

ORR & Network Rail

**Mandate AO/015: NR Bottom-Up
Benchmarking Programme Audit**

Final Report

Version 1.1 | January 2012

**** PUBLISHED VERSION ****

Ove Arup & Partners Ltd
13 Fitzroy Street
London
W1T 4BQ
United Kingdom
www.arup.com

This report takes into account the particular instructions and requirements of our client.

It is not intended for and should not be relied upon by any third party and no responsibility is undertaken to any third party.

Job number 209830-15

ARUP

Contents

	Page
0	Executive summary 2
0.1	Introduction 2
0.2	Bottom-up benchmarking approaches in other sectors 3
0.3	NR bottom-up benchmarking –programme structure and objectives 7
0.4	NR engagement with comparator organisations 8
0.5	NR’s approach to collecting and analysing benchmarking data10
0.6	Bottom-up benchmarking data review 11
0.7	Application of bottom-up benchmarking programme outputs to CP5 efficiency proposals 16
0.8	Conclusions 17
0.9	Recommendations 21
1	Introduction 24
1.1	Background and objectives 24
1.2	Our approach 24
1.3	ORR mandate – reference report sections 26
1.4	Report structure 29
1.5	Acknowledgements 29
2	Background and context 30
2.1	Bottom-up benchmarking definition and objectives 30
2.2	CP5 determination process overview 30
3	Bottom-up benchmarking approaches 34
3.1	Introduction 34
3.2	Water sector (England & Wales) benchmarking 35
3.3	Electricity distribution (UK) benchmarking 38
3.4	Gas distribution (UK) benchmarking 39
3.5	Electricity & Gas Transmission (UK) benchmarking 40
3.6	International metro benchmarking 41
3.7	Lessons learned from other sectors 42
4	Developing effective benchmarking groups 44
4.1	Introduction 44
4.2	Success factors for benchmarking programmes 44
4.3	Key stages and timescales for establishment of a successful benchmarking programme 47
4.4	Inherent challenges and limitations 48

5	NR bottom-up benchmarking –programme level review	51
5.1	Introduction	51
5.2	Programme objectives & strategy	52
5.3	Programme approach	55
5.4	Selection of comparator organisations	57
5.5	Programme data	62
5.6	Programme outputs	66
6	NR bottom-up benchmarking – workstreams & data analysis	70
6.1	Introduction	70
6.2	Track benchmarking	71
6.3	Signalling benchmarking	81
6.4	Telecoms benchmarking	91
6.5	Electrical Power & Fixed Plant benchmarking	94
6.6	Civils benchmarking (buildings, structures, earthworks)	98
6.7	Maintenance benchmarking	118
7	Recommendations	126

Glossary

ABB	Activity Based Benchmarking
AIS	Asset Information Strategy
AMA	Asset Management Assessment
ANOVA	Analysis of variance between groups
B&C	Buildings & Civils
CoMET	Community of Metros
CP4	Control Period Four, the regulator's current control period
CP5	Control Period Five, the regulator's next control period
CSV	Composite scale variable
DNO	(Power) Distribution Network Operator
ERTMS	European Rail Traffic Management System
E&P	Electrification and Power
Imperial	Imperial College London
IIP	Initial Industry Plan
IM	Infrastructure managers
INESS	Integrated European Signalling System
IP	Investment projects
KPI	Key performance indicator
LICB	Lasting infrastructure cost benchmarking
M	Maintenance (where appropriate)
MDU	Maintenance delivery unit
M&R	Maintenance and renewals
MUC	Maintenance unit costs
NR	Network Rail
Ofgem	Office of Gas and Electricity Markets
Ofwat	Water Services Regulation Authority
OLS	Ordinary Least Squares (regression)
ORR	Office of Rail Regulation
OVB	Omitted variable bias
PR13	Network Rail Periodic Review 2013
RTSC	Railway and Transport Strategy Centre
RVfM	Rail Value for Money Study
SBB	"Swiss Federal Railways"
SBP	Network Rail Strategic Business Plan
SNCF	"National Corporation of French Railways"
S&C	Switches & Crossings
UIC	Union Internationale des Chemins de Fer
DfT	Department for Transport

0 Executive summary

Version edited for publication

Please note: for reasons of commercial confidentiality, a small number of edits / redactions have been made to this report.

0.1 Introduction

In accordance with the Independent Reporter mandate AO/015 (issued 9th May 2011), Arup has been commissioned to undertake a review of Network Rail's bottom-up benchmarking workstreams for maintenance and renewal activities. This report presents the findings of Arup's review.

0.1.1 Background

Bottom-up benchmarking activities are of significant importance to both Network Rail and the Office of Rail Regulation (ORR). Network Rail (NR) began a bottom-up benchmarking programme in 2010, focusing on operations, maintenance and renewals activities and costs. Its aim has been to "understand differences in specific activities and their drivers, including quantifying impacts of structural factors and identifying improvement initiatives." The work is part of a wider efficiency effort within the current and upcoming regulatory periods. NR has indicated that it is aiming to use benchmarking to gain "information to drive performance and delivery of CP4 outputs" and to develop "informed and evidenced plans" for CP5.

The Office of Rail Regulation considers bottom-up benchmarking to be an important means by which Network Rail should "demonstrate to its stakeholders that its plans have fully considered available efficiencies". The ORR's approach to assessing efficiency within the context of the regulatory determination includes a combination of both top-down and bottom-up cost / efficiency analyses. For the purposes of the PR13 determination process, the ORR has indicated that it considers that the detailed and specific outputs from bottom-up benchmarking to be of particular importance, given the narrowing of the estimated "efficiency gap" between Network Rail's expenditure levels and the target efficiency "frontier".

The ORR mandate for this assignment comprises both a programme level review of the benchmarking activities in terms of overall scope, approach and strategy, and a review and analysis of benchmarking data quality, coverage, robustness and its application to Network Rail's CP5 efficiency proposals. Owing to the current status of the programme, and the fact that much of the data presented are yet to be turned into quantitative outputs, it has not been possible to address all of the points set in the way envisaged by the mandate (Section 1.3 provides further details).

0.1.2 Approach

The purpose of this assignment is to provide assurance to the ORR by assessing the robustness of NR's benchmarking approach, analysing the quality and comprehensiveness of the data obtained by NR and reviewing the application of these data to NR's CP5 efficiency proposals.

Arup's approach has combined direct engagement with project participants and stakeholders within NR, independent analysis of data and processes and provision of recommendations and feedback, drawing upon our understanding of experience within other regulated industries and best practice.

We have also undertaken an analysis of benchmarking experience in other regulated industries. In particular, we have focused on UK utility sectors in which benchmarking is an established practice. We have compared these various approaches with NR's programme, and have sought to identify areas of improvement and best practices that can incentivise efficiency improvements.

We have undertaken a detailed review of documentation and data provided by NR, which has included, where available, NR project documentation, original benchmarking data, presentations and reports and information relating to asset policies. We have also reviewed third-party data (studies, publications etc.) informing NR's programme and our own analysis. A full list of external documentation reviewed is included as **Error! Reference source not found.**

0.2 Bottom-up benchmarking approaches in other sectors

0.2.1 Overview

Benchmarking techniques are utilised in a range of regulated industries to improve understanding of firms' relative cost efficiency and performance levels. We have reviewed benchmarking approaches used by regulators and companies in other (non-rail) UK regulated industry sectors - water supply, electricity distribution (UK), gas distribution (UK), and electricity and gas transmission (UK) - as well as an international benchmarking programme involving metro rail organisations. Drawing upon experiences in these areas, we have sought to identify key issues that we consider of relevance to NR and the ORR in relation to potential benefits and best practice.

A full review of benchmarking approaches for each sector is provided in Chapter 0 of this report. In addition, we provide a review of the principal elements of effective benchmarking groups in Chapter 4.

0.2.2 Relevant themes for NR and the ORR

Based on our review of benchmarking within other (non-rail) regulated UK market sectors, we have identified the following themes which we consider to be of key importance for NR:

- **Alignment of top-down and bottom-up approaches:** we have found that in other regulated UK industry sectors, bottom-up activities and top-down analysis are clearly linked from a regulatory perspective. A key area of focus within our review of NR's bottom-up benchmarking programme is the linkage between bottom-up activities undertaken (and the results obtained within individual workstreams), and the top-down and econometric analysis of relative efficiency at the programme level. We recommend within this report that NR implements measures to coordinate its bottom-up benchmarking data collection and analysis with the top-down assessment of the efficiency gap, gaining a comprehensive and coordinated overview of factors influencing relative efficiency levels, and integrating such factors into CP5 efficient expenditure proposals on a quantified basis. This recommendation is detailed (with reference to Recommendation 2011.BUB.1) in Section 0.9 below.
- **Robust forward-looking capex planning:** a key feature of bottom-up capex benchmarking in other industry sectors (e.g. water) is the level of information provided within companies' forward-looking planned capex programmes. Such programmes can facilitate effective benchmarking of bottom-up capex expenditure, by providing a clearer, longer-term overview of the linkage between expenditure levels and outputs specified within plans. This can complement the understanding of relative cost and efficiency levels gained through benchmarking activities retrospectively, enabling capex and outputs to be benchmarked against other infrastructure organisations on a more stable and consistent basis, as well as providing a means by which actual performance vs. plan can be gauged. We consider that for NR the development of stable, forward-looking, output-based projections for its renewals expenditures represents a major opportunity for improving visibility of renewals capex and enabling effective forward-looking benchmarking to be undertaken.¹
- **Consideration of impact of long-term structural change:** we consider that the implications of structural change for NR regulatory benchmarking warrant consideration from a medium to long term perspective. The most obvious structural factor likely to influence the regulatory process relates to NR's proposed organisation devolution. We recommend that the ORR and NR assess the implications of:
 - Opportunities to utilise cost and efficiency data benchmarked between NR operating entities to leverage competitive pressure and incentivise efficiencies.

¹We note that this is currently being explored in relation to Arup's other Independent Report mandates that relate to efficiency reporting.

- Risks in relation to continuity of previous benchmarking data, restructuring of cost and efficiency accounting, which may compromise the historical comparability of benchmarking data.
- **Long-term basis for benchmarking activities:** experience from other industry sectors, particularly international metro benchmarking, has shown that long-term, embedded, participative approaches to benchmarking have yielded more significant and meaningful results than short-term, limited or “one-off” studies or analyses. We consider that establishing and maintaining sustained contact with comparators is an essential prerequisite for the exchange of meaningful cost and efficiency data. This will take time and committed effort. NR’s bottom-up benchmarking should, on the whole, still be regarded as being at a comparatively early stage. Whilst NR’s programme has the potential to provide useful inputs for the PR13 process, we consider there will be limitations in the level and depth of benchmarking that can be achieved within the remaining PR13 timeframe. Subject to the necessary long-term commitment and resourcing of the programme on NR’s part, its bottom-up benchmarking should be considered not only in relation to the PR13 process, but also in terms of the longer-term value it may bring through CP5 and beyond. We set out what we regard as the key elements of effective international benchmarking groups to support this process in the next section and discuss this in greater detail in Chapter 4.

0.2.3 Key elements of effective international benchmarking groups

Alongside the review of benchmarking approaches within regulated UK utilities sectors, our review has included an analysis of effective international benchmarking groups. This has drawn upon on Imperial College London’s experience with the CoMet and Nova metro rail benchmarking consortia.

We set out below factors which have supported the successful long term operation of the CoMet and Nova benchmarking programmes (which we review in detail in Chapter4). These include:

- Confidentiality and trust supported by robust agreements.
- Effective collaboration.
- Speed of information exchange.
- Independence of participants from third-party interests.
- Properly resourced central co-ordination.
- Long-term approach / continuous annual cycles.
- Resources and Senior Level Support.
- Selection of member peers sufficient in number and representative of efficient practice.
- Clear structuring and hierarchy of benchmarking data systems.

- Face-to-face communication.
- Broad scope of benchmarking comparisons, supported by working groups.

In overall terms, the evidence indicates that a benchmarking programme will be successful if it has a clear purpose and objectives and if it follows basic principles that enable a free-flow of information and data.

0.2.4 Inherent challenges and limitations of bottom-up benchmarking

Short-term (up to PR13)

Notwithstanding our observations and recommendations above, there are significant challenges and limitations inherent within bottom-up benchmarking activities, particularly when viewed within the context of the PR13 determination process and timescales.

Experience indicates that successful benchmarking interaction requires both mutual interest and commitment of participants and time. Collating and interpreting benchmarking data itself is unlikely to be a simple and straightforward process, particularly when data are sourced from organisations operating in different countries under different organisational structures and regulatory regimes.

A related challenge is that the gaining of cooperation and trust of comparator organisations requires a significant level of discussion and personal interaction. Parties need to discuss the objectives of such activities, and clarify the nature of interaction and are likely to need to allay concerns relating to sensitive benchmarking information. Imperial College London's experience is that, typically, around two years of dialogue are needed before organisations become engaged in the process of exchanging data (although this may vary).

Medium to long term

In the medium to long term (i.e. beyond PR13 timescales), whilst the potential for meaningful benchmarking data will increase, the degree to which the data and outputs from such a programme will be useful to the ORR remains unclear. Underlying problems relating to limited data may persist if NR continues its programme approach on the present basis, because quantification and analysis of NR's overall relative efficiency level are not the primary objectives of the workstreams, and hence activities are not oriented toward this specific purpose.

Establishing a more formalised benchmarking group could enable a greater degree of more detailed comparative cost data to be shared. However, it is likely that such activity would need to be undertaken on the premise of strict confidentiality and non-disclosure outside the participant group, which could also by definition preclude disclosure by NR of the data obtained to any other party.²

² We understand that NR is currently exploring the possibility of establishing a formalised benchmarking club with other heavy rail operators. We understand that NR's concept would be to develop such a club on a similar basis to the CoMET and Nova benchmarking clubs— although these proposals are still in the very early stages. Whilst we consider the development of such a club would be likely to have positive implications from an overall efficiency perspective, and indeed we highlight in Chapters 3 and 4 of this report a number of factors likely to support

Reporter opinion on challenges associated with bottom-up benchmarking

In light of our understanding of the challenges associated with bottom-up benchmarking and the progress made to date by NR, we consider there is a significant risk that bottom-up benchmarking outputs will be not available in relation to some parts of NR's business in time to inform the preparation of NR's SBP and the ORR's determination at PR13.

