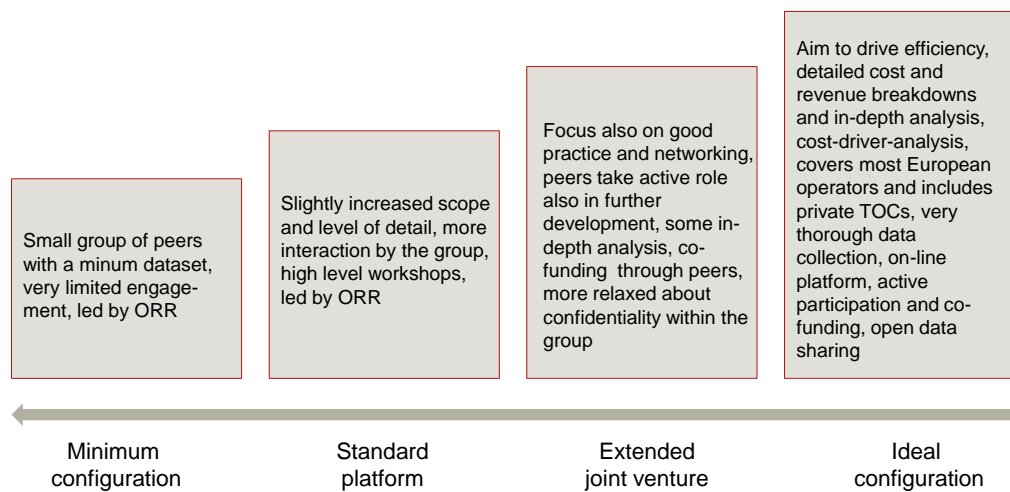
---

[1] The definition of a Train Operating Company (TOC) in this report includes a part of a vertically integrated national railway company

- What measures could be taken to **incentivise** participants to provide data and information and to participate actively in the exercise?

- What are the potential **scenarios** in which an ongoing data comparison platform could be developed, possibly starting out with a small number of operators and growing the exercise over time?

This report will discuss the elements of a number of potential scenarios, their advantages and disadvantages and the views of train operators from across Europe on the different options. The range of possible scenarios, demonstrating the key characteristics and dimensions a platform could take are shown in the Figure below, starting with a very lean, minimum configuration that might be seen as a starting point, and then extended step by step, further developing the different dimensions such as scope, size and mix of the peer group, level of detail and engagement.

Possible scenarios

| Minimum configuration | Standard platform | Extended joint venture | Ideal configuration |

Box contents (left to right):
- Small group of peers with a minum dataset, very limited engagement, led by ORR
- Slightly increased scope and level of detail, more interaction by the group, high level workshops, led by ORR
- Focus also on good practice and networking, peers take active role also in further development, some in-depth analysis, co-funding through peers, more relaxed about confidentiality within the group
- Aim to drive efficiency, detailed cost and revenue breakdowns and in-depth analysis, cost-driver-analysis, covers most European operators and includes private TOCs, very thorough data collection, on-line platform, active participation and co-funding, open data sharing

Figure 1: Platform scenarios

In an ideal configuration the platform would satisfy a wide range of goals, actively supporting the aspiration to increase efficiency in the European railway sector and focusing on a number of cost, revenue and performance impacting issues to better understand the opportunities. As this might be a longer term ambition, several interim solutions with a different mix of characteristics are imaginable.

This report presents: in chapter 3 an overview of the critical success factors which have been identified in the course of the study; chapter 4 sets out the potential peer group; chapter 5 discusses the findings from interviews with train

operators both in Great Britain and elsewhere in Europe; chapter 6 suggests the methodology that could be adopted for such an exercise; chapter 7 sets out the necessary project organisation and execution; chapter 8 provides a roadmap for the development of a comparative exercise, including aspects of planning and resourcing; and chapter 9 contains the conclusions and next steps.

3. Success critical factors of comparative analysis and benchmarking

There are a number of critical success factors which should be addressed in the establishment of an ongoing data comparison exercise, in order to maximise the likelihood of a successful exercise and to encourage a strong commitment and engagement by all peers, providing dedicated time and resources.

The most important critical success factors are set out below.

- All participants need to agree on a **common set of goals and issues** as these will ensure that the peers' interests are sufficiently included. It is essential that these are agreed in the preparatory stage as this will define the scope, the approach, time planning and the level of resources and funding needed. This includes alignment of the purpose of the group between ORR and participant TOCs both in Great Britain and elsewhere in Europe.
- A consistent participation and the provision of necessary resources will only be achieved if a strong **commitment of the senior management** of each TOC is provided.
- A clear and significant **benefit for each participating organisation** in order to achieve the necessary commitment to the exercise.
- The **approach should be interactive**, based on bilateral exchange and providing opportunities for networking.
- The exercise should be **transparent** and the group must be confident about the underlying methodology, data processing and the generation of results.
- Given the high level of sensitivity in dealing with commercially relevant information sufficient measures to ensure **full confidentiality** must be ensured.
- The **quality of the data** must be high, applying commonly agreed and well understood definitions and thoroughly validating data to achieve robust results.
- Depending on the objectives and scope the exercise will be **resource** intensive and require strong analytical capabilities, domain knowledge and ongoing commitment.

Given the importance of these critical success factors and the impact their fulfilment will have on the feasibility of a comparative data platform, each are considered in detail in the following chapters.

4. Peer group

ORR has an aspiration to establish a peer group consisting of TOCs from Great Britain and other European countries. To understand the feasibility of this, we have assessed all Western European non-GB TOCs with regards to their size, preparedness to participate and importance as a comparator. As part of the assessment several non-GB European TOCs as well as GB TOCs and the Rail Delivery Group have been contacted and interviewed to understand their interest in establishing an ongoing European dataset.

4.1 Overview of non-GB European TOCs

In total 15 Western European countries and their state railway companies have been assessed. In addition to the incumbent state operator, private operators should be considered in those (few) countries where a liberalised market environment has led to a significant number of private operators.

civity has grouped these railways into four categories (see illustration) which will be further described along with the railways allocated to each category.
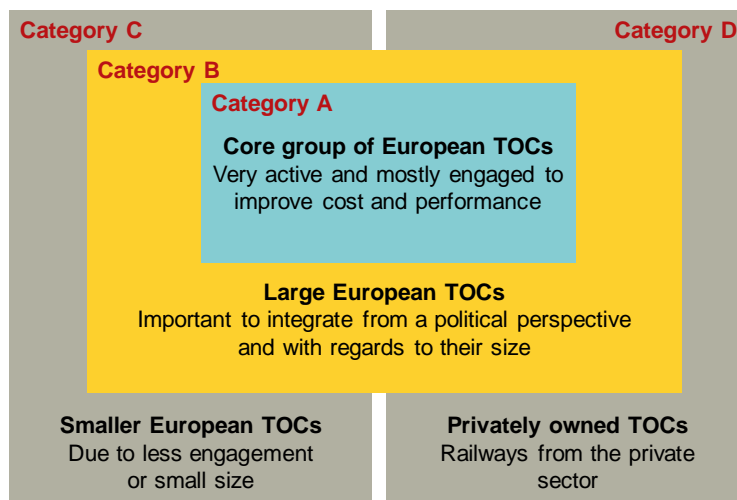
Figure 2: Categorisation of peer railways

Category A consists of European state railway operators, most of who have actively engaged in international studies in the past. These railways have a range of different levels of performance. They include countries with very dense traffic such as Switzerland and the Netherlands as well as countries with less complex and dense networks and lower utilisation (for example Sweden, Norway). All of these operators offer commuter, regional and long distance service.

Both the Netherlands and Belgium would provide good comparators with the more densely operated sections of the GB network such as the South East of England. The services provided in these countries are largely suburban and commuter services, interwoven with some longer distance regional services. Train utilisation tends to be relatively high, although low frequency, low utilisation rural lines also exist. Whilst only comparable to parts of the GB network, both of these countries would be valuable participants in a TOC to TOC comparison.

Denmark, Norway and Sweden offer a mix of commuter services around the major cities, together with long distance services linking the cities together. Whilst the countries are considerably less densely populated than the UK, and with intercity services covering longer distances, line speeds are comparable, and these countries provide an opportunity for TOC to TOC comparisons of long distance and commuter operations. The limited competition in Denmark, and rather more in Sweden have encouraged improved performance and increased efficiency in these countries. The participation of two or more Scandanavian railways should be encouraged.

Both Switzerland and Austria are transit countries and operate in a very different geography to the UK. However, both are recognised as high performing, good practice railways, operating a mix of longer distance, commuter and regional rail services, and although operating with limited competition would nonetheless make interesting comparators in a national, TOC-to-TOC or good practice exercise.