More time is needed to establish the benchmarking relationships that are likely to yield more significant meaningful benchmarking results. The PR13 timescales can be considered comparatively short-term in benchmarking terms. The level of progress to date, as documented in this report, is limited.

We recognise NR has adopted a range of approaches to seek to gain understanding of relative cost and efficiency factors within the time available, which we review in detail. We make recommendations throughout this report, to support both the ORR and Network Rail in maximising the benefits that can be gained from the programme.

0.3 NR bottom-up benchmarking –programme structure and objectives

0.3.1 Programme overview and structure

NR proposes to combine and reconcile bottom-up and top-down benchmarking analyses to achieve its goal of “identifying internal and external best practice and improving business performance.” NR believes that whilst the top-down benchmarking can provide an understanding of overall efficiency levels relative to comparators, identification of specific factors and practices will enable it to help close the efficiency gap, deliver CP4 targets and inform CP5 planning. The bottom-up benchmarking programme is being delivered through ten asset-specific workstreams, which fall into four larger categories. These are:

- Track (single workstream)
- Signalling, power and telecommunications, which includes:
 - Signalling
 - Telecoms
 - Electrical Power & Fixed Plant
- Civils and structures, which includes:
 - Civils costs benchmarking (led by IP team)
 - Civils best practice benchmarking (led by IP team)
 - Structures benchmarking (led by AM team)

successful benchmarking activities through such activities, such a forum may be of limited use for the ORR from the perspective of providing evidence substantiating the efficiency gap, given that confidentiality arrangements may preclude “third party” access to benchmarking data.

- Buildings benchmarking (led by AM team)
- Earthworks benchmarking (led by AM team)
- Maintenance (single workstream)

All of the workstreams share the same high level milestones which relate to development of CP5 efficiency initiatives in each area to support NR's SBP submission. NR has stated that progress against these high level milestones is reviewed on a weekly basis as part of the wider PR13 asset management weekly review meeting, and that the workstream teams also meet on a monthly basis at a benchmarking forum to share learning and coordinate activities.

However, the gathering and analysis of bottom-up benchmarking data is being undertaken on a decentralised basis with each workstream using or adopting an approach that it considers most appropriate for its particular asset area. Activities and deliverables (engagement with comparators, completion of analyses and reports) reflected in the programme plan are structured on an individual basis for each workstream. Specific activities or milestones often appear not to be closely linked, and the degree of progress and level of data achieved varies between them.³

0.3.2 Programme objectives

Our review has found NR's teams involved in the benchmarking workstreams are well motivated and working with a clear focus on identifying tangible efficiency opportunities.

Whilst NR has highlighted understanding of the efficiency gap at the programme level as an overall goal of its benchmarking activity, we have not had sight of specific objectives reflecting this goal at the workstream level. Activities at the workstream level are geared towards comparing activities and practices and identifying efficiency measures, rather than substantiating NR's relative efficiency level. There is a general lack of a significant pool of quantified numerical observations. This reflects in part the limitations inherent in bottom-up benchmarking activity. We consider there is potential for a greater number of quantified observations to be gathered to support the assessment of NR's efficiency level going forward, e.g. by engaging further with comparator organisations to benchmark a wider range of activities. (This is discussed further in Section 0.8 below.) To support this process, we recommend that NR defines as a formal objective for each of its benchmarking workstreams the requirement to capture numerical data to support quantification of its efficiency gap.

0.4 NR engagement with comparator organisations

NR has identified a range of different comparator organisations through its respective workstreams and sought to establish meaningful contact with them. The degree of benchmarking data sharing has been, for the most part, limited.

³ NR has indicated that it consider the decentralised programme approach to be in line with a key finding from the top down benchmarking analysis conducted by NR in collaboration with ORR and the Institute of Transport Studies which confirmed that the key structural factors influencing maintenance and renewal spend vary widely across assets.

We discuss below our findings in relation to the nature of NR's engagement with comparator organisations, taking into account what we see as the inherent difficulties in carrying out bottom-up benchmarking of this nature before focusing specifically on the nature of data obtained.

0.4.1 Comparator selection

NR's identification of comparator organisations for bottom-up benchmarking began with the comparison of high-level maintenance and renewals expenditure figures from other European national infrastructure managers.

NR undertook a detailed process of normalisation and analysis of the high-level cost data from the LICB dataset, with the aim of identifying comparators that are managing their networks at lower cost. On this basis, NR has been able to establish contact with seven European national rail infrastructure managers identified as appropriate comparators through this analysis - although the level of engagement and cooperation has varied.

NR also has identified a number of non-heavy rail comparators including light rail or metro rail infrastructure operators, operators of non-rail transport infrastructure, utilities and telecoms companies, and other operators of built assets (e.g. public buildings; car parks).

Arup / Imperial has supported NR by providing contacts a number of other infrastructure operators globally.

We consider NR's process for identifying and establishing contact with comparators, including identification of efficient rail comparators via top-down analysis through analysis of the LICB dataset to be based on sound analysis. NR has for the most part targeted those rail organisations identified as efficient for comparison.⁴ We consider NR should seek to extend the selection of rail organisations with whom benchmarking activities are being undertaken to include a greater range of comparators outside Europe. A number of potential benchmarking contacts in other rail organisations have been provided by Imperial / Arup.

0.4.2 Nature of engagement with comparators to date

Whilst NR has coordinated engagement with comparators centrally, taking care to avoid multiple uncoordinated contacts into the same organisations, each bottom-up benchmarking workstream in practice has had to engage with its comparator organisations on an individual basis (see Chapter 6). The nature of interaction and degree of information and data shared / obtained differs between the workstreams. No formal club or forum (new or existing (e.g. LICB)) has been used. Workstream leads do routinely discuss their experience and share and coordinate contacts made where appropriate, and NR also has a central database of all visit reports.

⁴ NR has also indicated that German rail infrastructure operator Deutsche Bahn (DB), not presently part of the LICB dataset, is a very active member of the RTC benchmarking group. NR considers that, in demonstrating the benefits that this brings for Telecoms, it may well lead to wider collaboration with DB across different functions

NR has invested considerable time and effort to engage with comparator organisations. The level of quantified benchmarking data obtained so far has been limited. As discussed in the previous section, it is likely that this is in part due to the comparatively recent timescales for many workstreams and the time that is likely to be needed to establish the relationship, identify and facilitate data, gain understanding of cost factors, etc. Political sensitivities also affect the behaviour European rail organisations; a general reluctance exists to share cost-related data, with concerns that information may be utilised by government or regulatory bodies.⁵ In general, comparator organisations have so far shown a reluctance to share cost-related data in the first instance.

As a result NR has had to take a cautious approach in engaging with comparators, avoiding demands for detailed cost data and instead developing a number of methodologies (see Section 0.5) which draw upon information that is likely to be seen as “less sensitive” by the comparator organisations .

Overall, we consider NR has taken a proactive approach to engaging with engagement with comparator organisations. Personal interaction and in-depth meetings and discussions have formed an essential basis for sharing data and for potentially establishing a benchmarking relationship in the longer term.

To support a stronger level of comparator engagement, we recommend that NR develops plans to deepen and improve comparator engagement, and maintain such interaction on a longer-term basis. Mutually beneficial interaction, including sharing by NR of results and analysis of its programme with partner organisations, is an essential element of comparator engagement. (See recommendation 2011.BUB.2).

0.5 NR’s approach to collecting and analysing benchmarking data

Teams leading the respective workstreams are largely responsible for establishing their own approach and identifying those cost / efficiency aspects they see as most relevant for their particular area.

Bottom-up benchmarking workstreams are being led primarily by teams from NR’s Asset Management division with experience in front-line management. We consider the decentralised structure of the benchmarking programme to be appropriate. Team members’ knowledge and understanding of cost drivers and potential efficiency measures gained through front-line delivery of the asset are important factors in informing the benchmarking approach taken for their particular area.

Because of difficulties in gaining direct comparative cost data, In light of the inherent limitations and difficulties in gaining direct comparative cost data, NR’s benchmarking teams have adopted a range of alternative comparative

⁵ European rail organisations’ reluctance for cost-related information to get into the hands of regulatory / government bodies appears to be reflected in recent departures from the LICB dataset by operators such as Deutsche Bahn. Conversely, the weakening of the LICB could logically lead to potential increased attractiveness of an alternative forum such as that proposed by NR, if – as Imperial College London’s experience with metro rail benchmarking suggests – participants can see the benefits of a “closed” club where the confidentiality of data sharing can be guaranteed.

methodologies across the different workstreams to gain an insight into relative cost levels and efficiency factors, including:

- Comparing individual jobs / activities on a site-by-site basis, including quantitative comparisons of time and resourcing and qualitative assessment of processes and procedures;
- Obtaining cost estimates from comparator organisations for a pre-defined package of work; and
- Analysing policies and strategies for managing and delivering rail infrastructure, focusing on how these drive activity scope.

On the whole, we have found that NR benchmarking teams consider this to be a long-term process, the results of which will remain relevant through CP5 and beyond. This assumption appears reasonable. Although difficult, we consider in the context of the PR13 determination process that the ORR would benefit from gaining greater clarity with regard to the nature of information NR expects to be able to obtain, its timing, and how the data will be used for the purposes of supporting the development of CP5 efficient expenditure proposals. For example, it is not yet clear what quantitative data NR will collect from many of the comparators with which it is in discussion. Nor is it clear at this stage, what performance indicators might be calculated with data, or how data might have to be normalized in order to present relevant comparisons.

We provide a detailed review of the approach taken for each of the benchmarking workstreams in Chapter 6 of this report.

0.6 Bottom-up benchmarking data review

0.6.1 Nature of data obtained

NR has acquired a mixture of qualitative and quantitative comparative data through its bottom-up benchmarking programme, with varying levels of detail and differences in the overall scope of information obtained. When its programme commenced, NR articulated a number of high-level expectations about the areas to be explored by workstream leads. Detailed specifications relating to the types or volumes of data required for individual workstreams have not been provided.

NR has been able to obtain direct comparative cost data in one asset area (civils). Some workstreams (such as track and maintenance) have obtained quantified (non-cost) benchmarking data or observations that have been used to inform quantified CP5 efficiency proposals. For other workstreams, such as buildings and telecoms, we understand no quantified data have yet been obtained.

Overall, the quantified benchmarking data obtained to date is not yet of a sufficient volume or level of detail to meaningfully support the high-level quantified analysis of NR's relative efficiency level or the efficiency gap.

We recommend that NR should explore means through which the quantity of data and detail are increased from its existing comparator contacts (e.g. by increasing numbers of observations). As part of this work, NR should develop a set of coordinated benchmarking data specifications / criteria, taking programme level

requirements (in relation to any overall efficiency gap) and setting these out as explicit data requirements for each workstream.

0.6.2 Workstream data progress

The overall level of progress in obtaining and analysing bottom-up benchmarking data for each bottom-up benchmarking workstream is summarised below.

- **Track:** detailed, qualitative sets of site-based observations have been obtained through around fifty individual site visits, and a range of salient efficiency factors have been collated and explored. NR has presented quantified proposals for achievement of efficiency in S&C renewal rates over CP5 (accounting for approximately a quarter of total track renewals expenditure). We understand NR is presently developing similar proposals for both conventional and high-output Plain Line renewals.
- **Signalling:** of the workstreams we have reviewed, signalling has adopted the most comprehensive approach of any area. Benchmarking data have been obtained through three distinct workstreams combining different methodologies. A detailed set of initial efficiency opportunities has been presented. The outputs presented have been present in qualitative terms, and for confidentiality reasons we have been unable to review quantified cost data obtained through LICB benchmarking. We understand that NR is proposing to progress the workstreams in order to develop a more quantified insight in terms of efficiency.
- **Telecoms:** one of the less advanced workstreams, with no benchmarking data yet provided for review. We consider that the proposed approach, focusing on telecoms from a service-based perspective, to be appropriate and reflective of the evolving nature of telecoms provision.
- **Electrification & power:** another of the less advanced workstreams for which, like telecoms, no benchmarking data have yet been provided for review.
- **Buildings and civils:**
 - **Civils costs benchmarking (led by IP team):** one of the more advanced workstreams, with detailed analysis and breakdown of costs comparing NR with three comparators across three schemes. We consider the methodology adopted provides useful insights into relevant cost factors, but also suggest that the scope of this analysis should be widened to include further peers across a wider range of schemes to increase the explanatory power and robustness of results.
 - **Civils best practice benchmarking (led by IP team):** a wide-ranging workstream with potential for identifying the impact of relevant asset management, project and procurement-related factors. The way in which the initial findings have been analysed and applied to actual expenditure proposals is unclear from the version of the report provided (from March 2011). We understand work is ongoing.

- **Structures benchmarking (led by AM team):** one of the less advanced workstreams for which no benchmarking data have yet been provided for review.
- **Buildings benchmarking (led by AM team):** another of the less advanced workstreams for which, like structures, no benchmarking data have yet been provided for review.
- **Earthworks benchmarking (led by AM team):** This comparatively new workstream has involved only a limited level of qualitative benchmarking based on observations from two site visits, NR has already drawn upon the results to explore potential implementation of remote condition monitoring as a key efficiency initiative. The degree to which the workstream will progress to a level that enables quantified benchmarking to be undertaken is presently unclear.
- NR has provided a high-level breakdown of CP5 efficiency factors applicable to both buildings (operational property) and civils, on the basis of which overall unit cost efficiencies set out in the IIP have been projected, with a quantified estimation of each efficiency measure listed.
- **Maintenance benchmarking:** quantified comparative analysis has been undertaken using an activity based benchmarking approach, which has produced direct comparisons of cost-rates (plant, labour, materials) and outputs for specific activities. This has also supported analysis of wider factors relevant to efficiency such as policy, planning, competencies of personnel and contracting structures. The scope of benchmarking undertaken using this methodology has been limited to a handful of individual visits. We recommend extension of study scope to include additional comparators and case studies for the given type of activity. It is not evident how far specific efficiency initiatives have been developed for the maintenance workstream.

For a detailed review of the above workstreams please see Chapter 6.

0.6.3 Numerical benchmarking data assessment

Numerical data coverage

We set out in *Table 1* an overview of the areas of expenditure across the different categories provided to Arup for review.