Category B includes the large European states and their state railways, e.g. France, Germany, Italy and Spain. Although experience has shown that it is more difficult to agree participation with these railways, it would improve the comparator group to have one or two state railways contributing to a European panel dataset.

For a comparison at a national level between Great Britain and other national railways it would be important to have participation from some of the big four railways in Germany, France, Italy and Spain, which are comparable in scale to Great Britain and have a full range of long distance, commuter and regional services. There are considerable differences between the arrangements for provision of rail services in these four countries: Germany has a large number of private rail operators in the regional and commuter market, with a smaller number in Italy and very little competition to SNCF and RENFE in the French and Spanish markets. Therefore in order to have participation from railways operating in a competitive or semi-competitive environment we would recommend participation from at least Deutsche Bahn and Trenitalia. However based

on previous experience, SNCF might be a willing participant, given its engagement in the McNulty study and in ORR's European TOC Benchmarking.

Smaller railways such as Ireland and Luxembourg have been grouped in category C. They have been considered as well but are not of the highest priority.

Category D represents private operators, playing a significant role in liberalised markets such as Germany, Sweden and The Netherlands. It is recommended to engage with these operators as well, as a next step, although this will be difficult as they are usually either already represented by operator groups providing services in Great Britain and/or they are likely to be reluctant to participate on grounds of commercially sensitivity. However in regard of their level of cost efficiency and performance, their participation might be very desirable; e.g. some operate under more competitive labour arrangements than the incumbent state railways.

A summary table showing each national train operator in Western Europe is provided below, together with an assessment of each railway's interest in participation and its value as a comparator with Great Britain, based on a combination of our interviews with these railways and our wider knowledge and experience. Although the preferred group of non-GB European comparator railways will vary to some extent depending on the objectives and scope of any exercise.

Section 4.3 below provides more detail on how differences in organisational structures and funding between GB TOCs and European comparators might affect comparability.

	Size of country (t square km)	Service offer (m train-km)	Passenger Demand (bn pass.-km)	Interest in participation	Value as comparator
A-Countries					
Austria – ÖBB				★☆☆	★★★
Belgium – SNCB				★★★	★★☆
Switzerland – SBB				★☆☆	★★☆
Denmark – DSB				★★★	★★☆
Netherlands – NSR				★★☆	★★☆
Norway – NSB				★☆☆	★☆☆
Sweden – SJ				☆☆☆	★★☆
Finland – VR				☆☆☆	☆☆☆
B-Countries					
Germany – DB				☆☆☆	★★★
Spain – RENFE				☆☆☆	★★☆
France – SNCF				★☆☆	★★☆
Italy – Trenitalia				☆☆☆	★★★
C-Countries					
Ireland – Irish Rail				☆☆☆	★☆☆
Luxemburg – CFL				☆☆☆	☆☆☆
Portugal – CP				☆☆☆	☆☆☆

Scale: 0 250 500 750 | 0 300 600 900 | 0 30 60 90

Table 1: European railways

4.2 Overview of GB TOCs

At this stage, it is envisaged that the train operators involved in the exercise from Great Britain would consist of those railway operators holding franchises with the Department of Transport or Transport Scotland, as well as London Overground and Merseyrail. However, this could be extended to open access operators and concessions offered by Transport for London. The franchised train operators and the associated operating group are summarised in the figure below.

Owning Group	TOC
Abellio	Greater Anglia
Arriva UK Trains	Arriva Trains Wales
Arriva UK Trains	Chiltern Railways
Arriva UK Trains	Cross Country
Arriva UK Trains (50%) & MTR Corporation (50%)	London Overground
Directly Operated Railways	East Coast
First Group	First Capital Connect
First Group	First Great Western
First Group	First ScotRail
First Group (55%) & Keolis (45%)	First Transpennine express
GoVia	London Midland
GoVia	Southeastern
GoVia	Southern
National Express	C2C
Serco-Abellio	Mersey Rail
Serco-Abellio	Northern
Stagecoach Group	East Midlands Trains
Stagecoach Group	Stagecoach Southwestern
Virgin Group (51%) & Stagecoach Group (49%)	Virgin Trains
Total	

Table 2: Franchised TOCs in Great Britain (April 2014)[2]

4.3 Structural differences and comparability

European comparators and GB TOCs are different with regards to organisational structures, outsourcing, funding and other aspects. The following table illustrates how these differences between non-GB and GB TOCs might affect comparability.

Aspect	Explanation
Services provided	It is recommended that train operators' costs, revenues and any other aspects are differentiated by applying a categorisation of service offers, e.g. commuter, regional and long distance services. This is consistent with the GB rail market which is seg-

---

[2] TOC franchises and owning groups at time of interviews in April 2014; includes Merseyrail and TfL concession LOROL

Aspect	Explanation
	mented, broadly along the lines of the franchises, between commuter, regional and long distance services. However, elsewhere in Europe the mix of these services differs very much from country to country; some have very densely operated networks with a high share of regional and commuter services and few long distance services (for example the Netherlands and Belgium) whilst others have a more balanced mix of services. Good comparators for long distance, intercity operators in GB would be, for example, Germany and Italy, which have comparable long distance, non-high speed trains.
	In many cases the mix of services is not reflected in the organisation of the railway, and accounting systems and statistical information do not make a distinction. As a consequence significant effort is needed to manually split data and allocate them to each service category, and in some cases, it is not possible to separate service segments. For example: in Denmark commuter services are managed in one business unit while another business unit is in charge of both regional and long distance services; in Austria and Switzerland, long distance data can be separated from regional services, but not from commuter; and in Norway, all segments can be separated, as should be possible in Germany, to the extent that the data is available.
Rolling stock - ownership	Throughout Europe we find different models of ownership and contractual arrangements with providers of fleet and related services.
	GB is the only country in Europe which collectively leases its entire fleet from a number of different rolling stock companies (ROSCOs) which in turn provide different scopes of service (e.g. dry lease, wet lease, etc).
	Other European train operators such as NS Reizigers own and finance their fleet and fully inven-

Aspect	Explanation
	torise it. Others such as DSB own most of their own fleet but also have a number of smaller leasing contracts. In Austria the ownership of fleet is split between two ÖBB subsidiaries: ÖBB Passenger Services and ÖBB Production.
	These different structures vary significantly in terms of their complexity and in the transparency of the costs, with different combinations of bought and leased fleets, owned by railways, railway holding companies and government agencies.
	There are also differences in accounting standards, and depending on the accounting procedures capital costs (depreciation) which can be based on different asset lives. Often financial lifetimes do not correspond to technical and actual lifetimes, thus requiring some adjustments. This can be a very intensive task given the variety and size of each fleet.
	Any comparative analysis would also need to look at the efficiency of each arrangement and hence the actual costs. This is covered in more detail in other industry reports.
	Overall the specificities of the arrangements in each country mean that there is no one good or bad comparator with GB, and that the complexities should be assessed and understood on a case by case basis.
Rolling stock – maintenance	As with rolling stock capital costs, maintenance is organised in different ways by each different railway. Even in GB, different train operators have different maintenance arrangements with their ROSCOs, utilising different combinations of light and heavy maintenance within and outside the lease agreement. Therefore there is no obvious director comparator with GB amongst other European railways.
	Many continental European railways have created subsidiaries to carry out fleet maintenance (e.g. DSB, NS, SNCB, DB, ÖBB etc.) For example, heavy

Aspect	Explanation
	maintenance of rolling stock in Germany is carried out by a central DB function, and DB subsidiaries carry out light maintenance. Nedtrain, a directly owner subsidiary of NS, carries out all fleet maintenance.
	In addition, these subsidiaries may serve other clients as well, and these costs need to be separated from the costs for the operator's own fleet. If the analysis is only looking at genuine maintenance costs then possible overhead costs and margins need to be excluded, too. This is further complicated by internal charging of costs e.g. for depots, and outsourcing of e.g. cleaning (which is a mix of in and out of house).
Train crew	In all European countries the train-crews are the railways' own staff. However the range of tasks that each member of staff has to carry out varies. For example, to what extent is a driver required to carry out train preparation tasks or hand-overs; or are these carried out by maintenance staff? The number of train crew on board trains varies between countries, both by number and roles carried out – e.g. driver, guard, cleaning, catering, revenue protection staff.
	The cost of train crews, consisting of drivers and on-board staff such as conductors requires a full and detailed understanding of their tasks. In some countries, such as Denmark, train drivers have additional tasks such as vehicle preparation and cleaning which train drivers in other countries are not required to carry out.
	Some railways have a considerable share of drivers in training which results in higher headcounts but lower levels of productivity.
	The number of on-board staff requires an understanding of internal rules, quality parameters and legal requirements, setting minimum levels for on-

Aspect	Explanation
	board staff. Here as well it needs to be understood how the tasks of a crew are defined (e.g. revenue protection, ticket sales, customer information, catering services). These functions differ from country to country and also by the type of train service. Even once the staff and their roles have been clearly identified it is necessary to make sure that the costs are comparable, and adjusted for differentiated costs, and social payments, holidays etc. Furthermore, there are country-specific particularities which need to be taken account of - for example the fact that in Switzerland, staff costs are compensated when employees carry out their military reserve duties. What this means is that overall no country or countries are more or less similar to GB, and therefore the challenge is to understand the data from each country in detail and make the necessary adjustments to make them comparable with data from GB.
Station staff	The ownership and the responsibility for station services are organised very differently in different European countries. In GB the major stations are owned and managed by Network Rail, with the others leased as part of the franchise agreements and managed by the relevant TOCs. In continental Europe the split of responsibilities between the operator and the infrastructure manager varies and is just as different. The roles of the train operator and infrastructure manager in respect of stations also varies by country, with different parties responsible for ticketing, train dispatching, information provision, and in some cases station security. Therefore, in a train operations benchmarking, in extremis some operators might not show any costs for station staff. For example, in Norway station staff are employed by Jernbaneverket the infrastructure manager; and this is also the case in Austria. In con-

Aspect	Explanation
	trast, in the Netherlands station staff are employed by the operator NS Reizigers. At Deutsche Bahn in Germany, staff are provided by both the infrastructure manager (DB Netze) and a separate service company (DB Station & Service). So, again, there is no one country that provides a good, clear comparator to GB, and costs would need to be collected from both the train operator and the infrastructure manager, and in Germany from the infrastructure manager subsidiary, DB Station & Service.
Station management	Station management – which is typically services such as station maintenance, cleaning, security and third party management services, such as station retail and parking - is more complex as different functions come into play. Here the question is who is in charge of the station (i.e. infrastructure manager, train operator), if the role includes additional services such as parking, and if there are significant commercial activities (shopping areas). The management and maintenance responsibility for each element of a station also differs between railways, e.g. who owns the station buildings, the platforms, etc. Thus any comparative activity needs to define clearly which functions of a station shall be covered and who owns and manages them. Another particularity is security in stations which in the UK and in France is provided by a dedicated police force, funded separately. Other countries do not have such a set-up, with more costs borne by the railways themselves.
Overhead	Any comparison or benchmarking activity requires a thorough and appropriate allocation of overhead costs, to ensure that each railway and each function bears the appropriate overhead costs. The administrative overhead consists of classical overhead func-

Aspect	Explanation
	tions such as finance, controlling, HR, and procurement, and operational overhead such as service planning, fleet and staff planning, dispatching, maintenance management etc. Different degrees of centralisation and different organisational concepts make this attribution of overheads very complex.
	If the overhead functions are strongly decentralised to operational units they all need to be analysed in detail to fully capture the resources deployed in these units.
	From an organisational perspective, outsourcing and shared services make comparability more difficult. If for example maintenance services are outsourced they will include a share of overhead which often is not transparent. If a railway holding company provides overhead services to different subsidiaries, the costs that are created by train operations need to be understood. For example in Germany, DB has a number of regional service centres, as well as internal suppliers who specialise in providing certain services (e.g. DB Systel for IT).
	Comparisons with GB can be challenging. GB TOCs are typically very lean, with many central functions incorporated into the owning group, and some activities which are typically carried out by the railway, carried out by DfT.
	Similar issues exist in many continental European TOCs with the responsibility for certain activities with the central holding company. The costs of these central activities is commonly held in the central function and needs to allocated to the different operating units including passenger and freight rail operations and bus operators. In the case of outsourced activities such as car parking and cleaning, it is also necessary, although typically not easy, to separate out overhead costs in order to maximise comparability.
	DSB in Denmark and SNCB in Belgium both have

Aspect	Explanation
	relatively straightforward structures, without a holding company, and therefore in these cases it should be relatively straightforward to identify and allocate overhead costs. DB in Germany and ÖBB in Austria both have complex structures with holding companies, multiple business units, and different legal entities, covering multiple logistics functions including passenger and freight train operations, buses and technical support functions. In France there is a complex flow of activities and costs between the holding company SNCF and the infrastructure manager RFF. NS in the Netherlands has a holding company but a relatively simple structure.
	Therefore again, there is no one country that provides a good, clear comparator to GB, and each country and each railway would need to be assessed individually to understand how to make the best possible comparison to the GB railway.

5. Findings from interviews with railways

5.1 Non-GB European railways

The findings from non-GB European railways are based on interviews with and comments from senior managers from seven railways across Europe: Austria, Belgium, Denmark, France, The Netherlands, Norway, and Switzerland. We received responses to our request for interviews from ten railways; however three of these decided not to participate in the interviews and are therefore not included in the analysis.

Overall there is interest in participation in such a comparative exercise from at least four or five Non-GB European TOCs, in particular DSB (Denmark), SNCB (Belgium) and NSR (the Netherlands). ÖBB (Austria) and SBB (Switzerland) might also be persuaded to participate. Efforts should be made to encourage participation by the larger European railways (i.e. in France, Germany, Italy and Spain), and smaller, efficient TOCs operating in a competitive environment. The degree of engagement depends in the end on the balance be-tween added value that will be generated for each peer and the effort and resource required.

The possible set of objectives has been discussed with those European railway operators interview, and can be summarised as follows:

• Position peers' actual costs, revenues and performance in an international context

• Highlight strengths and weaknesses and identify key levers and actionable insights that support an improvement of the peers' performance

• Establish a platform to network with colleagues; share, discuss and transfer experience and good practices

• Inform discussions with stakeholders such as governments, regulatory bod-ies, client organisations etc.

• Develop a long term set of consistent and comparable data to monitor pro-gress over time and analyse trends

Overall the comments from the non-GB European TOCs can be summarised as follows:

• There is a wide range of opinions on the approach to the creation of a pan-European dataset and comparative exercise that would be most beneficial. The key issue for all railways is how much value and benefit can be provided

through such an exercise and how this relates to the level of effort needed for data collection and to develop a sound methodology.

- Most railways have expressed an interest in developing a small dataset covering costs and revenues and relating these to outputs such as traffic volumes, railway performance and service quality.

- It has been stressed by every railway contacted that an accurate definition of input data as well as ensuring a high degree of comparability are of utmost importance and should be taken care of when establishing such a panel dataset. The willingness to provide historic time series will depend to a large extent on the extent of the dataset, the number of years requested and the efforts needed.

- In order to achieve a high level of quality, railways advised that any exercise should be supported by external experts. According to their experience this has helped in the past to better understand each other's input and challenge the quality of the data submitted.

- Some non-GB European TOCs would like to have robust evidence about the correlation between the existing market structure (degree of liberalisation, level of competition) and national performance.

- Some railways have identified the benefits of detailed good practice benchmarking but in many cases this is already carried out through bi- or multilateral exchanges between operators. Some do recognise the benefit but do not yet prioritise it.

- The composition of the peer group has been raised as an important issue by some peers. Some railways would prefer a group of railways that are comparable in size and structure, e.g. Scandinavian railways, Belgium, the Netherlands etc. For one railway it would be of the highest importance to include best practice operators that might be potential competitors in future bidding activities.

- Any further development of the panel dataset, once established, including in-depth analysis should be led by the railways themselves, and not solely by ORR or another regulatory body.

- All the non-GB European TOCS require the use of anonymous results, and a high degree of confidentiality, as the data is seen as commercially sensitive.

5.2 GB TOCs

civity believed that it was essential to have good engagement with the GB train operators in order to understand the feasibility of such a study, and to define an approach that would be acceptable and ideally beneficial to the GB train operators.

We asked ORR to write to all the major GB train operators to ask for their views on the exercise and their interest in participating in such a study, and to understand what set-up would be most beneficial to them and under what conditions they could envisage participation.

We met with senior managers from three GB train operators as well as the Rail Delivery Group (RDG) and this section summarises their opinions on the feasibility study.

All the TOCs with whom we spoke indicated that they would agree to participate if such an exercise was established, but there was a mixed set of views on the genuine benefits of such an exercise, with only one TOC expressing an active interest in participating and that it would be a good thing to do.

There were a number of consistent messages, and these are summarised below:

- Any comparative exercise needs to have agreed objectives and it should be clear from the beginning what the output looks like

- It was commonly recognised that whatever form a comparative exercise is to take it needs to be cognisant of the competitive environment in which train operators function, both in Great Britain and continental Europe.

- There is a common view that any comparative exercise should not be about creating a 'league table'.

- None of the operators with whom we spoke would be comfortable with a direct TOC to TOC comparison exercise.