Expenditure category	Annual expenditure level (FY 11/12) (£m)	Numerically benchmarked activities	Indicative coverage level (£m / % overall spend)	
Renewals expenditure				
Track renewals	605	- S&C renewal - High-output track renewal	251 ⁶	42% ⁷
Signalling	373	- ERTMS renewal	18	5%
Telecoms	248		0	0%
Electrical power & fixed plant	108		0	0%
Civils (structures & earthworks)	356	- Renewal of metallic underbridge ⁸ - Renewal of footbridge - Earthwork renewals works	115	32%
Operational property	272		0	0%
Other renewals categories	272		0	n/a ⁹
Total renewals expenditure	2,234		385	17%
Maintenance expenditure				
Total maintenance expenditure	1,068	- Re-railing - Re-sleepering	46	4%

Table 1 - indicative coverage level of bottom-up numerical benchmarking data

The coverage figures provided in **Table 1** are to be considered as indicative. We have not been able to form a definitive judgment of coverage levels. Further assessment is needed by NR to examine the degree to which benchmarking data are sufficiently representative (e.g. taking into account comparator data depth, breadth and comparability).

Numerical data quality and robustness

We summarise the results of our assessment of data quality and robustness in **Table 2** overleaf:

⁶ We note that Arup has not been given sight of quantified comparative data obtained for track renewals.

⁷ Indicative coverage figure assumes that High-Output track renewal accounts for 25% of total Plain Line renewal expenditure.

⁸ Indicative coverage based on estimated renewals expenditure on metallic underbridges of 40% of total underbridge renewals expenditure.

⁹ Other renewals categories have not been reviewed by Arup, therefore the level of data obtained for such categories through the bottom-up benchmarking programme is unknown.

Workstream	Data quality	Data robustness
Track	N/A – data not provided for review.	Limited. Arup recommends increased number of comparators and observations*.
Signalling	N/A – data not provided for review.	Reasonable. Arup considers the dataset size, scope and number of participants should be sufficient to ensure a reasonable level of robustness*. ¹⁰
Civils (structures & earthworks)	Reasonable. Data sourcing methodology described, normalisation and adjustment calculations demonstrated. On this basis, Arup considers the quality of civils benchmarking data appears be reasonable. ¹¹	Limited. Arup recommends increased number of comparators and observations.
Maintenance	Reasonable. Source observations of activity time and resourcing individuals’ site observations demonstrated. Data breakdown, adjustment and allocation of cost factors clearly set out. Demonstrates reasonable data quality.	Limited. Arup recommends increased number of comparators and observations*.

Table 2 - Arup high-level assessment of data quality and robustness

* Note: Recommendations for increased robustness of comparator data for each asset area can be found in the respective sections of Chapter 6.

Whilst the methodology adopted provides useful insights into relevant cost factors, the scope of this analysis could be widened to include further peers across a wider range of schemes. This would increase the explanatory power and robustness of results.

We recommend that NR specifies the comparative cost / efficiency data it expects to obtain going forwards, and its judgment as to the usefulness of such data for the CP5 determination process and its contribution towards the understanding and assessment of the efficiency gap.

¹⁰ This assessment is based on descriptions of the nature and scope of numerical data provided by the signalling workstream. Arup has not been provided with numerical benchmarking data for direct review for this workstream.

¹¹ This assessment is based on NR’s presentation of adjusted figures, together with explanations of normalisation and adjustment factors applied. Arup has not been provided with original source cost data provided by the comparator organisations.

0.7 Application of bottom-up benchmarking programme outputs to CP5 efficiency proposals

0.7.1 NR CP5 efficiency proposals set out in the IIP

NR's bottom-up benchmarking programme has been established not only to "leverage" off the elements contributing to the efficiency gap but also to identify positive drivers and initiatives to improve the organisation's efficiency and performance. Arup's mandate is based around the following two NR submissions, which represent key milestones within the overall PR13 determination process:

- The Initial Industry Plan ("IIP"), published in September 2011, sets out NR's current expenditure and efficiency proposals.
- The Strategic Business Plan ("SBP"). To be published in January 2013, this will set out NR's expenditure and efficiency proposals in detail.

The IIP projections draw upon "top-down" projections of NR efficiency for CP5, drawing upon target efficiency levels set out in the McNulty Rail Value for Money (RVfM) study and PR08 determination as the basis for efficiency proposals.

In forming its judgements of projected efficiency, NR has also drawn upon the emerging findings from the bottom-up benchmarking programme. Alongside the main published IIP document, NR has provided supporting documentation that includes qualitative assessments of a number of efficiency initiatives and improvements identified through the bottom-up benchmarking programme. The IIP sets out proposals for the achievement of an efficiency level over CP5 of 16% compared to CP4.

Arup has only been provided with a quantified breakdown linking IIP efficiency projections with bottom-up benchmarking findings for two asset areas – track and civils – which we discuss further below.

We understand that NR intends to substantiate further the existing IIP efficiency projections and to identify potential opportunities for efficiency above and beyond the 16% level as it develops its proposals for the SBP over the next year.

We recommend that each benchmarking workstream provides a detailed explanatory document setting out in full the numerical calculations through which bottom-up benchmarking outputs are applied to the relevant CP5 efficiency projections, together with the key assumptions and rationale supporting such calculations. This document should develop and evolve as benchmarking activities progress and NR further develops its efficiency proposals. We note that the track team provided its model showing its efficiency calculation after we completed our review. We may be able to review this model in detail in due course as part of our progressive assurance process (subject to discussions with NR and the ORR).

0.7.2 CP5 efficiency proposals based on bottom-up benchmarking outputs

As indicated above, NR has provided quantified breakdown of CP5 efficiencies based on bottom-up benchmarking findings for track and buildings & civils.

Track

NR has developed a quantified assessment of efficiency to be achieved for S&C renewals over the course of CP5. S&C renewals account for approximately one quarter of total track renewals expenditure.

NR has set out in qualitative terms a number of factors that will result in improved S&C renewals efficiency. These factors relate predominantly to improved contractor management. NR projects a 23.4% unit cost efficiency for S&C renewal, derived primarily from reduced contractor site costs expected as a result of the measures identified. Arup has not been provided with the numerical calculations or assumptions on which this figure is based, so we cannot give an assessment of how bottom-up benchmarking was numerically factored into the efficiency figure presented.

NR's S&C efficiency proposals represent the most substantively quantified proposals for CP5 that we have reviewed. Further analysis is needed to substantiate proposals, assess risk (e.g. associated with delivery) and firm-up implementation costs. It will also be essential for NR to continue to engage directly with contractors and the broader industry to ensure buy-in to the efficiency proposals as they develop.

Civils

NR has provided a high-level numerical breakdown of the projected IIP scope and unit cost efficiencies for buildings and civils assets. NR has presented a number of asset management-related measures to support the projected 5% scope efficiency. For the 11% projected unit cost efficiency, NR presented a quantified estimation of the efficiency impact of eight specific efficiency measures.

We consider that the level of efficiency gains associated with the proposed measures appears reasonable. We have not been provided with detailed calculations and assumptions upon which the estimated efficiency factors may be based.

0.8 Conclusions

0.8.1 Applicability of current bottom-up benchmarking data for CP5 efficiency projections

We have reviewed in detail NR's bottom-up benchmarking programme, including the approach of individual workstreams, the nature of benchmarking data obtained and the outputs generated by NR so far to support its CP5 efficiency proposals.

We consider that the data and outputs we have reviewed to date would be of limited use to the ORR in the context of its CP5 determination. We regard the numerical benchmarking data to be limited, both in terms of coverage (with numerical comparisons obtained for only a limited proportion of M&R activities) and in terms of robustness (with only a small number of comparisons for any given cost element).

With respect to numerical data quality, we have reviewed the maintenance benchmarking data and have concluded that this is of a reasonable quality. Arup has not been given sufficient access to data during the timescales of this assignment that allows us to make a definitive assessment of numerical data quality for any other asset area.

Presently, we consider the bottom-up benchmarking data insufficient in form and scope to meaningfully inform a gap analysis or other quantified assessment of NR's relative efficiency level in the asset areas reviewed.

0.8.2 Applicability for PR13

We summarise in this section our assessment of the level of bottom-up benchmarking data and outputs that are likely to emerge in the run-up to PR13, and their applicability to the ORR's PR13 determination of CP5 efficiency.

Our assessment has been based on the following two key assumptions:

- NR will continue its bottom-up benchmarking activities on a generally consistent basis for the remainder of the period (i.e. workstreams will continue their activities as set out in the programme plan, programme resources will remain in place).
- NR will make reasonable endeavours (within current resourcing and time constraints) to adopt Arup's recommendations relating to widening comparator engagement and the maximisation of the scope and breadth of observations it gains through its interactions.

On the basis of the above assumptions, we expect NR's bottom-up benchmarking programme to produce data relating to a greater number of expenditure areas than at present – with data likely to be obtained to enable quantified comparisons to be made for at least a proportion of expenditure elements within each of the asset areas covered by the programme. We expect this will enable NR to develop and substantiate CP5 efficiency proposals for a greater number of expenditure elements. We also expect NR to be able to increase the number of numerical observations which will increase the robustness of its benchmarking analysis.

We do not consider it possible at the present time to give a quantified estimation of likely coverage or applicability as this will depend on the scope and breadth of data / observations gained over the next year. NR has not defined in quantified terms the specific areas of expenditure that it expects to cover over this period. Our expectation is that the scope and coverage of bottom-up benchmarking data will remain limited, not least due to the inherent challenges involved in engaging in benchmarking activities of this nature that we describe above. It is therefore unlikely that a bottom-up benchmarking dataset will be in place that covers the majority of M&R expenditure.

With regard to quantified gap analysis or other high-level assessment of NR's relative efficiency level in any given asset area, we do not anticipate the bottom-up benchmarking programme to yield data of sufficient scope or breadth to inform such analysis in a comprehensive way.

Overall, we recommend that ORR considers alternative comparative analytical approaches to support its assessment of NR's relative efficiency level and CP5 targets. This is discussed further below.

0.8.3 Potential alternative approaches for efficiency analysis

We discuss in this section other potential approaches for the ORR to analyse NR's relative efficiency level and CP5 targets, as an alternative to usage of data / analysis of NR's bottom-up benchmarking programme.

We consider that initiation of benchmarking activity with European rail transport organisations by the ORR is unlikely to be feasible, due first and foremost to practical limitations in undertaking such activities (as being experienced by NR), as well as the reluctance of comparator organisations to engage with ORR as a regulatory body.

However, we do consider there to be potential for ORR to explore the following:

- Analysis of efficiency potential using internal NR cost data.
- Qualification of renewals capex according to quality of proposals.
- Bottom-up engineering models for technical comparison.

Analysis of efficiency potential using internal NR cost data

Drawing upon our prior understanding of NR's business and internal reporting, we consider there is potential for ORR to undertake its own targeted benchmarking analysis utilising comparative cost data from within NR. It should be feasible for data to be sourced to a considerable level of detail, and for this data to be cut and compared in a number of different ways. ORR may be able to draw upon a range of techniques including regression analysis, statistical calculations and qualitative reviews. The cooperation of NR itself would be an essential prerequisite for such benchmarking techniques to be taken forward. We recommend that ORR explores potential internal cost benchmarking options with Network Rail further.

Qualification of renewals capex according to quality of proposals.

Another approach to support the ORR's analysis of efficient NR expenditure in CP5 could be based around the ORR making capex funding availability contingent on the quality and robustness of its renewals expenditure proposals. Drawing on the experience of water industry regulator OfWAT, there may be potential for ORR to explore the analysis of NR's CP5 renewals expenditure proposals for each asset area in relation to a number of criteria such as:

- stakeholder engagement;
- leadership;
- policy and strategy;
- management;
- processes;
- systems;
- data;
- analysis; and
- reporting.

Bottom-up engineering models for technical comparison

We recommend that ORR explores the efficient cost and delivery of rail asset maintenance and renewals outputs utilising bottom-up engineering models based

on engineering judgment of technical specifications and output requirements of the specific infrastructure to be delivered, and the associated costs. This has been explored by OfWAT. Such an analysis could be particularly useful when assessing the efficient cost of delivery of a specific type of asset or asset group with particular characteristics, for which the availability of comparative cost data from other asset organisations is limited. In addition, such analysis could be utilised to challenge from a technical perspective the strength of the needs case made for the given activities, to “chip away” at capital and operating expenditure projections.

0.8.4 Recommended next steps for IR review

The ORR may need to consider further independent review of NR’s bottom-up under this mandate. A possible approach could be for emerging data and outputs to be reviewed on quarterly basis.

To ensure an efficient review process and to avoid duplication we would suggest focusing only on new data and analysis that emerge. These reviews should enable the ORR and NR to assess the relevance and the degree to which such data can be utilised for the purposes of PR13

As well as reviewing any new material from the programme, it is recommended that parts of the original mandate that it has not been possible to fully address during this review (see Section 1.3) are considered for inclusion.