- A total country level comparison looking at how operators perform overall in different market structures, over time, and with varying structural factors would be of interest to all the TOCs with whom this was discussed.

- All the TOCs interviewed stated that a credible exercise needs a strong methodology and robust and comparable data to make reliable and meaningful comparisons.

- Any exercise should recognise the considerable resource requirements needed both by the TOCs providing the data and explanations and by any third party involved in consolidating and analysing the data and information

to make good comparisons. Data and resource requirements should be min-imised and where possible aligned with data collected and used for other purposes, such as the provision of data to the DfT under the terms of the franchises.

- All TOCs noted the importance of the involvement of a third party to manage the process and be responsible for the data.

- The importance of confidentiality, how this would be managed and what data, results and analysis would be made available in the public domain were also raised as important issues.

Other comments were as follows:

- There was a strongly held view that the costs included in any study should be as full as possible, and adjusted to take account of structural differences between countries, including wage rates, pension contributions, national in-surance etc.

- The benefit of a comparative exercise increases significantly with the collec-tion of time series data, and therefore any such study should not be one-off exercise with all the resource requirements that this takes with little benefit, but an ongoing exercise.

5.3   Possible approaches

From our discussions to date with TOCs in Great Britain and continental Eu-rope it appears that there are three distinct approaches to benchmarking and dataset collection that could be envisaged:

1. National comparisons to identify and understand differences in the overall costs and performance of train operations in each country; how these relate to the form and structure of the industry in each country; observations on the effect of different cost and performance drivers on each country's rail opera-tions.

2. Cost and performance comparisons between individual train operators in different countries.

3. Identification and sharing of good practices between TOCs with drill down studies to look at specific topics.
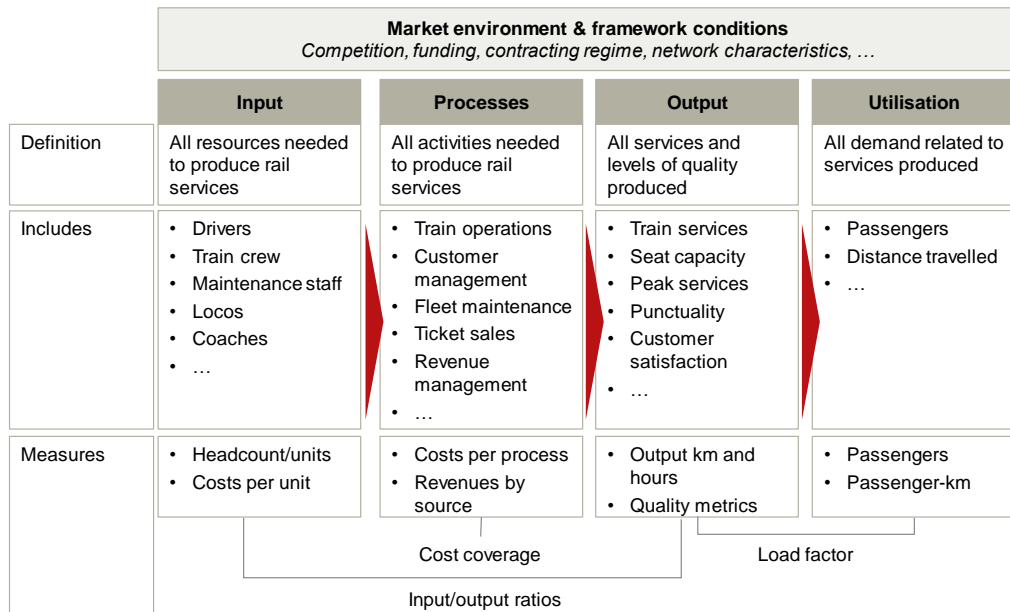
6. Analytical approach

6.1 Scoping

The scope of a European comparative analysis and benchmarking platform will be determined on the basis of the ORR's and the peer group's objectives and the railways' willingness to support the project with available and committed resources. Any dataset could be structured in a number of ways, and in this section we describe the constituent building blocks of a dataset along the value chain of a railway operator. A possible approach to normalise the data for better comparability and data sources, and the option to collect multi-annual datasets, are also evaluated.

It should be emphasised, based on our knowledge and experience that no individual railways have cost, performance or operations data that is notably comparable with data from the UK, or is readily prepared for a comparative exercise. A considerable amount of work will be required to prepare the data from every TOC to make it comparable and ensure its robustness. The detailed review and analysis of the data can only take place as a part of the comparative exercise.

6.1.1 Building blocks

The dataset for a comparative analysis project can be structured along the value chain, as shown in Figure 3, below.

In order to provide railway services, factor **inputs** are needed; two of the most important being human resources and fleet. These can be measured by staff numbers (full time employees), fleet units and costs per unit.

	Input	**Processes**	**Output**	**Utilisation**
Market environment & framework conditions *Competition, funding, contracting regime, network characteristics, …*				
Definition	All resources needed to produce rail services	All activities needed to produce rail services	All services and levels of quality produced	All demand related to services produced
Includes	• Drivers • Train crew • Maintenance staff • Locos • Coaches • …	• Train operations • Customer management • Fleet maintenance • Ticket sales • Revenue management • …	• Train services • Seat capacity • Peak services • Punctuality • Customer satisfaction • …	• Passengers • Distance travelled • …
Measures	• Headcount/units • Costs per unit	• Costs per process • Revenues by source	• Output km and hours • Quality metrics	• Passengers • Passenger-km
		Cost coverage		Load factor
	Input/output ratios			

Figure 3: Sub-datasets and measures of train operations comparisons

The **processes** of transforming these inputs into train services (the output) need to be described in a way such that the operators can allocate their cost and revenue data. Typically a functional cost and revenue structure is applied to facilitate a consistent data collection.

The inputs as well as these total costs are then related to **outputs** which lead to measures indicating a railway's productivity (e.g. Driver FTEs per train-km) or unit costs (e.g. train operations costs per train-km). In addition to quantitative output measures qualitative indicators can be collected, describing for example the level of customer satisfaction collected from surveys about the railway services.

Finally, the level of **utilisation** of train services by passengers is captured by providing the number of passengers, journeys, distances travelled etc. In combination with revenues or capacity figures this indicates the strength of a railway's revenue generation per passenger or the utilisation of train capacity.

Each of these sub-datasets can be further disaggregated, starting from high level, aggregated measures and breaking them further down, for example into detailed cost or revenue elements. The following tables provide an overview on the data items which are likely to be relevant for a realistic comparative analysis exercise, including comments on aspects that need to be considered to ensure data consistency and comparability, given that significant differences exist in data definitions and structures.

6.1.2 Focus areas of comparative analysis

In order to determine the data input required, the set of questions to be answered by the comparative analysis needs to be well defined. Based on the four building blocks described above, ORR and the peer group need to decide on the purpose and use of the dataset and the issues to be addressed.

Possible questions relating to **cost analysis**:

- What are the full costs to produce a train kilometre and what are the costs per function (train staff, fleet maintenance etc.) compared to other operators?

- What are the underlying cost drivers per function, e.g. fleet structure, annual running performance, circulation speed, labour cost levels?

- To what extent do various factors influence a railway's cost position? For example, concerning train drivers.

    − What is the effect of annual labour cost, annual working hours and driving time on staff cost per train hour?

    − What are the practices in different organisations concerning paid break times, the length of shifts, preparation and hand-over times, influencing the productivity of drivers?

- What is the mix of the operators' fleet with regards to size, technology, etc. and what is its impact on operational and capital costs?

- These and similar questions could be explored for other functions such as station and customer services, revenue management, overhead, etc.

Possible questions relating to **revenue analysis**:

- What income is yielded in total per train km and to what extent does it cover the cost?

- What is the breakdown of the operator's income? How much revenue is generated from ticket sales and ancillary business (real estate, advertisement, catering, …)?

- To what extent does the operator rely on subsidies from the government?

- What is the level of ticket prices per passenger km? How much does a passenger pay for comparable services (a typical journey, week day, peak time, 2nd class …)?

- What is the structure of tickets? (single ticket, smart card, senior tickets, discounts, services included …).

- Is there a potential to increase the level of fares? What is the financial effect?

- How do load factors (utilisation of train capacity) impact the potential to increase revenues?

Possible questions relating to **productivity** analysis:

- How many driver full time employees are needed to produce a train hour or train kilometre?

- How large is the driver reserve to compensate for operational disruption or to optimise fleet maintenance?

- How much time of a driver's time available is dedicated to absenteeism, driving or other productive work?

- How much service staff is on board of trains?

- How much maintenance staff is needed to produce a train kilometre?

Possible questions relating to **quality and service**:

- What is the operator's philosophy concerning capacity (e.g. length of trains, double stack trains, seats offered)?