0.9 Recommendations

We set out in *Table 3* below our -level recommendations in relation to this assignment. We draw together recommendations from both within the Executive Summary, and throughout the remainder of the report. We categorise our recommendations under the following six sub-headings:

- Alignment of objectives
- Engagement with comparators
- Benchmarking dataset
- Benchmarking data analysis
- Workstream-specific recommendations
- Alternative benchmarking approaches for analysis by ORR

Ref.	Recommendation to Network Rail
<i>Alignment of objectives</i>	
2011.BUB.1	<p>We recommend that a requirement to obtain data to support and substantiate the analysis of NR's efficiency gap is defined as an explicit objective for each workstream.</p> <p>This should influence the approach taken, the scope of analysis and the utilisation of the data obtained by the given workstream. It should also help to improve visibility of efficiency factors identified by NR (including initiatives already identified) and may provide additional insight in relation to how and why a given efficiency factor / opportunity is being prioritised, thereby helping justify its implementation.</p>
<i>Engagement with comparators</i>	
2011.BUB.2	<p>We recommend that NR develops a detailed engagement plan for each workstream, setting out specific steps through which it plans to initiate, establish and maintain contact. This should include both proposals for maintaining long-term, mutual engagement, and implementing procedures that ensure outputs are shared and mutual interests served, thereby maximising the prospects for obtaining meaningful data.</p>
2011.BUB.3	<p>We recommend that NR develops ties with further comparator organisations and looks more widely than the present comparator pool which is heavily oriented towards European rail organisations. In particular, we consider NR should focus on benchmarking with light-rail / metro organisations both within the UK and internationally, for which we consider there to be a significant level of potential comparability in spite of differences in infrastructure characteristics. We also consider NR should explore contacts with non-European heavy rail organisations. (Note: recommended comparators and contact details have been provided by Arup/Imperial through feedback sessions provided for each workstream. Further details to be provided in Chapter 6).</p>

Ref.	Recommendation to Network Rail
<i>Benchmarking dataset</i>	
2011.BUB.4	We recommend that NR should explore means through which the amount of data - and the level of detail - can be increased from existing comparators. This should include exploring means through which existing datasets can be deepened and more observations obtained, e.g. through a greater number of activities, increased use of questionnaires / automated data sharing.
2011.BUB.5	We recommend that NR develops a set of coordinated benchmarking data specifications / criteria, taking programme level objectives (in relation to overall efficiency gap) and setting these out as explicit data requirements for each workstream. This should include criteria to ensure a sufficiently representative dataset – e.g. for breadth of data in relation to areas of expenditure, level of depth and detail, sample size, and nature of comparative data. This should also account for incremental progress as the benchmarking progresses and increasing level of data are obtained (timetable / “roadmap”).
<i>Benchmarking data analysis</i>	
2011.BUB.6	<p>Recommendation: We recommend that each benchmarking workstream lead provides a detailed explanatory document setting out the key assumptions and rationale relating to benchmarking data outputs. This should include:</p> <ul style="list-style-type: none"> • details of incoming data adjustments and normalisation; • an explanation of the process by which key cost / efficiency factors have been identified and prioritised and those factors excluded / marginalised from the analysis; and • details of the application and extrapolation of the data to higher-level expenditure and efficiency proposals / projections.
<i>Workstream-specific recommendations</i>	
2011.BUB.7	Track: we recommend NR undertakes further analysis to support its CP5 efficiency proposals for S&C renewals. This should include analysis of implementation cost relating to the initiatives proposed, and a detailed assessment of risks associated with the proposals.
2011.BUB.8	Track: we recommend that NR undertakes comparative analysis between itself and its peers by costing work banks seen overseas as if they were undertaken in the UK.
2011.BUB.9	Buildings & civils: we recommend NR undertakes further analysis to support its CP5 efficiency proposals for Buildings & Civils activities. This should include analysis of implementation cost relating to the efficiency initiatives proposed, and a detailed assessment of risks associated with the proposals.

Ref.	Recommendation to Network Rail
2011.BUB.10	<p>Maintenance: To deepen its analysis, we recommend NR to use time series data from other railways to identify where these organisations have driven down maintenance costs. This should allow NR to explore in detail whether these cost savings arise from efficiency, deferred maintenance or the impact of renewals or enhancement spending.</p>
2011.BUB.11	<p>Maintenance: to support NR’s internal maintenance benchmarking activity we suggest the merits of a Bayesian approach. To ensure transparency, we recommend that inputs & assumptions are clearly laid-out to support such analysis.</p>
2011.BUB.12	<p>Maintenance: For present and future internal benchmarking analysis of efficiency across MDUs, we recommend that NR considers the following statistical techniques:</p> <ul style="list-style-type: none"> • Semi/non-parametric estimation. • Confounding / omitted variable bias (OVB). • Normalised dependent variable. • Cost function with price data. • Use of temporal as well as cross-sectional analysis. • Analysis by category of spending. • ANOVA for grouping analysis.
<p><i>Alternative benchmarking approaches for analysis by ORR</i></p>	
2011.BUB.13	<p>We recommend that ORR considers alternative comparative analytical approaches to support its analysis of NR’s efficiency level and CP5 targets for the PR13 review process. We recommend the ORR considers the following:</p> <ul style="list-style-type: none"> • Analysis of efficiency potential using internal NR cost data. • Qualification of renewals capex according to quality of proposals. • Bottom-up engineering models for technical comparison.

Table 3 - NR bottom-up benchmarking: recommendations

1 Introduction

1.1 Background and objectives

This report presents the findings of Arup's review and audit of Network Rail's programme bottom-up benchmarking activities, intended to inform Network Rail's CP5 expenditure and efficiency proposals as part of the Periodic Review 2013 (PR13) determination process. This work has been undertaken in accordance with the Independent Reporter mandate AO/015: Audit of Network Rail's bottom-up benchmarking workstreams for maintenance and renewal (issued 9th May 2011). A copy of the mandate is included as Appendix A of this document.

The purpose of this assignment is to provide assurance to the Office of Rail Regulation (ORR) through an assessment of the robustness of Network Rail's benchmarking approach, analysing the quality and comprehensiveness of the data obtained by NR and reviewing the application of these data to NR's efficiency proposals.

Network Rail's bottom-up benchmarking activities are part of a longer-term programme to gain an understanding of comparative efficiency levels and to develop efficiency initiatives into CP5 and beyond. The programme continues to develop and evolve. Therefore our review has focused not only on benchmarking activities undertaken to date (including in relation to the recently completed IIP) but also the approach and focus of the programme going forward. We have also sought to draw on best practice and to provide recommendations to help Network Rail improve its benchmarking activities.

1.2 Our approach

Our approach entails a combination of direct engagement with project participants and stakeholders within NR, independent analysis of data and systems and provision of recommendations and feedback, drawing upon our understanding of experience within other regulated industries and best practice.

1.2.1 Meetings with programme participants

We have held a number of meetings and discussions with teams within Network Rail involved in the programme. This has included detailed discussions of activities undertaken and data obtained with each of the asset-led workstreams, as well as discussion and review of programme leaders regarding the programme strategy and objectives and the coordinating of the programme outputs within the overall context of Network Rail's CP5 expenditure and efficiency proposals currently being developed.

We attach as **Error! Reference source not found.** a list of the meetings held.

1.2.2 Documentation and data review

We have undertaken a detailed review of documentation and data provided by Network Rail to establish the scope and range of activities carried out so far and to inform our overall analysis of programme approach and outputs. This has included:

- Project documentation, setting out approach, aims, objectives of the bottom-up benchmarking activities at both programme and individual workstream level.
- Original / “raw” benchmarking data collected by NR in both qualitative and quantitative format (e.g. asset-specific observations, records / notes of site-visits, review of input figures / sources and calculations undertaken, etc.)
- Analytical presentations and reports setting out areas of focus for each asset-based workstream, including specific efficiency initiatives and proposals derived from the respective workstreams.
- Information relating to asset policies, costs and outputs, published plans (e.g. IIP and SBP), other non-benchmarking specific reports and documentation.
- Third-party sources (studies, publications, etc.) which Network Rail has drawn upon as part of its analysis.

As well as documentation provided prior to the submission of our initial draft report on 27th September 2011, this version of our report takes into account documentation and clarifications subsequently provided. Recent documents include NR’s programme level reports “International Bottom up Maintenance and Renewal Railway Benchmarking: Draft November 2011” and “CP5 Maintenance and Renewal Scope and Unit Cost Efficiencies: summary and progress report” as well as documents containing quantified breakdowns of projected CP5 efficiencies based on bottom-up benchmarking findings for track and buildings & civils assets. A full listing all of the documentation received is included in **Error! Reference source not found.. Benchmarking experience and best practice from other industries**

We have undertaken an analysis of benchmarking experience from other regulated industry sectors, with a particular focus on UK utilities sectors within which benchmarking is established practice. We have compared the various approaches taken with Network Rail’s programme and sought to identify areas of potential improvement and best practice that can be applied to support and incentivise efficiency improvements.

1.2.4 Constructive feedback and recommendations

Over the course of our review our aim has been to provide clear and constructive feedback. Drawing in particular on the extensive international benchmarking

experience and expertise from Imperial College London we have sought to develop clear, workable recommendations. This process has included:

- Programme-level feedback provided to Network Rail and the ORR on a monthly basis at Steering Group meetings.
- Detailed findings and recommendations provided at feedback meetings with each NR asset-based workstream.

We have also provided Network Rail with contact details for other organisations, including organisations outside Europe with which we consider Network Rail may be able to establish contact and share benchmarking data. A list of the contacts provided is included in **Error! Reference source not found.**

1.3 ORR mandate – reference report sections

Table 1 shows how the various sections of this report address the scope set out in the mandate for this assignment. As a consequence the extent of progress made by NR so far, and the data the company have been able to share with us, some elements of the scope have not been addressed in the way envisaged by the mandate.

ORR mandate requirement	Reference report sections
<p><i>Scope (general)</i></p> <p>Programme-level review of the benchmarking scope, approach and strategy.</p> <p>Analysis of data quality, coverage and robustness, comprising:</p> <ul style="list-style-type: none"> • external bottom-up benchmarking data analysis; and • internal benchmarking data analysis, which draws upon but does not duplicate the Reporter’s Q4 work covering MUC and CAF unit metrics. <p>Appraisal of data reworking, normalisation etc.</p> <p>Appraisal of outputs, including implications for ORR’s use of NR’s benchmarking data for its Periodic Review activities.</p> <p>Assessment of Network Rail’s methodology to incorporate the results of each benchmarking workstream into its plans for CP5.</p>	<p>0.3, 5.2, 5.3</p> <p>0.6.3, 5.5, 6 (workstream data sections)</p> <p>Maintenance: 6.7.3</p> <p>Other asset areas: Not feasible to address this requirement</p> <p>0.7, 5.6</p> <p>0.7, 6 (workstream outputs sections)</p>

ORR mandate requirement	Reference report sections
<p>Development of recommendations for programme and/or workstream changes leading up to PR13.</p>	<p>0.9, 0</p>
<p><i>Programme level review</i></p>	
<p>Assess and document the overall approach and strategy of Network Rail’s bottom-up benchmarking programme, and how this relates to Network Rail’s CP5 efficiency plans.</p>	<p>0.3, 0.5, 0.8, 5.2, 5.3</p>
<p>Assess the programme structure, timescales and linkages between respective workstreams, and how these relate to the efficiency and expenditure determination process for CP5.</p>	<p>0.3, 5.2</p>
<p>Examine the approach and defined objectives of each individual workstream, taking into account the scope, breadth and range of comparators, as well as the timescales and outputs (existing and future) of the given workstream, including in relation to PR13 timescales.</p>	<p>0.2, 0.3, 6 (workstream approach sections)</p>
<p>Assess the linkage of the programme with other studies and workstreams including:</p> <ul style="list-style-type: none"> • Unit Cost Framework. • Asset Management policies. • Other internal efficiency / best-practice initiatives. 	
<p>Review and compare the benchmarking programme approach and scope in relation to benchmarking experience from other industries.</p>	<p>0.2, 4, 5</p>
<p>Specify the key facets of an idealised benchmarking approach, based on best-practice and that which is considered appropriate and realistically achievable for Network Rail’s benchmarking work, if different, at IIP and SBP (i.e. “what does good look like?”); establish the level of improvement required within the NR benchmarking programme to achieve this.</p>	<p>0.2, 3.7, 4.2, 4.3</p>
<p>Assess Network Rail’s methodology to incorporate the results of each bottom-up benchmarking workstream and top-down benchmarking into its efficiency plans for CP5.</p>	<p>0.7, 5.6, 6</p>
<p>If appropriate, make recommendations for measures to improve the benchmarking programme, its workstreams, coverage and potential to produce useful information to inform Network Rail’s efficiency plans.</p>	<p>0.9, 6 (Reporter opinion sections), 0</p>
<p><i>Benchmarking data analysis</i></p>	
<p>For each workstream, review the appropriateness of the benchmarking approach adopted, its strengths and weaknesses and its comparison to best practice.</p>	<p>0.6.2 6 (workstream approach, reporter</p>

ORR mandate requirement	Reference report sections
<p>Carry out a review of the quality, suitability and robustness of benchmarking data (both internal and external and including the collection of data for normalisation purposes) for each key category of maintenance and renewals expenditure used in the determination process. Review the extent to which internal data are available on a regional basis suitable for benchmarking, with particular emphasis on operating routes, Scotland, England, Wales and MerseyRail.</p> <p>Review the methodology by which data are collated and normalised, taking into account the sourcing of input data and the scope of data reworking that may be required.</p> <p>Assess linkage between external input data and Network Rail’s internal unit cost metrics; analyse data granularity, level of comparability, normalisation of comparator source data, and measures taken to ensure consistency of approach.</p> <p>Review data coverage levels and use of extrapolation, assess whether the outputs produced are sufficiently comprehensive and representative, and review the process by which individual outputs are linked to overall expenditure at the macro level.</p> <p>Provide a view on the quality and range of certainty of the benchmarking outputs.</p> <p>Review how unit cost benchmarking data can be meaningfully assessed alongside “real” maintenance and renewal items that NR will be using in its Asset Management Lifecycle Plans.</p> <p>Establish the fitness for purpose of the output data for determining available efficiencies in CP5 in each main asset expenditure category and by operating route, Scotland, England, Wales and MerseyRail.</p> <p>If appropriate, make recommendations for measures to improve input and output data quality, robustness and coverage, and identify and assess potential alternative sources of comparator data.</p> <p>Provide to ORR all data and documentation submitted by NR and comparator organisations in its original format wherever possible and explanations where not.</p>	<p>opinion) 0.6.2, 6</p> <p>6 (workstream data sections)</p> <p>Maintenance:6.7.3, 6.7.3, 6.7.5</p> <p>Other asset areas: Not feasible to address this requirement</p> <p>Maintenance: 6.7.3, 6.7.5</p> <p>Other asset areas: Not feasible to address this requirement</p> <p>0.6.3, 0.7.2, 5.5, 6.2.5 track projections), 6.6.3.3 (B&C projections)</p> <p>0.6.3, 0.7.2, 0.8, 5.5</p> <p>Not feasible to address this requirement</p> <p>0.7.2, 0.8.2, 5.5, 6 (workstream output sections)</p> <p>0.9 6 (Reporter opinion sections) 0.9, 6, 0</p> <p>Data to be transferred to ORR as required</p>

ORR mandate requirement	Reference report sections
	(subject to NR agreement).

Table 4: ORR mandate: relevant report sections

1.4 Report structure

We set out our findings in this report as follows:

- Chapter 2 Background and context: this explains the background and underlying purpose of NR’s bottom-up benchmarking programme, and reviews programme timescales within the context of the PR13 determination process.
- Chapter 3: we outline approaches taken from other infrastructure sectors, including where those benchmarking programmes have been used by regulators for in the context of price determination
- Chapter 4 Developing effective benchmarking groups: we discuss the key elements and principles for establishing and sustaining successful benchmarking groups, drawing primarily on Imperial College London’s experience of the CoMet and Nova benchmarking consortia.
- Chapter 5 NR bottom-up benchmarking –programme level review: this chapter sets out our findings with respect to the overall programme-level approach, objectives and structure as well as the nature of programme outputs and how these relate to the PR13 determination process.
- Chapter 6 NR bottom-up benchmarking – workstreams & data analysis: this chapter presents our detailed review of each of the ten asset-based benchmarking workstreams, the data and outputs that they produce, and how these are applied to NR’s CP5 expenditure and efficiency proposals.
- Chapter 0: we provide a summary of our recommendations.

1.5 Acknowledgements

We are grateful to Network Rail staff for making themselves available to assist us with our work and their continuing co-operation in providing us with material, arranging meetings and other assistance.

2 Background and context

2.1 Bottom-up benchmarking definition and objectives

In mid-2010 Network Rail began a bottom-up benchmarking programme. This involves the assessment and comparison of its own operations, maintenance and renewals costs, activities and processes with other organisations on a bottom-up basis, i.e. looking in detail at individual cost drivers / elements.

Bottom-up benchmarking is being undertaken by Network Rail to “understand differences in specific activities and their drivers, including quantifying impacts of structural factors and identifying improvement initiatives.” This is part of a wider exercise to utilise benchmarking as a means both to gain “information to drive performance and delivery of CP4 outputs” (i.e. outputs within the present regulatory control period), and to develop “informed and evidenced plans” for CP5. The ORR considers bottom-up benchmarking to be an important means by which Network Rail should “demonstrate to its stakeholders that its plans have fully considered available efficiencies”¹³. The ORR’s approach to assessing efficiency within the context of regulatory determination includes a combination of both top-down and bottom-up cost / efficiency analyses. The ORR regards detailed and specific outputs from bottom-up benchmarking to be of particular importance, given the narrowing of the estimated “efficiency gap” between Network Rail’s expenditure levels and target efficiency “frontier”.

The outputs from the bottom-up benchmarking activities are therefore of significant importance to both NR and the ORR. The focus of this mandate is to review Network Rail’s bottom-up benchmarking within the specific context of the PR13 regulatory determination process, through which the target efficient expenditure level for the next regulatory control period (CP5) will be determined.

2.2 CP5 determination process overview

The CP5 determination process involves submission of a number of key documents by NR, the ORR and the UK Department for Transport (DfT). It is proposed that NR’s bottom-up benchmarking will represent one of a number of key inputs into this process.

We set out in Figure 1 below an overview key milestones of the determination process leading up to CP5 and the timescales of NR’s bottom-up benchmarking programme in relation to it, as well as the timing of Arup’s independent review.

¹² “Overall Benchmarking Approach” presentation, “Introduction of NR benchmarking Prog.ppt”, 21st January 2011.

¹³ As set out in the assignment mandate (included in **Error! Reference source not found.** to this document).

¹⁴ See for instance the reference to combining bottom-up and top-down comparative analyses in the ORR’s PR08 document (October 2008), Section 7.21 (p.114)

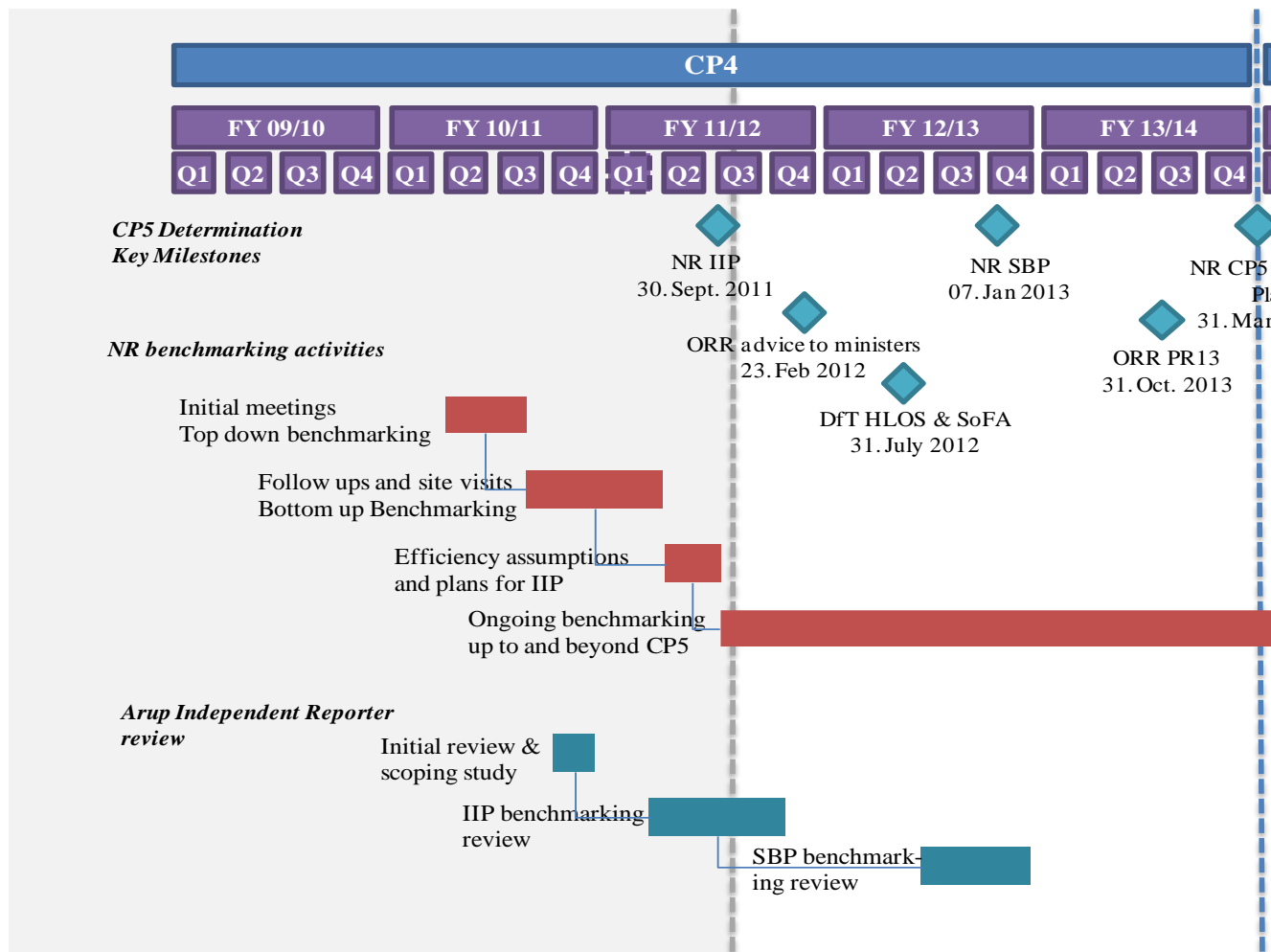


Figure 1: Overview of CP5 determination process and NR benchmarking programme

2.2.1 CP5 determination key milestones

The first key milestone in this process has been the submission of the Initial Industry Plan (IIP), published on 30th September 2011. The IIP is a document produced jointly by NR, UK passenger and freight train operators and industry suppliers, and includes NR's initial CP5 expenditure projections and its proposals for measures to achieve required efficiency targets. The way in which NR's bottom-up benchmarking data and analysis have been factored into the IIP proposals is a principal area of focus for this report.

Going forward, the CP5 determination process includes the following key stages:

- ORR advice to ministers (Feb. 2012): following a period of public consultation in response to the IIP (currently ongoing, due to be completed in late 2011), the ORR will in publish “advice to ministers and decisions on the framework for setting outputs and access charges” and will also “consult on detailed financial issues.”
- Publication by DfT of “high level output specification” (HLOS) and “statement of funds available” (SoFA) (July 2012): the purpose of these documents is to “ensure the railway industry has clear and timely information about the strategic outputs that Governments want the railway to deliver for the public funds they are prepared to make available.”
- NR publication of Strategic Business Plan (SBP) (Jan. 2013):
 - the SBP will represent NR's principal submission to ORR's 2013 periodic review;
 - building on the high-level expenditure and efficiency proposals contained in the IIP, the SBP will be based on more detailed expenditure and efficiency inputs, with more comprehensive and wider supporting analysis and evidence.
- PR13 (Periodic Review: Oct. 2013): this represents the finalisation of the ORR's determination of Network Rail's outputs and access charges/funding for CP5 which will begin on 1 April 2014. This takes into account both NR's SBP submission and a range of other inputs (studies, analyses, consultations, etc.).
- Taking into account both the company's own original SBP proposals, and the outputs and provisions set out in PR13, NR is required to produce a finalised and detailed overview of proposed expenditure and outputs for the Control Period through its CP5 Delivery Plan.

The role of bottom-up benchmarking within the context of ongoing steps in the PR13 determination process is an area of focus both for this report, and for future stages of this assignment.

2.2.2 Network Rail benchmarking activities and Arup review

NR's bottom-up benchmarking programme entails the following, as reflected out in *Figure 1*:

- Following initial top-down benchmarking activities and analysis, the bottom-up programme began in Q3 FY10/11 (October 2010).
- NR proposed to combine outputs from both the top-down and bottom-up workstreams in order to inform the efficiency assumptions and plans feeding into the recently completed IIP.
- the bottom-up benchmarking programme will continue for the remainder of the CP5 determination process and beyond.

As illustrated in *Figure 1*, Arup's review is based around three key stages:

- Our present analysis builds upon a brief initial review undertaken during Q4 FY10/11 (completed February 2011), which focused on NR's initial bottom-up benchmarking activities and which set out the proposed scope for the current IIP benchmarking review.
- Our present review, the results of which are presented in this report, focuses on NR's bottom-up benchmarking data that have informed expenditure projections and efficiency proposals within the IIP.
- It is anticipated that Arup will undertake review of data and outputs emerging from the bottom-up benchmarking on a quarterly basis, as these are expected to inform NR's development of its SBP. The SBP is expected to include a greater level of input and analysis from NR's bottom-up benchmarking programme, compared to the IIP.

3 Bottom-up benchmarking approaches

3.1 Introduction

Chapter overview

This section of the report outlines approaches taken to bottom-up benchmarking from other infrastructure sectors, including where those benchmarking has been used by regulators in the context of price determination.

We present an overview the utilisation of bottom-up benchmarking in the following (non-heavy rail) industry sectors:

- Water supply (England & Wales): bottom-up benchmarking is an established core element of the regulator's price setting activities.
- UK electricity distribution: bottom-up benchmarking has been utilised by the regulator to supplement top-down econometric analyses to assess comparative efficiency.
- UK gas distribution: bottom-up benchmarking has been utilised to a limited extent to support the regulator's price review process.
- UK electricity and gas transmission: although utilisation of benchmarking to date has been limited, the regulator has begun engaging with overseas counterparts to obtain comparative benchmark data over the longer-term.
- International metro benchmarking: this relates to an international benchmarking body in which comparator organisations participate for their own benefit / interest, but not for regulatory purposes.

We also set out in **Error! Reference source not found.** of this report the types of data collected and monitored through benchmarking activities. We cite as case studies the international metro benchmarking programme referred to above, together with other benchmarking programmes undertaken in the UK water and electricity transmission sectors.

Utilisation of benchmarking for regulatory purposes

Regulators face well established principal-agent problems when setting price limits for regulated firms. One significant problem is an information asymmetry: regulated firms will always better understand the potential for cost reduction and service improvement. This asymmetry can lead to a moral hazard issue, whereby firms have an incentive to understate scope for cost reduction and service improvement.

Benchmarking can help the regulator reduce information imperfections by generating information that is independent (or more independent) of the regulated company.

Regulators can benchmark many measures, such as prices, costs, service levels, environmental performance, financial performance, technology and processes. In view of the focus and approach of NR's present benchmarking programme, we have considered primarily cost benchmarking and process benchmarking. We contrast 'bottom-up' methods with whole-of-market, top-down econometric

methodologies, paying more limited attention to the latter for each sector reviewed.

3.2 Water sector (England & Wales) benchmarking

3.2.1 Overview

The water sector in England and Wales was privatised in 1989 with considerable scope for cost efficiency improvement. The Water Services Regulation Authority (Ofwat) identified ‘comparative competition’ between the then 39 water companies as a mechanism to deliver cost reduction, creating data collection systems and surveys of cost reduction techniques to monitor improvement.

3.2.2 Utilisation of econometric analysis

Ofwat has developed econometric analyses to assess companies’ relative operating costs and capital expenditure. It has conducted separate econometric analyses of water and sewage services annually since 1994. At the first price review (PR94) Ofwat assessed relative operating cost efficiency separately for water and sewage services by:

- carrying out single year cross-sectional analyses via corrected ordinary least squares (COLS) regressions at a range of levels to determine relative efficiency;
- assigning companies to relative efficiency bands, depending on their distance from the most efficient; and
- making “catch-up” assumptions accordingly, eventually assigning companies to bands ‘A’ (<5% variance) through ‘E’ (10% per band from bottom of ‘A’).

Ofwat began reporting comparative efficiency in 1996/97, and in 2009 set out the requirement for firms to close 60% of the difference between the middle of their band and the benchmark within five years of reporting. In 2009 (PR09), Ofwat extended its econometric approach to capital expenditure, and developed four additional econometric assessment techniques (e.g. distribution as a function of length of water main per number of properties).

Ofwat’s use of cross-sectional OLS regressions to assess efficiency has been questioned from an academic perspective. It has been argued that the omission of explanatory factors and reliance on a small, single-year cross section of data greatly limits the statistical significance of the regressions and, therefore, their regulatory value. Ofwat is considering changing its statistical models and using panel (multi-year) data to improve their predictive power.

3.2.3 Utilisation of benchmarking

Ofwat’s principal tool for assessing capital expenditure improvements is the “Cost Base” tool. This is utilised every five years, for price reviews only, to assess capital expenditure and process benchmarking. As before, the analysis relies on a single year’s data and treats sewage and water facilities separately.