- What is the operator's service offer, considering train frequencies, travel speed, the number of stops, the classes and on-board services offered etc.?

- What is the level of service quality? How many trains are delayed or cancelled and what are the root causes?

- What is the quality perceived by the customer and how is customer satisfaction measured?

- What level of safety and security is achieved in the railway operation?

Possible questions on the **utilisation** of the system:

- How much capacity is offered per train kilometre? How much of that capacity is seat capacity?

- What is the overall level of utilisation in different passenger services?

- To what extent does utilisation vary by time (e.g. in peak and off-peak-hours)?

6.1.3 Datasets

Reflecting feedback from the train operators, this dataset outlined in this chapter could be used for both comparisons of cost, revenue and performance elements at a national level and at a TOC-by-TOC level. Some of the elements might be too detailed for some train operators to be able or willing to provide and might therefore only be shared by a subset of participants, e.g. data on reliability and availability of fleet or customer satisfaction measures.

An example of the outline dataset is presented below. The balance and emphasis of this dataset will clearly depend on the focus areas chosen by ORR and the participant railways.

It is only once the scope of the exercise has been largely determined by ORR and the peer group, and the key issues to be addressed and questions to be asked have been defined, that the specific data required to support and inform that specific scope of work can be finalised. At this stage it will then be possible and necessary to work with the peer group of railways to define common data definitions and to collect and cut each railway's data to align with the common definitions and ensure good comparability. Any particular issues that will need to be addressed to generate consistent data are explained for each data item.

Input data

Data items	Remarks regarding consistency
1. Total headcount	
1.1. Drivers	In-/exclude drivers in training
1.2. Train crew	Include only conductors, no other service staff such as catering
1.3. Maintenance staff	Depending on level of outsourcing, needs to be aligned to operational maintenance costs
1.4. Overhead staff	Needs precise definition of overhead and full understanding of where overheads are allocated to; align to overhead costs
2. Total number of fleet	
2.1 Fleet size	Should understand peak demand, maintenance and operational reserve
2.2 Composition of fleet	Very different mix of fleet: diesel vs. electric, multiple units vs. locomotives and coaches

Data items	Remarks regarding consistency
2.3 Vehicle capacity	Different length and size of vehicles (double deck cars)
3. Total number of stations	Very difficult to compare as responsibilities for stations and platforms vary strongly by country
3.1 Classification of stations	Difficult to compare as classification varies by country

Cost data

Data items	Remarks regarding consistency
4. Total costs	Even if aggregation seems most feasible it requires an understanding of the cost elements/services/activities included
4.1 Operations	
4.1.1 Train drivers	Should include full labour costs; pension schemes, contributions to social security differ from country to country
4.1.2 Train crew	Subject to different staffing and service concepts
4.1.3 Energy	Energy costs include different levels of taxation
4.1.4 Operational management	
4.1.5 User fees	Non controllable costs charged by infrastructure managers; needs to be seen in the context of government subsidies
4.1.6 Station fees	Non controllable costs charged by infrastructure managers, unless operator is in charge of stations or parts thereof
4.2 Rolling stock	
4.2.1 Maintenance	Need to ensure comparability of in- and outsourced maintenance costs; maintenance

Data items	Remarks regarding consistency
	can be included in leasing rates; various definitions of heavy maintenance and hence allocation to maintenance or capital costs
4.2.2 Cleaning	
4.2.3 Preparation	Preparation is sometimes part of the drivers' responsibility
4.2.4 Capital costs	Very different schemes of ownership, on- and off-balance sheet fleet, lifetimes and depreciation vary, cost of financing not always borne by railway, various leasing models, leasing models can include maintenance services
4.3 Support functions	
4.3.1 Overhead	Ensure that overhead is fully understood and correctly allocated (can be provided by holding, core business units, shared service units etc.)
4.3.2 Marketing, sales and planning	Might need further disaggregation, for example if cost efficiency of sales is of importance

Revenue data

Data items	Remarks regarding consistency
5. Total revenues	
5.1 Farebox revenues	
5.2 Government contributions	Needs to be seen in connection with track access charges ("communicating vessels")
5.3 Ancillary business	Should be specified by each railway as scope and level of income are very different

Output data

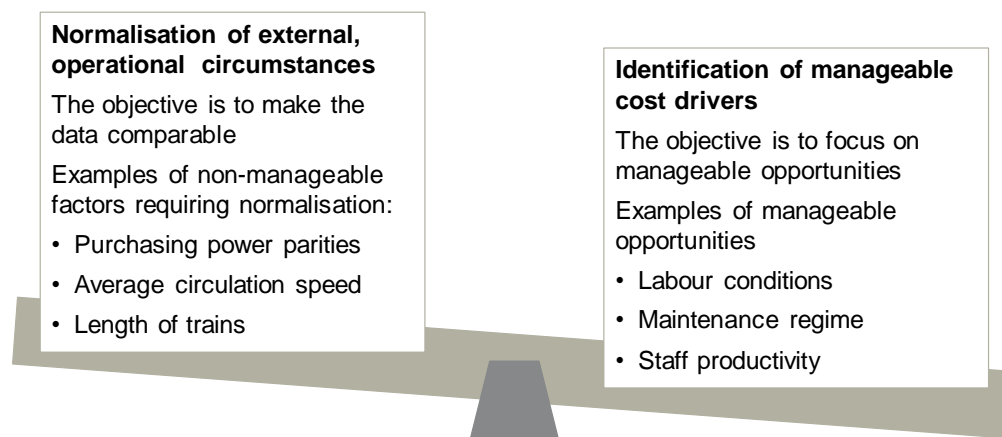| Data items	Remarks regarding consistency
6. Quantitative output	
6.1 Train kilometres	In-/exclusion of empty rides, cross-border services, time-tabled vs. actual
6.2 Car kilometres	Different length of cars
6.3 Seat kilometres	Treatment of standing areas, different levels of quality (available seats)
6.4 Train hours	
6.5 Operated route length	Different counting of track (e.g. sidings)
7. Qualitative output	
7.1 Customer satisfaction	Subjectivity of surveys and country specific levels of expectation
7.1.1 Punctuality	Idem
7.1.2 Safety	Idem
7.1.3 Information	Idem
7.1.4 Service	Idem
7.2 Reliability	
7.2.1 MDBF, MTTR	Different definitions of defects
7.2.2 Defects	Different definitions of defects
7.3 Punctuality	Different thresholds and measurement concepts

Utilisation data

Data items	Remarks regarding consistency
8.1 Passengers	
8.2 Passenger kilometres	

6.2 Requirements for normalisation

It has been highlighted in almost every interview that comparability is a major issue that needs to be addressed fully. Even if the methodology is robust and tested it should be discussed and reviewed thoroughly with the peer group. It also needs to be taken into consideration that a more sophisticated method of normalisation will require more data and information from peers.

To achieve better comparability some circumstances which are outside the short or medium term control of the operator can be taken account of through normalisation. This normalisation process leads to adjustments in the dataset and delivers a simulation within which all the peers are operating under similar conditions.

Normalisation of external, operational circumstances

The objective is to make the data comparable

Examples of non-manageable factors requiring normalisation:

• Purchasing power parities
• Average circulation speed
• Length of trains

Identification of manageable cost drivers

The objective is to focus on manageable opportunities

Examples of manageable opportunities

• Labour conditions
• Maintenance regime
• Staff productivity

▶ **Creating an intelligent comparability and finding the differences which are changeable at the same time**

Figure 4: Normalisation of datasets

Cost and revenue data should be at least normalised for different economic circumstances in different countries within Europe by applying purchasing power parities to adjust. At the same time currencies are converted to one comparable currency. This step requires data which is publicly available.

Rolling stock maintenance costs have fixed, time based elements and running variable cost which are related to a train's annual distance operated. The share of fixed cost is normalised by taking the annual distance operated into account. This requires data about the annual distance operated by train-sets.

As travel speed on the network varies (even within service categories such as commuter, regional or long distance services), the demand of train personnel varies. To compensate for these differences train staff costs are adjusted by a

normalisation of travel speeds. This speed can usually be provided by the railways.

Further normalisation could be undertaken to take account of other differences between railways.

6.3 Data collection

As a robust and well-defined data input is key to delivering quality results the collection of data directly from the railways is highly recommended. If possible, existing datasets should be used and expanded upon (e.g. data provided by GB TOCs to the DfT as a franchise obligation). We have not identified any other datasets which seem to be usable, especially to cover the scope set out in this study. The provision of a multi-annual dataset right from the beginning of such an exercise does not seem to be a feasible option, but could be developed over time.