Capital expenditure

The capital expenditure cost base analysis involves the following three key stages:

- Firstly, companies submit forecasts of unit costs for the next five years. Ofwat audits and challenges the forecasts, checking for consistency and compliance with its reporting guidelines.
- Second, should disagreements arise, visits to the relevant company by engineering consultants, who probe material differences, then take place. Companies submit final cost base submissions one year later.
- Finally, Ofwat assesses relative efficiency of capital expenditure through:
 - normalisation of costs to allow for regional variation;
 - comparisons of the percentage differences between a company's cost and the median cost for each standardised project;
 - weighting these percentage differences according to the composition of the capital expenditure forecast in each company's regulatory business plan; and
 - arriving at an overall efficiency assessment for each company, under the four categories water infrastructure, water non-infrastructure, sewerage infrastructure and sewerage non-infrastructure.

Ofwat uses the assessment to increase (or decrease) a company's baseline capital expenditure permitted under the Capex Incentive Scheme.

Process benchmarking

To help set these capital expenditure allowances, Ofwat has carried out a form of process benchmarking for the water sector, through utilisation of an Asset Management Assessment (AMA). Rather than assessing individual investment proposals, AMA examines the process governing overall asset planning and investment activity.

The AMA process is inherently subjective, relying on Ofwat's scoring and judgment:

- Ofwat reads companies' business plans and reporters' submissions and judges the quality of the asset planning process as revealed by the plan.
- In PR09, Ofwat scored each company's plan from 0 (well below expectation) to 5 (well above expectation).
- Assessment of each plan occurred for 28 different dimensions, grouped under nine headings. Criteria included "stakeholder engagement", "leadership", "policy and strategy", "management", "processes", "systems", "data", "analysis" and "reporting".

Companies that scored highly then received a smaller reduction to their capital maintenance proposals than those that scored poorly.

3.2.4 Company-led benchmarking activities

Such extensive regulatory analysis has arguably tended to overshadow or even crowd out benchmarking initiatives by individual water companies. Analysis by the regulator also has focused companies' attention on the wants and needs of the regulator - Ofwat's efficiency targets - rather than consumer cost and performance which are arguably of equal or greater importance.

Nevertheless, some companies in the sector undertake benchmark internally, comparing performance between divisions or units and many benchmark informally by visiting each other to learn about best practice. Water UK (the industry's trade association) facilitates the ad-hoc collection of comparative data on specific issues.

3.2.5 Challenges and possible developments for benchmarking in the water sector

The water sector experience highlights a number of challenges:

- Volume of data required: acquiring sufficient and robust comparative data are essential to good benchmarking but can be challenging to achieve.
- Administrative burden: Ofwat completes annual assessments using data collected in the June Return. Completing the June Return is a substantial undertaking for each company, requiring thousands of items of data and taking over a year.
- Benchmarking limitations / reliance on judgment: process benchmarking through the AMA relies almost completely on judgments made by Ofwat staff and could be seen as more a test of companies' ability to explain their asset management and investment processes than the processes themselves.

Ofwat is reviewing its approach to setting price limits, and more details about proposed future changes are expected by the end of 2011. Some possible future developments include:

- Sub-company analysis: as part of its accounting separation project, Ofwat requires water companies to provide information for nine "business segments." The benchmarking of specific business segments may help improve cost transparency in contestable parts of the value chain, support determination of access prices and help set regulatory backstops.
- Bottom up analysis: Ofwat is also exploring the possibility of developing further bottom-up benchmarking activities, which could include building up company-specific engineering models and comparing costs derived with actual costs. This approach is complex and data intensive and may be used on a case-by-case basis to start with.

3.3 Electricity distribution (UK) benchmarking

3.3.1 Overview

Comparative analysis has a longer history in electricity distribution than in other utilities sectors, because 12 regional electricity companies have been operational within the sector from the outset (of privatised sector operation). Even so, in the early 1990s the benchmarking focus remained macro-economic, and the first micro-level attempt at comparative benchmarking came in 1999. This was aimed at determining a basis for individual firms to catch-up with a market-average efficiency frontier.

3.3.2 Utilisation of econometric analysis

Market regulator Ofgem has analysed operating expenditure through both a top-down, econometric approach and bottom-up consultant studies. Notably, Ofgem has used the two approaches to cross-check each other, finding broadly consistent results.

Ofgem's modelling of operating expenditure in the electricity market demonstrates use of increasingly complex statistical analysis:

- Analysis in DPCR3 (1999) consisted of a top-down OLS regression of controllable operating costs for 14 networks against a composite scale variable (CSV), accounting for customer numbers, units distributed and network length. Data were cross-sectional (single year – 1997/98).
- DPCR5 went further by using multi-year panel data to model operating costs at three different levels: disaggregated, single-group, and single, top-down OLS.

Ofgem concluded that the disaggregated analysis approaches were most cost effective and so focused on these, resulting in more than 40 separate efficiency scores for each network. For network operating costs, the frontier was defined as the upper third percentile and for indirect costs the upper quartile.

3.3.3 Utilisation of benchmarking

Operating expenditure

OfGEM has commissioned bottom-up consultant studies of operating expenditure to supplement top-down econometric analysis. The aim has been to ascertain the efficient base-year level of operating expenditure theoretically achievable, given the application of “best practices” within the sector.

- Best practice and cost metrics have been defined as the activities and costs of the best-performing companies (based on OLS analysis).
- Activities include the main electricity distribution activities (i.e. engineering, meter operations, corporate and administrative). Costs include the operating costs associated with these activities.

- The analyses have provided an estimated range of the efficiency savings achievable for each component activity, and these were combined to give an overall level of operating cost savings possible for each company.

Ofgem also commissioned Ernst & Young to assess qualitatively network companies' overheads and corporate costs. The study produced results broadly consistent with the top down benchmarking.

Capital expenditure

To assess capital expenditure, Ofgem commissioned consultants PB Power to review the companies' proposals and develop models for load- and non-load related expenditure:

- PB modelled load-related expenditure by benchmarking each company's forecast and historic spend as a percentage of MEAV per customer, drawing upon 15 years' worth of data from 1995 to 2010. The median ratio across the industry was used to determine an allowance for load-related expenditure.
- PB also modelled non-load expenditure by benchmarking companies' asset replacement profiles, to arrive at a standard profile for each category. PB calculated allowances by applying the standard to each network's asset population at March 2003 and multiplying by standard unit costs.

More recently, Ofgem benchmarked capital expenditure in relation to asset characteristics and unit costs. For example, for asset replacement it benchmarked each Distribution Network Operator's (DNO's) forecast against the forecasts of other DNOs, taking into account the age profile of assets on the individual networks and a standardised asset replacement model developed and shared with the industry. A key difference from DPCR4 was the reduced reliance on external consultants, with more of this work brought in-house and delivered by Ofgem directly.

3.4 Gas distribution (UK) benchmarking

3.4.1 Overview

The utilisation of benchmarking within the UK gas distribution sector has been extensive than in other UK utilities sectors. The gas industry was privatised in 1986 as a single entity with a single price control. A separate transmission and distribution price control was set in the early 1990s, but it was not until 2001 that transmission and distribution price controls were separated, and not until the sale of gas distribution networks in 2004 that comparative analysis between gas distribution companies was possible.

3.4.2 Utilisation of benchmarking

GDPCR1 in 2007 was the first gas review that involved significant benchmarking between companies. Because of the lack of data for top-down analysis, the regulator (Ofgem) opted to benchmark the efficiency of individual activities comparing both direct and indirect cost areas.

Ofgem has aggregated bottom-up measures to complete basic econometric analysis. It has assessed direct operating cost efficiency individually, including measures of work management, emergency services, repair and maintenance, using an OLS regression corrected to the upper quartile level of efficiency.

Ofgem commissioned external consultants to assess capital and replacement expenditure requirements by benchmarking across gas distribution networks. Econometrically, this involved running OLS regressions on single-year, cross-sectional data (2006/07) for connections and mains reinforcement capital expenditure, corrected to the upper quartile. Additionally, consultants conducted project-specific reviews to assess the need for investment and the “appropriate” unit costs.

3.5 Electricity & Gas Transmission (UK) benchmarking

3.5.1 Overview of costs benchmarking

Because of the lack of comparators, electricity and gas transmission price controls have traditionally been set without reliance on benchmarking. Instead, the approach has been to employ technical consultants to challenge the strength of the needs case made by the companies and to “chip away” at their capital and operating expenditure projections.

3.5.2 Company-led benchmarking activities

Although useful, it is hard to see companies’ benchmarking activities as a replacement for benchmarking carried out by the regulator; such activities have tended to focus on operational issues rather than overall efficiency. National Grid’s benchmarking focuses on operating costs.

National Grid has benchmarked its electricity transmission through the International Transmission Operations and Maintenance Study (comparing direct maintenance operating costs and facilitating the exchange of best practice) as well as the International Comparison of Transmission System Operators (involving annual surveys to inform on best practice in systems operation).

For gas transmission, National Grid is involved in European benchmarking through the Gas Transmission Benchmarking Initiative.

3.5.3 Challenges and possible developments

The electricity and gas transmission markets highlight several issues for consideration, including:

- The appropriate level of disaggregation and detail. Current thinking in Ofgem is that DPCR5 was too detailed and that at future reviews it should be sufficient to rely on about eight core regressions.
- The need for aggregated, top-down assessment. Disaggregated bottom-up analysis does not allow for substitutability between cost categories which can be influenced by different decisions companies make about how to combine inputs to deliver outputs. This may lead to inconsistency.

- The lack of comparators for transmission. To date Ofgem has not been able to carry out significant benchmarking for electricity and gas transmission because of an absence of sufficient UK comparators.

Gas and electricity transmission (RIIO-T1) and gas distribution (RIIO-GD1) price control reviews are underway, and point to a number of benchmarking developments, including:

- Greater use of benchmarking for transmission. Ofgem is working with its European counterparts to build up a data set for comparative analysis. This is seen as a long-term project.
- Benchmarking is being considered as a means by which specific efficiency- and performance-related questions can be posed to companies.
- Benchmarking of future costs: Ofgem intends to benchmark companies' total expenditure forecasts in relation to outputs, to give comparative indications of value for money.
- Greater emphasis on the companies' own benchmarking: Ofgem is keen to shift the burden of proof to companies, so that firms themselves will have to justify their business plans.

3.6 International metro benchmarking

3.6.1 Overview of costs benchmarking

CoMET (the Community of Metros), and its sister group, named Nova, are the world's largest urban railway benchmarking groups, which have been active for more than 16 years. CoMET began in 1994 when five heavy metros came together to form the "Group of Five", which expanded and took on the name CoMET in 1996. Nova was established in 1998 to cater for medium-sized metros, but the two groups are part of the same wider benchmarking consortia led by the Railway and Transport Strategy Centre at Imperial College London. Combined, the two groups currently include urban railways in 27 world cities that collectively share information and best practices. CoMET and Nova's principle aim is to identify and share best practices in a confidential environment.

The objectives of the group remain largely unchanged since its inception in 1994. These are:

- To share knowledge and identify best practices in a confidential environment.
- To build systems of measures for use by management and to establish metro best practice.
- To provide comparative information both for the metro board and the government.
- To prioritise areas for improvement.

The success of the groups led Imperial to establish an international Bus Benchmarking Group for operators of large city networks (usually over 1,000 buses) in 1994. It was not known at the time whether the benchmarking process

that had been developed for metros was transferable to other public transport sectors, but this group was equally successful and is still in operation.

In 2009, Imperial established a similar group for Suburban Rail operators, further indicating that the group structure and framework was successful in its generic form.

3.6.2 Company-led benchmarking activities

The CoMET and Nova groups are owned and run by the members themselves, and managed by the Railway and Transport Strategy Centre (RTSC) at Imperial College. The members gather twice a year around the world and maintain continual contact through the group's secure website and online forum as well as through project management, benchmarking analysis, and coordination provided by the RTSC.

It is important to note that benchmarking is not merely a comparison of data points or the creation of league tables. Rather, benchmarking is a comprehensive effort that stimulates productive "why" questions to identify lines of inquiry and best practices, with a clear focus on implementable results.

CoMET and Nova's benchmarking process follows an annual cycle, including a range of activities. At the highest level, these include:

- One annual meeting and one steering group meeting each year for each group, hosted by member organizations on a rotating basis.
- A standardised Key Performance Indicator (KPI) system to assist in the identification of strong performance and best practices.
- Benchmarking case studies consisting of in-depth research and analysis on specific issues of mutual interest to members.
- A 'clearinghouse' study mechanism for the member-initiated collection of specific data and information in areas of immediate interest.
- An online forum and secure website to facilitate the rapid sharing of information amongst members.
- Expert groups and workshops.

3.7 Lessons learned from other sectors

The activities observed in other sectors provide important context for the development of NR's bottom up benchmarking programme, and consideration of how its outputs might be used.

Key points we note include:

- The use of bottom-up benchmarking outputs by regulators in the water sector (Ofwat) and for gas and electricity distribution networks (Ofgem). In those sectors, benchmarking is undertaken with reference to a range of UK comparators operating under a single regulatory framework. Participation in benchmarking activities in those sectors is supported by the regulatory framework within which those businesses operate.

- We understand that Ofgem does not draw significantly on benchmark data in its regulation of UK gas and electricity transmission networks. In this respect the fact that those assets are operated by a single national entity is relevant. This is also a relevant consideration in the context of NR's benchmarking activities. However, UK transmission companies do participate in company-led international benchmarking activities, supporting attempts to identify good practice and improve efficiency.
- Imperial College's metro benchmarking groups offer further examples of company-led international benchmarking programmes, within the broader rail / passenger transport sector.

We set out in **Error! Reference source not found.** of this report details of the types of data collected and monitored through benchmarking activities.

The next chapter of this report draws on the experiences of benchmarking documented in this chapter - and particularly experience of Metro benchmarking – to identify a number of key lessons on which NR may be able to draw in shaping the long term development of its bottom-up benchmarking activities.

4 Developing effective benchmarking groups

4.1 Introduction

We discuss in this chapter of the report the key elements and principles for establishing and sustaining successful benchmarking groups, drawing primarily on Imperial College London's experience of the CoMet and Nova benchmarking consortia. We also set out what we consider to be the inherent challenges and potential limitations of bottom-up benchmarking activity, viewed in the context of PR13 timescales.

If managed and set up properly, benchmarking groups can be an effective means of facilitating a free-exchange of benchmarking data and information, usually for the sole purposes of identifying best practices for performance improvement.

A benchmarking programme will be successful if it has a clear purpose, objectives, and follows some basic principles that enable a free-flow of information and data, ensuring benefits to participants.