All the TOCs interviewed call for a high level of comparability and methodological robustness, and data definition and comparability are seen as considerable but not insurmountable challenges that would need to be addressed by such an exercise.

Also ORR needs to provide confidence to the TOCs that the benchmarking can be carried out robustly with accurate comparisons e.g. infrastructure costs passed through as track access charges; ROSCO costs; etc.

GB TOCs are clear that the resources available for such an exercise would be limited and that all efforts would need to be made to remove duplication and minimise resources. At the same time, the peer group needs to understand that a high level of data quality requires an adequate level of internal resource and support.

Data collection for the datasets described can be carried out in two different ways:

• Research to collect data and information from publicly available sources such as national statistics, annual reports and other publications;

• Collection of data according to definitions through templates/spreadsheets and questionnaires directly from railways.

As most data needed is directly related to the TOC it is strongly recommended to collect this data first hand from each peer, starting with a small but robust dataset and then expanding this stepwise. Only this approach will deliver a sufficient level of completeness and quality.

ORR has requested the assessment of the possibility of creating **multi-annual datasets** which allow for trend analysis. As explained above, the possibility to establish a dataset including historic data based on previous benchmarking exercises is very limited. An alternative option is to ask railways for multi-annual datasets. Due to the lack of clarity about the scope and the likelihood that the collection of these datasets will be very resource intensive for each peer a collection of historical datasets will be difficult to start with, but could be developed over time.

6.3.1 External and publicly available data sources

For the datasets described above the following data **sources** are available:

- **Annual reports published by the railways**

  Although publicly available it is recommended not to build the analysis of costs and revenues on annual reports. They do not provide sufficient transparency about the definition of the required data elements and do not allow allocating costs and revenues systematically to a functional taxonomy. Furthermore, the figures are subject to extraordinary effects and changes (e.g. reorganisation, change in accounting principles). Even on a highly aggregated level of data a lot of uncertainties remain, e.g. it is often not clear what services are included or excluded (e.g. freight, high speed) and if cost and revenue data correspond to the needed outputs (train kilometres etc.)

- **Statistics provided by national or international institutions**

  Sources like Eurostat and OECD are valuable to use for macro-economic information such as inflation rates, currency exchange rates and purchasing power parities. It is recommended to use these sources for example in the process of normalisation.

- **Former benchmarking reports**

  Reports from previous benchmarking studies might be seen to be useful to derive historic data and start developing trend analysis. However, they differ in scope, the level of detail, the way data is defined and cut and the normalisation method applied. Furthermore they vary with regards to the peer group. Hence they should only be used if the underlying dataset can be accessed, if it is fully understood and if use is authorised and not conflicting with existing confidentiality agreements. Some of the relevant existing reports are:

- The European Railway Performance Index, linking an aggregated performance index with public costs (recurring subsidies and average public investments), the level of market liberalisation and the governance model.

- The Independent Regulators' Group – Rail "Annual Market Monitoring Report" which is providing information on the market structure in European countries, characteristics of the railway networks and its utilisation by passenger and freight services.

- The European Commission's report monitoring the development of the rail market, mainly exploring the share of rail, its volumes, performance and investments in infrastructure in EU member states

All of these reports are taking a very macro-economic view on the development of the railway sector. They do not seem to follow an approach comparable to the one described here, aiming at a detailed collection of cost and revenue data and including an appropriate methodology to ensure comparability among countries. None of these reports seems to be usable to assess the efficiency of operators and to identify opportunities for improvement. The only report trying to link performance with costs is the European Railway Performance Index. However, it is not transparent how cost data have been collected and how robust and comparable they are.

6.3.2   Data collection from participating railways

Data collected directly at and in close cooperation with each participating peer requires some efforts and resources but represents the most reliable and robust approach to build up a dataset. It has proven to be the only way that ensures data to be sufficiently understood and consistent with definitions.

It is our view that except for some general information that will be needed to complement the analysis, e.g. inflation rates and purchasing power parities, data should be gathered first hand and in close collaboration with each peer. There are different ways of engaging with the peer group in relation to data collection and quality assurance which will be described further below.

With regards to this approach an important concept should be set out: instead of relying on the railways' own allocation of their raw data to the predefined structure of the dataset it could be the project team's responsibility to undertake this allocation. Practically, raw data such as costs per cost centre, detailed staff lists with headcounts etc. would be delivered to the project team and allocated to the data structure in close cooperation with the railway. The advantage of this approach is that the project team has full transparency and a very deep understanding of the data and can discuss the allocation in detail with each

peer. It has become civity's standard approach in many urban transport benchmarking projects. Although more resource intensive it significantly supports quality assurance and comparability.

As a first step a collection of data through templates directly from participating peers is seen as the minimum requirement. In a next step this process can be extended to collecting more detailed raw data.

Technically data management can be dealt with by using standard tools such as spreadsheet files and questionnaires to capture qualitative information. This would be sufficient in an initial start-up phase. If the comparative analysis platform becomes a continuous activity, data management could be facilitated by an online solution. Such a web-based application would enable the peer group to:

- Enter the complete dataset on-line and in subsequent steps and go through first plausibility checks as data is entered into the system;

- Give access to different people involved in the process, providing different user rights;

- Analyse results through a flexible reporting tool and produce customised graphical illustrations of the results.

Such a web-based tool has been implemented for example by UIC to support the data management and processing for LICB.

Whatever option is chosen to manage the data continuous support, a dedicated project team will be needed. Its responsibility would be to:

- Manage the regular process of data collection and processing, including the provision and administration of the necessary support tools;

- Assist to clarify data definitions and hence avoid potential misunderstandings and misinterpretation;

- Execute the necessary quality control by carefully validating each peers data, checking completeness and consistency;

- Analyse the data according to the agreed methodology and prepare reports and presentations.

The importance of these tasks cannot be underestimated as they are success critical and absolutely crucial to ensuring continuity as well as comprehensive and accurate data. In order to ensure a well-functioning process ORR should ensure either internal or external resources are available to facilitate this process. These requirements will be described further in chapter 7.

7. Project organisation and execution

ORR should take the initiative and establish the project in its initial phase, with clearly articulated objectives and scope and with formal agreement and support from the participant railways, perhaps through the RDG. This initial phase should cover the establishment of the peer group and the development of the recommended continuous small benchmarking dataset.

Furthermore, ORR's role should be to ensure that there is a sufficiently robust methodology with a high level of comparability and methodological robustness; the desired level of confidence requires a substantial contribution by peers to define data items and collect them accordingly and therefore there needs to be realism on the level of resources required to achieve this.

The exercise would need to be put together and the data managed by a third party for purposes of confidentiality. Independent of where resources come from there needs to be a dedicated team providing the necessary expertise.

Any further activities drilling down into other subjects would have to be discussed and agreed between the railways once the platform has been established.

For the time being there is no or no clear commitment on funding. This issue should be re-addressed as the project gains momentum and generates recognisable benefits.

With regards to the organisation of a comparative analysis platform the following questions arise:

- **Lead** - who will take the overall responsibility to promote, structure, guide, develop and facilitate the complete process?

- **Execution** - who will from an operational perspective execute the necessary tasks, e.g. data collection, analysis and reporting, peer workshops etc.?

- **Data ownership** - who will own the datasets, have access to the sensitive data and guarantee full confidentiality?

- **Funding** - who will fund all necessary activities, especially external resources and services shared by the group?

Different options exist to respond to these questions as outlined in the table below. They can change and develop as the project moves from its start-up phase to a stabilised and more comprehensive platform.

First of all, the **leadership** of the project could come from ORR[3] or from ORR together with the participant railways, perhaps through the RDG, acting as a forum to co-ordinate the viewpoints and inputs of the different GB TOCs. It is our view that a coordinated approach would be more effective and produce better and quicker results.

In a start-up phase ORR would need to state clear objectives for the project, and might need to take the responsibility for establishing the project and convincing the TOCs of its benefits. As the engagement of the participant railways increases (or in order to foster the engagement of the participants) the leadership could be shared with the participants and the project could have joint leadership and responsibility. Practically this could happen through a steering group, meeting regularly and determining an annual or multi-annual programme. As ORR is interested in establishing a long term platform which benefits its own objectives, it would be important to remain committed and engaged with the exercise.

	Start-up	Evolution	Stabilised
Lead	ORR takes the initiative, states objectives and establishes project, ideally together with RDG	ORR and railways share responsibility for further development	ORR and railways share responsibility for further development
Project execution	Operational project management is outsourced to external project manager with sufficient benchmarking experience	Project management is outsourced to external project manager with benchmarking experience; start building up own resource at the ORR (similar to UIC model)	ORR and/or railway(s) is/are in charge of project management and provide continuous resource

---

[3] This report refers to ORR but it is recognised that this could be an exercise jointly undertaken by ORR and DfT

	Start-up	Evolution	Stabilised
Ownership of data	Data is handled by an external custodian to ensure confidentiality	Data is handled by an external custodian to ensure confidentiality; data is being transferred to ORR	ORR takes over responsibility for full data management (Similar to UIC model)
Funding	ORR ensures funding to start-up the project	Joint funding by ORR and railways	Joint funding by ORR and railways

Table 3: Summary of development stages

Project management and **execution** could be started using an external support which is more experienced and more flexible and scalable than creating a fixed internal ORR resource. In later stages when the platform becomes more established and develops ORR and RDG could establish their own resource. Experience has shown that a successful project needs dedicated staff with a strong background, expertise and management skills to ensure progress and continuity. The benefit of a third party project manager, is that it is neutral, trusted by all parties and is able to establish a common agenda amongst all the participants, but acts in all participants' best interests and is overseen by a steering group made up of all of the participants in the exercise.

Data ownership is a very sensitive issue as all data and information will have to be treated very confidentially. This issue is strongly linked to the previous one about project management and execution as the question is not only who holds the data but also who handles them and carries out the necessary analysis and reporting. There are clear benefits from the participation of an independent third-party organisation to manage and analyse the data, to provide confidence to all participants.

Funding could be provided through ORR, for a first phase. This would very much help to facilitate starting up the platform. It would also reduce the complexity of negotiations. However, as peers want to have more influence and shape the exercise to meet their needs a co-funding through the railways would be preferable. It could also be an option for sub-groups to be established, dealing with specific topics and undertaking separate analysis which could then be funded separately by that sub-group.

Overall it is our view that the exercise, including leadership and funding, should be developed as a joint exercise between ORR and the participant railways, with clear and agreed objectives from the start. A work plan would be set out to

meet these common objectives, which creates sufficient benefit for all parties involved. If the influence and the benefits are shared, the costs and funding should be borne by the parties as well. In contrast, the operational programme management and execution of the activities, including handling the highly sensitive datasets, should be in the hands of an external and independent organisation, protected by strict confidentiality agreements.

7.1 Incentives

One of the key recommendations to ensure commitment of peers is ORR's early engagement in developing the platform. It will require senior management attention and a dialogue with managers from all parties involved. Furthermore, generating value and demonstrating the benefits of the project will be a key driver for commitment.

Previous comparison and benchmarking projects have shown that there are a number of risks which should be taken into account, e.g. participants do not

- Support the project with enough management attention; tasks are delegated and lose momentum due to conflicting day-to-day activities;

- Deliver complete datasets, leading to insufficient data population and causing additional efforts to collect data;

- Dedicate enough resources to data validation, risking the consistency of data and equally requiring more resource to check data;

- Actively engage in development of the platform, weakening the opportunities to exchange information in the group and networking;

- Trust the adequate use of data and express concerns about matters of confidentiality which can block a more in depth analysis.

A number of **provisions** can be put in place to mitigate these risks and set incentives for the participants:

- Increase the level of management attention by raising the engagement on the project to a top management level; this can be achieved by initiating it on a high management level and maintaining the contact about the development and key outputs on this level

- Give participants an active role so they can define their own goals and needs, ensuring that they will receive a valuable output. This should happen in the early stage of the project, for example by the means of a joint start-up workshop

- Involve peers in the development of the methodology; this avoids the process being a non-transparent black box and makes results more acceptable. Working group meetings used to explain the options and agree on the approach have proven to be helpful and increase identification with the project.

- Make it interactive, providing result oriented reports, smart tools and state-of-the-art interfaces; here as well, review sessions and workshops with peers to discuss results, share good practices and determine future steps are valuable

- Request a financial contribution from peers to increase their obligation to produce substantial outputs justifying the budget they spend; this should happen in combination with a contractual agreement that clearly sets out the rules, including everyone's rights and liabilities.

7.2 Confidentiality

TOC to TOC comparisons would not be welcomed by GB train operators as they would run into issues of commercial confidentiality, and TOCs would not be prepared to share data at this level of granularity.

All the non-GB European TOCs require the use of anonymous results, and a high degree of confidentiality, as the data is seen as commercially sensitive.

Due to the nature of the data and information involved, all data and information needs to be handled completely confidentially.

A distinction needs to be made regarding data exchange within the group and communication towards the public.

Within the group different possibilities exist:

- Show fully anonymised reports where peers can only identify their own position

- Partially disclose some of the data and information, be it qualitative or quantitative and discuss it openly in the peer group

- Fully disclose data and information within the group and share reports openly.

Some of the results of the group might want to be shared with a public audience. In the past, the ORR has published anonymised reports which have been authorised by each peer before publication. The acceptance of peers will very much depend on the kind of data and the level of detail that shall be disclosed, requiring a discussion with the group on a case by case basis.

8. Roadmap

8.1 Activities

The establishment of a dataset will require a number of activities from both ORR and participating railways. An initial set of activities is described here as it provides an input for further planning of roles and responsibilities, time planning and the provision of resources.

Activities to be carried out by ORR or a designated project team:

• Overall project management and organisation of all activities, including contractual arrangements, funding, confidentiality agreements

• Structuring of relevant issues and definition of work packages

• Development and dissemination of interview guides

• Specification of data and information required

• Interviews at the railways, data collection, quality assurance

• Evaluation of data, analysis, conclusions

• Preparation, moderation and wrap-up of meetings/workshops

• Documentation of all results, preparation and coordination of reports

Activities to be carried out by the representatives of the participating railways (GB and non-GB European TOCs):

• Nomination of key contacts responsible for the project

• Final definition of topics and depth of analysis

• Provision of relevant data and information, including definitions

• Provision of key contacts and interviewees to discuss data and information

• Participation in workshops and review of documents/reports

• Contribution of presentations at workshop sessions.

8.2   Ongoing process

The section describes an ongoing process which should be installed to ensure a continuous management of the dataset. This process consists of five recurring elements which could form the basis for an annual planning and programming of all activities.

It starts out with step 1, the definition of the work programme which should be a joint exercise of ORR and all participating railways to develop a common understanding of the project's goals and the issues that shall be addressed by the dataset. This work programme needs to be translated into a time, resource and budget plan which will be binding to all parties involved.
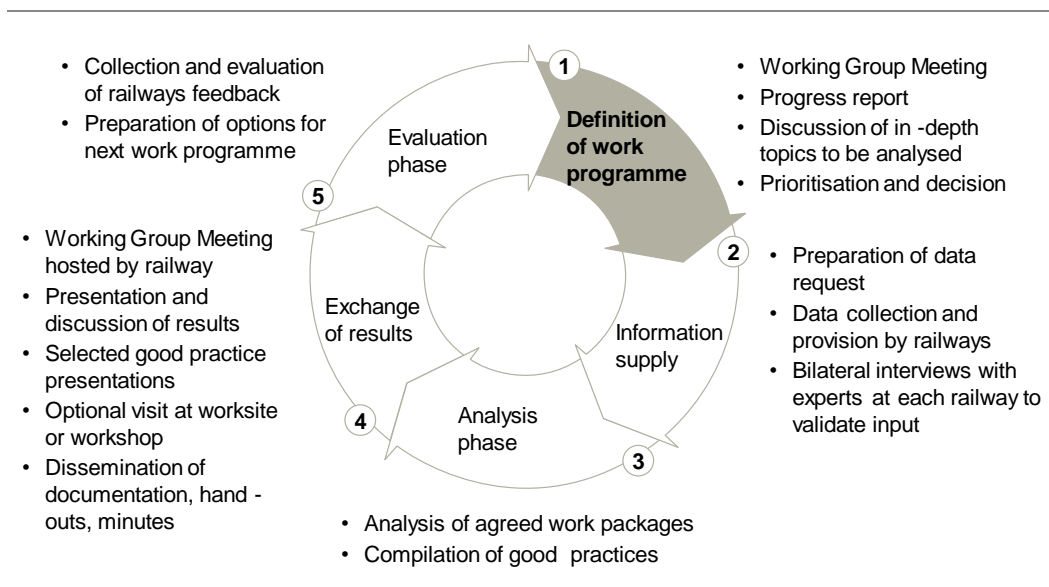
Figure 5: Annual process (schematic)

In a second step, data would be collected according to the approach set out above. This step covers the complete process of collecting, validating and discussing the data with the railways' experts, including ensuring comparability and consistency of definitions.

In step 3 data would be analysed by the project team and prepared for discussion by the peer group. It would include a description of each operator's operating environment and key characteristics, comparisons with and without normalisation, trend functions, the explanation of differences and good practices. This documentation would serve as an input for the discussion with the group.