Experience suggests that a successful benchmarking process need not be overly theoretical, should provide insights which add value, identifies best practice, focuses on transferable and implementable results, facilitates regular information flow between participants, encourages members to think as an industry but requires a supportive culture and systems from the participants.

Internationally, the greatest barrier to a formal process for benchmarking of infrastructure managers such as the European national rail operators, is that they are typically large, monolithic organisations, highly politicised and fearful of the consequences of information falling into the wrong hands.

Experience indicates that confidential benchmarking consortia can provide a successful route to lowering this barrier. However, nervousness in sharing results with regulatory / governing bodies will always persist and regulatory bodies may need to step aside to allow the benefits of best practice sharing to be realised, even if this comes at the expense of providing regulators with detailed information. As discussed in Section 4.4, such challenges must be considered when assessing NR's bottom-up benchmarking programme – particularly considering the limited timescales of the CP5 determination process, into which the programme outputs are expected to be factored.

4.2 Success factors for benchmarking programmes

We detail in this section the twelve key factors that have supported the successful long term operation of the CoMet and Nova benchmarking programmes:

- Confidentiality and trust.
- Collaboration.
- Speed of information exchange.
- Independence.
- Properly resourced central co-ordination.
- Long-term approach / continuous annual cycles.
- Resources and senior level support.

- Selection of member peers.
- Hierarchy of benchmarking data systems.
- Hierarchy of information exchange.
- Face-to-face communication.
- Broad scope supported by working groups.

4.2.1 Confidentiality and trust

Confidentiality and trust are essential for benchmarking processes to work effectively. The success of the programme relies on uninhibited free-sharing of data and information between benchmarking partners. Data exchange is likely to be supported by formal confidentiality agreements but also relies on the long-term building of trust between partners, and a clear mandate from top-level management that information sharing is permitted and encouraged. Consortia-oriented, centralised approaches can overcome such barriers.

NR has told us that it has often found European rail infrastructure managers (IMs) reluctant to share detailed data about costs and resources. The UIC LICB process was designed to achieve this but we understand that it was not intended to allow for additional data exchange ‘on-demand’. It appears that the level of trust amongst participants is limited (as demonstrated for example by the withdrawal of DB Netz¹⁵ from the process).

4.2.2 Collaboration

Benchmarking activities are defined by all participants and not led by any one member. There must be processes that permit any participant to access particular information that they need at a particular time. A key principle is that members help each other to improve their respective businesses. The benchmarking process that supports this requires unimpeded information exchange.

4.2.3 Speed of information exchange

Speed of information exchange is important to maximise value. The benchmarking process should have both longer term initiatives (such as major studies on a particular topic) but should also support additional quick information exchange on an as-needed basis. Participants should not be frustrated by bureaucratic processes such as the signing off questionnaires or detailed study scopes at steering group meetings.

4.2.4 Independence

The process should be owned solely by the member IMs and not by authorities, governments or regulators. This independence permits the group to pursue its own agenda and focus its efforts in areas of most immediate interest to members. In a railway context, there is the inevitable temptation by governing authorities and regulators to want to access databases. If this is to happen, it must be handled with care to support the continued exchange of information, with full anonymisation of the results as they relate to other participants.

¹⁵ German national rail infrastructure operator.

4.2.5 Properly Resourced central coordination

Although full independence is ideal, a centralised, independent coordinator and analyst can add credence to the benchmarking, provide dedicated resources and ensure objectivity. The central body can facilitate and administer the benchmarking process, undertake analysis and provide benchmarking expertise.

4.2.6 Long-term approach / continuous annual cycles

One-off benchmarking studies rarely deliver long term value; more value can be secured through a long-term approach, with continuous efforts operating over an annual cycle. In particular, setting up a system of metrics with clear definitions for the provision of both high level and detailed data, takes time.

Imperial's experience in developing and managing metro and other benchmarking programmes is that it may take around four years of development to build a high degree of confidence in the comparability of benchmarking outputs.

4.2.7 Resources and Senior Level Support

In order to maximise the value gained from the benchmarking process, each organisation must commit its own (not insignificant) resources to managing collaboration, data collection, meeting attendance and internal communication of benchmarking results. Naturally, experience has shown that the organisations who gain the most benefit from benchmarking are also those who put the most effort into the process, and those who have truly integrated the initiative as a business improvement tool into their organisations.

The benchmarking process for individual participants should be owned by a senior representative with sufficient influence to spread the benefits of benchmarking within the company. In a railway context this is often the Operations, Engineering director or another senior director. This representative should attend all major group meetings and be able to commit enough time to ensure that the company gains maximum benefit from participation in the group.

In addition, each organisation should appoint a data coordinator and facilitator to carry out most of the work in gathering data and information from the different managers and sources within the organisation. Functional managers should be available for occasional interviews and the completion of benchmarking data and case study questionnaires as required.

4.2.8 Selection of Member Peers

There must be a critical mass of participants. A rule of thumb used by Imperial is that a minimum of eight active participants is required for an effective group. The process of generating best practice ideas will be more effective if the peers are truly international and do not include only European entities. Diverse cultures lend themselves to different and innovative approaches. It is important to ensure that some of the participants are seen to be 'good practice / efficient'.

4.2.9 Hierarchy of benchmarking data systems

Imperial's experience in developing and managing metro and other benchmarking programmes is that it may take around four years of development to build a high degree of confidence in the comparability of benchmarking outputs. Across the initial (annual) benchmarking cycles, KPI systems can be developed, data disaggregated, econometric analysis undertaken where it can add value and case studies launched to 'drill down' further to explore important issues or significant areas of apparent difference.

4.2.10 Hierarchy of Information Exchange

Although benchmarking outputs will inevitably generate significant focus on KPIs, other types of information exchange can also add significant insight. In particular, Imperial's experience is that detailed case studies on asset-specific areas are important, as are other smaller ad-hoc studies. Participants in the Metro benchmarking consortia also value fast information exchange using web / forum-based tools.

4.2.11 Face-to-face communication

In addition to remote participation in benchmarking activities, Imperial's experience is also that face-to-face communication plays an important role in the continuing success of the benchmarking consortia it manages. In practice, it is necessary for members to meet on a regular basis, to steer the process and to hear and discuss results of the benchmarking analysis.

4.2.12 Broad scope supported by working groups

A successful benchmarking process will be resource intensive and maximising value for all participants is a key aim. Consortia should not be focused only on a single issue (e.g. costs), but on all elements of the business (including safety and system performance). For very large organisations, there should be not only a high-level central co-ordinating committee, attended by director-level representatives, but also sub-groups that can undertake benchmarking activities at a function or asset-specific level, although these sub-groups must be centrally coordinated.

4.3 Key stages and timescales for establishment of a successful benchmarking programme

Drawing on the experiences referred to above Table 5 overleaf sets out an outline timetable for establishing and securing value from benchmarking of the type considered by NR, pursuing a consortium approach.

Indicative timeframe	Activities
0-6 Months	Establish support from an industry or central coordinating body, and approximately 3 initiating railways, including one non-EU. Start-up funding from 3 initiating railways.
6-12 Months	Develop details of proposed group framework using start-up funding. First draft of initial data systems. Synthesis of previous benchmarking.
Year 2	Start-up phase focused on attracting 10 initial members Some initial benchmarking to gain immediate value.
Year 3	First full annual phase of the group. First draft of initial data systems. First detailed drill/down bottom/up case studies. Attract further members.
Year 4	Second full annual phase. Develop sub-groups from last year's case studies as continuous groups. 3 detailed case studies in functional areas. Second cycle of refinement of data systems. Attract further members.
Year 5	Continuous process of development.

Table 5: Key stages and timescales for benchmarking programme

The full value of the benchmarking programme may not be realised for around four years. It is important to note that outputs secured before then can still be expected to deliver significant value.

When considering how bottom-up benchmarking might deliver value to NR and ORR in the context of NR's next price determination, we note that NR will be two to three years into the relevant programmes of work at the time that the most substantive discussions around PR13 are likely to take place.

4.4 Inherent challenges and limitations

We consider there to be challenges and limitations inherent within bottom-up benchmarking activities, which are particularly relevant when viewed within the context of the PR13 determination process and timescales.

4.4.1 Short-term challenges and limitations

As discussed in the Section 4.3, experience suggests that successful benchmarking requires both mutual interest and commitment of participants, and time.

Collating and interpreting that data are unlikely to be a simple and straightforward process, particular when these data are sourced from organisations operating in different countries under different organisational structures and regulatory

regimes. A significant level of scrutiny of supporting information and detailed explanations are needed in order to understand influencing factors driving bottom-up cost factors, and to gain meaningful results.

A related challenge is gaining the cooperation and trust of comparator organisations. As discussed in the previous sections, the process of establishing contact, gauging interest and identifying mutual goals and interests requires a significant level of discussion and personal interaction. Parties need to discuss the objectives of such activities, and clarify the nature of interaction and are likely to need to allay concerns relating to sensitive benchmarking information. Imperial College London's experience showed typically around two years of dialogue were needed before organisations become engaged in the process of exchanging data (though this may vary (see Section 4.3)).

Taking the above two factors into account, it is likely that information shared in the early stages of benchmarking engagement is unlikely to be of significant depth and detail. Time will be needed both to establish common interest and trust with comparator organisations, and to obtain and analyse further information and insight to deepen the analysis.

4.4.2 Medium- to long term challenges and limitations

Assuming NR continues the benchmarking programme in the mid- to long-term (i.e. beyond PR13), we consider the potential for a greater volume of meaningful benchmarking data to be obtained will increase.

On the one hand, NR may widen the scope of numerical observations and data it obtains. However the underlying problem of conflicting objectives is likely to remain, whereby NR focuses on discussing and identifying potential efficiency factors, and continues engaging with comparator organisations on the same basis. As a result, problems relating to limited amounts of detailed comparative cost data of sufficient breadth across a sufficient range of comparators are likely to persist.

On the other hand, NR may choose to establish a more formalised benchmarking group which, as Imperial College London's experience suggests, may facilitate sharing of more direct cost data to a detailed and granular level. However, this may also be problematic for the ORR, as such activity would probably need to be based on strict rules and procedures amongst participants to ensure confidentiality; this would preclude disclosure by NR of the data obtained to any other party.

Overall, the degree to which the data and outputs will be useful for the ORR's purposes of gauging and regulating NR's relative efficiency remains uncertain.

4.4.3 Potential bottom-up benchmarking with European comparators by ORR

Noting that NR's programme of bottom-up benchmarking is unlikely to have made the degree of progress desired by the ORR in time for PR13, we recognise that there may be a temptation for the regulator to pursue a bottom-up benchmarking exercise on its own.

We consider that any attempt by the ORR itself to initiate benchmarking activity with European rail transport organisations is unlikely to be feasible for the following reasons:

- First, as discussed above (and as demonstrated by NR), any bottom-up benchmarking engagement, requires more time than is now available to ORR in the run up to PR13,
- Second, we consider there to be inherent differences in objectives between the ORR and European rail infrastructure operators; whilst ORR is seeking to gauge relative efficiency levels, European rail operators are likely to be engaging in benchmarking for purposes of internal business improvement, innovations, etc. We consider it unlikely that such organisations would wish such data to be used to support regulatory activity.¹⁶

Overall, this suggests that direct engagement by ORR with European operators to gain benchmarking data are unlikely to be successful.

4.4.4 Conclusion

We have set out in this section of the report what we regard as the inherent challenges and limitations that typically apply to bottom-up benchmarking activities.

We consider that, from the ORR's perspective, NR's engagement with comparators is unlikely to progress to the level that facilitates sharing of robust and comprehensive bottom-up benchmarking data to a level that can meaningfully inform the ORR's PR13 analysis of the efficiency gap. Experience suggests significant time is needed to establish the benchmarking relationships that are likely to yield more significant meaningful benchmarking results. The PR13 timescales can be considered comparatively short-term in benchmarking terms, and indeed the level of progress to date, as documented in this report, is limited.

Nevertheless, we recognise NR has adopted a range of approaches to gain understanding of relative cost and efficiency factors within the time available, which we review in detail later in this report. We make recommendations throughout this report, to support both the ORR and Network Rail in maximising the benefits that can be gained from the programme during this period.

¹⁶ We understand that utilisation of LICB data by government / regulatory authorities has led to dissatisfaction of European rail participants, and subsequent departure of some members from the group. This appears to illustrate the point that European rail organisations are unlikely to wish benchmarking data to be utilised for regulatory purposes.

5 NR bottom-up benchmarking –programme level review

5.1 Introduction

We present in this chapter the programme-level findings of our review of NR’s bottom-up benchmarking activities. We set out our findings under the following sub-sections:

- Programme objectives and strategy
- Programme approach
- Selection of comparator organisations
- Programme data
- Programme outputs

At the end of each sub-section we provide an initial Reporter opinion in relation to the given area.

5.2 Programme objectives & strategy

5.2.1 High-level purpose and context of NR benchmarking activities

Network Rail considers the bottom-up benchmarking activities to represent an integral part of its overall programme of benchmarking activities, the purposes of which are to “support the delivery of our CP4 outputs” and “inform and provide evidence for the CP5 planning process.”¹⁷

NR is proposing to combine and reconcile bottom-up and “top-down” benchmarking analysis to achieve the overall purpose of “identifying internal and external best practice and improving business performance.” We reproduce below a chart provided by NR, setting out the alignment and combination of top-down and bottom-up activities.