As shown in step 4 the results would be provided to and discussed with the working group, consisting of all participating railways. It could be complemented by additional presentations, for example by some peers on selected topics and good practice. It might also be fruitful to organise site visits to facilitate the ex-

change on certain topics (e.g. visiting an optimised rolling stock maintenance facility). The results of this meeting would be incorporated in the analysis, documented and disseminated according to the agreed standards and rules of confidentiality.

The evaluation phase in step 5 closes the loop and is intended to collect the feedback that is necessary to enter into a next stage of planning and to shape the programme for the next cycle. Options for further analysis will be collected and prepared for the discussion in the following step 1.

9.   Conclusions and next steps

9.1   Conclusions

From our discussions with train operators in Great Britain and continental Europe we suggest that there are three distinct approaches to the creation of a pan-European database:

1. National comparisons to identify and understand differences in the overall costs and performance of train operations in each country, and the reasons for and drivers of these differences.

2. Cost and performance comparisons between individual train operators both within and between different countries.

3. Identification and sharing of good practices between TOCs with drill down studies looking in detail at specific topics.

Each of these approaches would have different objectives and a different scope and would therefore produce different outputs. The preferred approach is therefore very much dependent on the objectives for the exercise from ORR and the participant railways.

The table below presents a comparison of the key elements of each of the three approaches. However it is important to stress that the overriding driver of success of a comparative exercise of this kind, whichever approach is taken, is a commitment to the objectives and scope of that approach by the senior management of participant railways and the allocation of sufficient resources to the exercise.

	1. National comparisons	2. TOC to TOC comparisons	3. Good practice sharing
Participation (size and diversity)	No minimum number of participants, but should include at least two of the large European countries to make a meaningful comparison with GB, including national and private operators. Would require most GB operator groups to provide	Suggested minimum of five participants to balance value of outputs with resource required. Segmentation of types of TOC would enable comparisons of part of GB (e.g. NSE) with whole countries in Europe (e.g. Belgium, Nether-	No minimum number of participants for good practice sharing, but would benefit from railways which are recognised to be industry leaders in different aspects of their operations, including Switzerland and possibly Asian railways (e.g.

	1. National comparisons	2. TOC to TOC comparisons	3. Good practice sharing
	sufficient coverage of GB rail industry	lands)	Japan)
Length of comparisons	Meaningful comparative analysis requires time-series data with a minimum of three years and ideally five or more. However it is often difficult to back-calculate data to new definitions and therefore a comparison platform should be established with the intention of developing a panel data set over a number of years.		Good practice sharing does not require time series data, although it needs to include innovative railways with emerging technologies and processes.
Consistency of definitions and robustness of data	This will require considerable effort by all participants to provide data which is consistent with agreed definitions. National comparisons would require consolidation of data between operators in each country as well as between countries. There is no structural difference in the level of comparability between railways.	This will require considerable effort by all participants to provide data which is consistent with agreed definitions. It would be preferable to agree outline definitions between GB TOCs and then discuss these with non-GB participants. There is no structural difference in the level of comparability between railways.	Good practice exchange requires significantly less effort to make inputs consistent. The focus here is on appropriateness of practices and processes in different countries and environments, how these can be modified to make implementation most effective, and how the impact can be monitored and compared.
Coverage of data and information	A country to country comparison will focus on the outputs that are achieved (i.e. patronage, performance, quality) compared to the inputs (i.e. costs, labour, resources) given the industry set-up, funding and regulatory structures and industry constraints in different countries.	This is likely to focus on those aspects of a train operator's business which are within management contol. The detail will depend on the preferences and priorities of ORR and participants and could cover all aspects of a train operator's business including costs, performance, quali-	This depends on preferences and priorities of ORR and participants and could cover all aspects of a train operator's business, as well as approaches to tendering/ franchising and interaction with infrastructure manager and suppliers.

	1. National comparisons	2. TOC to TOC comparisons	3. Good practice sharing
		ty, support functions etc.	
Outputs	An analysis of the structural differences between the rail industry in different European countries, describing inputs and outputs and indicating the effectiveness and efficiency of each country's structure	A comparison of the costs and performance of the participating TOCs, normalised for structural differences between railways and countries, indicating opportunities for improvement	A set of good practices and processes which have been successfully applied in different countries, with guidance on how these might be successfully implemented in different countries and environments
Methodology	A detailed discussion and outline of how a comparative platform could be established is provided in sections 7 and 8 of this report. The most appropriate methodology will be determined based on clear and agreed objectives and scope and discussions with the specific needs of the participants in any particular exercise. The immediate next steps to achieve this are set out in section 9.2 below		
Risks	The biggest risk is the willingness of TOCs to participate, and their ongoing commitment, engagement and resource commitment once involved. Discussions have indicated that a national comparison would be the preferred option for GB TOCs. Robust consolidation of data within each country to create a representative national dataset also presents a major risk	The biggest risk is the willingness of TOCs to participate, and their ongoing commitment, engagement and resource commitment once involved. Discussions have indicated that TOC to TOC comparisons would not be the preferred option for GB TOCs	The biggest risk is the willingness of TOCs to participate, and their ongoing commitment, engagement and resource commitment once involved. The risk of willingness to share industry leading practices and ng processes, particularly by high performing railways is always a potential issue
Timescale	A comparative exercise, particularly one involving good quality data, should be established on an ongoing basis in order to		

	1. National comparisons	2. TOC to TOC comparisons	3. Good practice sharing
	achieve meaningful results and significant benefit for its participants. The time taken to establish such an exercise depends on the commitment of the potential participants and the resources made available, but should be able to be achieved in between six and nine months		
Costs and resources	The exact financial costs and resources required to establish a comparative exercise depend wholly on the scope of the exercise, the number of participants and the balance of resource expended between the participants and any third party project manager and data custodian/analyst. This will need to be determined once the objectives, scope and participants are agreed		

Table 4: Comparison of approaches

A detailed TOC to TOC comparative exercise would be preferable from the perspective of sharing good practices and identifying improvement opportunities. However, based on feedback from both European and GB TOCs, there is a lack of willingness to participate in a TOC to TOC exercise, but there is a potential appetite for a comparison of aggregate national datasets.

Therefore we believe that the approach which has the greatest likelihood of achieving commitment from the largest number of participants would be a comparison of aggregate national datasets, based on an initial limited dataset which includes costs, revenues and some performance data. The proposed comparisons would include explanatory factors on the railway industry in each country, including competition; flexibility; labour relations; government constraints. The results would be presented anonymously on a national level based on a set of jointly agreed definitions.

9.2   Next steps

Based on the findings and conclusions of this feasibility study, we suggest that there are a number of next steps to be undertaken to develop successfully a comparative exercise. These are set out below.

1. ORR to set out clear objectives for a comparative exercise

2. ORR to confirm its preferred scope, or scopes, based on this feasibility study

3. ORR to identify preferred participants from the non-GB European railways

4. Review of GB TOC management accounts to understand to what extent these, in their current form, could be fit for purpose for a comparative exercise

5. Outline project plan to be proposed, based on the four steps above, to RDG for discussion and agreement with GB TOCs through the Franchise Strategy Group

6. Outline project plan to be proposed to preferred non-GB European railways for discussion

7. Agree funding model for first phase of comparisons – ideally by GB rail industry: combination of ORR, DfT and GB TOCs

8. Finalise and confirm scope and key outputs from comparative exercise with ORR and all committed participants

Appendix

Guiding questions discussed with train operators

Criteria	Guiding questions
Overall interest	Would your organisation in general be interested in joining an ongoing comparison platform with other European peers? Would the CEO, the board or other managers of your organisation support this position?
Goals	What would you like to get from such a comparative exercise, e.g. a positioning, a trend analysis, information for stakeholders, knowledge about good practice?
Scope	What aspects should be addressed in such a project: costs, revenues, service levels and quality and how detailed should these aspects be covered? Do you find a segmentation by service categories helpful/necessary? Can you imagine to provide historical data as well? For how many years?
Peer group	What is the minimum size for such a peer group? Do you have any preferences regarding the configuration of the peer group?
Involvement	How intensively would you engage with your organisation, e.g. data delivery, work on methodology, regular workshops and networking, in-depth analysis?
Ownership/lead	Would you accept the ownership of the project including its datasets by a regulator? What would your preferred set-up be?
Funding	Would your organisation be willing to contribute financially and fund these activities?
Framework	To what extent and under what circumstances are you willing to provide and discuss data and information?

Frank Zschoche

phone: +49.40.181 22 36-66
mobile: +49.171 77 11 790
fax: +49.40.181 22 36-99
email: frank.zschoche@civity.de

William Adeney

phone: +44.20.3478 2237
mobile: +44 7771 850564
fax: +49.40.181 22 36-99
email: william.adeney@civity.co.uk