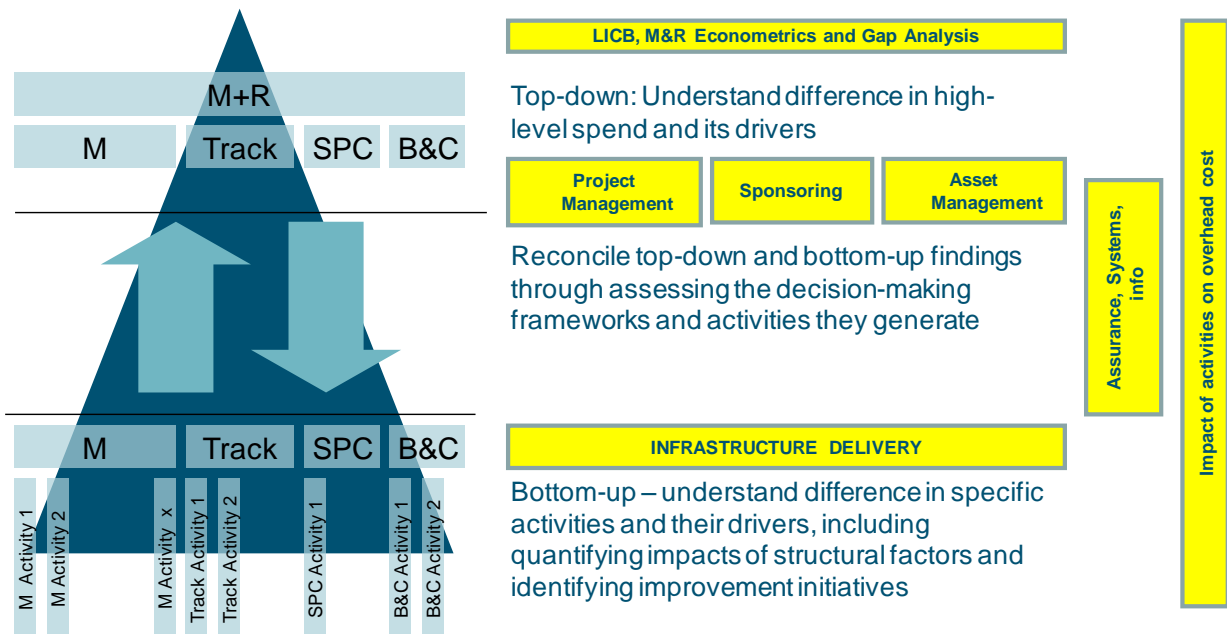


Figure 2: NR outline of top-down and bottom-up benchmarking alignment

At the centre of the ORR regulatory determination process is the assessment of the “efficiency gap” between NR’s expenditure levels and those of the most efficient / “frontier” rail infrastructure organisations. NR is proposing to utilise bottom-up benchmarking to help identify specific measures through which the gap can be closed. This is illustrated diagrammatically in Figure 3 below.

¹⁷ In January 2011 NR gave a presentation setting out its benchmarking activities, approach, aims and objectives. This included a number of statements and diagrams that we cite directly in this section of our report. A copy of the full set of presentation slides is included in **Error! Reference source not found.** of this document.

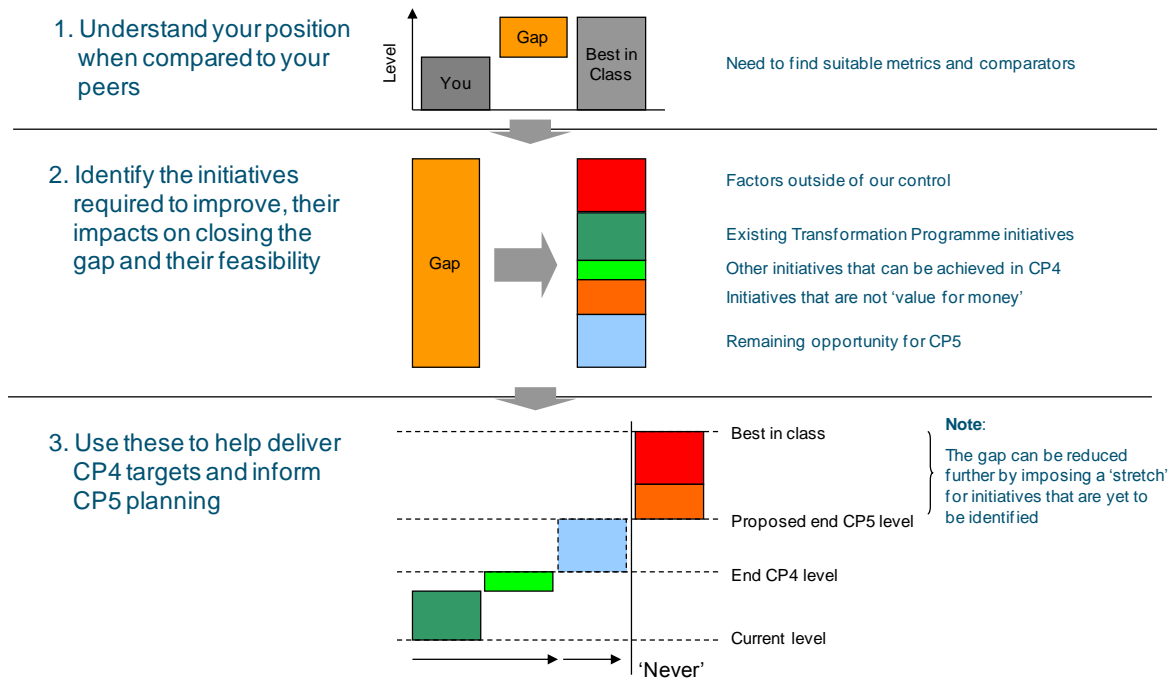


Figure 3: NR outline of benchmarking analysis of the efficiency gap

NR considers that whilst the top-down benchmarking provides an understanding of the overall efficiency levels relative to comparators, it is the identification of specific factors and initiatives which make up the efficiency gap, identified through bottom-up benchmarking which will then help deliver CP4 targets and inform CP5 planning, as illustrated in Figure 3.

5.2.2 Workstream-level objectives

The bottom-up benchmarking programme is being delivered through ten individual, asset-specific workstreams (see Section 5.3.1).

NR has indicated that the principal objective at the workstream level is to identify and develop efficiency opportunities relevant to that given asset area; each team is responsible for establishing its own approach and identifying those cost / efficiency aspects it sees as most relevant.

Although the workstream leads consider their findings to be an important input into the CP5 expenditure determination process, they do not regard the purpose of the exercise as being to provide bottom-up quantification of the efficiency gap. The focus is on improvements and efficiencies both in the short-term (within CP4) and the longer-term (CP5 and beyond).

5.2.3 Reporter opinion

Principal observations

NR and ORR regard the bottom-up benchmarking both as a programme of work which can inform the determination at PR13 and a means to support the long term objective to improve efficiency on the railway.

Our review has found that teams involved in the benchmarking workstreams are clearly well motivated and working with a clear focus on identifying genuine and meaningful comparisons to drive improvements and best practice.

However, although NR has highlighted establishing a greater understanding of the efficiency gap at the programme level, we have not found any specific, written objectives reflecting this objective at the workstream level.

NR is not pursuing this programme as a stand-alone exercise to provide a comprehensive picture of every inefficiency which might exist in its business. Clearly an effort to gain an absolutely comprehensive / all-encompassing overview of every bottom-up cost element contributing to the efficiency gap is unlikely to be workable or realistic. However, the orientation of workstream analysis predominantly towards opportunities means that certain factors influencing relative cost or efficiency could be seen as outside scope of potential efficiencies and risk being excluded from the analysis.

Recommendation

- We recommend that the requirement to obtain data and undertake analysis to support and substantiate the analysis of NR's efficiency gap is formalised and implemented by NR as an objective to be followed at the workstream level. We consider that the formalisation of this objective for each workstream should influence the approach taken, the scope of analysis and the utilisation of the data obtained by the given workstream. We also consider that this should improve visibility of efficiency factors identified by NR (including those initiatives already identified), and may provide additional insight in relation to how and why a given efficiency factor / opportunity is being prioritised, thereby helping justify its implementation.

5.3 Programme approach

5.3.1 Bottom-up benchmarking workstreams

Network Rail's bottom-up benchmarking programme is broken down into number of asset-specific workstreams each of which is led by the asset management / IP / maintenance departments responsible for delivery of the given asset group. NR has indicated that such departments are best placed to challenge cost drivers and identify efficiency opportunities, given their responsibility for day-to-day management and delivery of the infrastructure.

The benchmarking programme consists of a total of ten workstreams, which can be summed-up into four main groupings, as follows:

- Track: delivered by track Asset Management team (lead: Nick deBellaigue)
- Signalling, Power & Telecommunications: delivered by Asset Management teams (oversight: Alan Ross), with specific workstream leaders:
 - Signalling (Rob Ireland)
 - Telecommunications (Fraser Allan)
 - Power & Electrification (Adrian Murray)
- Civils & structures: delivered by both Asset Management and Investment Projects:
 - Initial benchmarking study (undertaken in 2010) comparing costs for three specific civils schemes, and subsequent workstreams comparing UK and European best practice, are being delivered by Investment Projects (lead: Andy Lundberg)
 - Benchmarking activities focusing on policy related aspects, specific to the three civils sub-categories (structures, buildings and earthworks) are being delivered by Asset Management (lead: Mark Evans)
- Maintenance (all asset groups): delivered by NR maintenance division (lead: David Wynne)

Programme-level meetings are held on a weekly / two-weekly basis. We understand the main purpose of these meetings is to discuss progress, share findings and best-practice, gain insight into respective levels of progress and to share contacts.

5.3.2 Approach to data gathering and analysis

The approach followed by individual workstreams has varied considerably, with differing levels of progress in the benchmarking data gathering and analysis undertaken.

Each workstream carries out its benchmarking activities on a generally autonomous basis. NR's benchmarking workstreams have adopted a range of methodologies through which comparator data have been compared and analysed.

For the most part, undertaking a straight comparison of costs for specific activities / cost factors has not proven to be practicable or realistic, and NR has been able to gain only a limited amount of direct cost data obtained from comparator organisations in a handful of areas (documented further in Chapter 6). Experience from elsewhere suggests that establishing relationships with comparator organisations through voluntary benchmarking programmes requires considerable time and effort. The level of information shared early on in the process has tended to be limited in quantified terms, and not directly costed. (See Chapter 4 for detailed discussion around the process of establishing benchmarking groups).

To get around these problems, NR has adopted a number of other approaches to obtain and analyse comparator data, including:

- Site-based comparisons of individual jobs / activities, with quantitative comparisons of time and resourcing, and qualitative assessment of processes and procedures.
- Obtaining a cost estimate from comparator organisations for a pre-defined package of work.
- Analysis of policies and strategies for managing and delivering rail infrastructure, in particular how these drive activity scope (volumes and frequencies of work, process and organisational requirements, etc.).

We discuss the different approaches of the various workstreams in detail in Chapter 6.

5.3.3 Reporter opinion

Principal observations

We consider that NR's devolved benchmarking approach should enable teams with experience in front-line management and delivery of the assets utilise their knowledge and understanding of cost drivers and potential efficiency measures to inform the benchmarking approach. However, we do have concerns with regard to the alignment of individual workstreams with programme level objectives, as documented in the previous section.

We have found in general an open-minded and flexible approach is being taken to the benchmarking analysis by the benchmarking workstreams, with a clear focus on identifying efficiency factors relevant and meaningful for the particular characteristics of asset management and deliver in the given asset area.

On the whole, we have also found that NR benchmarking teams consider this to be a long-term process, whereby benchmarking analysis will be used to inform efficiency and performance improvements in both the short and the long-term, including beyond CP5. We support the long-term approach, and indeed consider that long-term commitment is an essential pre-requisite for deepening the level of cooperation and the extent to which data can be shared.

However, we also consider clarity is needed in the shorter-term period, leading up to the PR13 review with regard to the nature of benchmarking data that is and that

can be made available for the purposes of establishing CP5 efficient expenditure levels. We explore the benchmarking data requirements for PR13 further in Section 5.5).

5.4 Selection of comparator organisations

5.4.1 Top-down comparative analysis and LICB dataset

NR's identification of comparator organisations for bottom-up benchmarking began with the comparison of high-level maintenance and renewals expenditure figures from other European national infrastructure managers. Data were obtained by NR principally through the Lasting Infrastructure Costs Benchmarking group (LICB)¹⁸ but also through bilateral contact (e.g. data from France). This selection process has not included consideration of non-European rail organisations.

NR undertook a detailed process of normalisation and analysis of the high-level data, with the aim of identifying specific comparators that are managing their networks at lower cost. This involved undertaking a comparative analysis of renewals and maintenance costs across the main asset types (track, signalling, civils, electrification, etc.).

Whilst the level of relative efficiency across the different infrastructure organisations differed between asset categories, NR was able to identify the following general trends:

- Sweden, Germany and Belgium were identified as having generally the lowest infrastructure costs in relative terms.
- France, Italy, Holland, Switzerland and Norway were identified as having relative cost levels that vary across different asset categories. In terms of overall efficiency, these organisations are considered to be broadly comparable to Network Rail.

NR has established at least some degree of meaningful contact with all of the organisations listed above with the exception of Germany, which has indicated it does not wish to participate. However, perhaps inevitably, the level of cooperation has varied, resulting in differing levels of engagement with comparator organisations across the different workstreams (discussed in detail in Chapter 0).¹⁹

A number of other European rail infrastructure managers such as Ireland, Portugal and Luxembourg are not on the whole considered by NR to be suitable comparators, and have therefore not generally been targeted for comparison.

¹⁸ The LICB is a European rail benchmarking group involving fourteen European rail infrastructure operators managers who share benchmarking data annually. LICB was established through the UIC (Union Internationale des Chemins de Fer), a European rail industry association.

¹⁹ NR has indicated that it was unable to establish contact with the Spanish rail infrastructure operator (Adif).

5.4.2 Non-heavy rail comparator organisations

Alongside the national rail infrastructure operators, NR has also identified a number of non-heavy rail comparators, with which it intends to undertake benchmarking analysis. These comparators can broadly be grouped under the following four categories:

- Light rail or metro rail infrastructure operators.
- Operators of non-rail transport infrastructure (highways, airports).
- Utilities and telecoms companies operating fixed assets / networks.
- Other operators of built assets / infrastructure (public buildings, car parks, etc.).

Some non-heavy rail comparators (e.g. London Underground, BAA, Highways Agency) were identified at the commencement of the bottom-up benchmarking programme. Other organisations of relevance to specific asset groups have been identified by the workstreams themselves (e.g. communications companies identified by the telecoms workstream).

Since commencement of this assignment, Imperial College London (Imperial) and Arup have provided NR with a number of additional contacts for various organisations. This has mainly related to metro rail infrastructure operators, both in the UK and internationally (outside Europe). In particular, Imperial has identified relevant contacts through its work in the CoMET and Nova benchmarking groups. We include as **Error! Reference source not found.** a list of contacts provided by Imperial / Arup to date.

Whilst the original objective of comparator identification was to target those organisations considered more efficient for the given asset area, the workstreams have also in many cases sought to engage organisations that are deemed less efficient within the given asset area. NR considers such analysis may still enable useful information and insight into certain elements of efficiency and best practice to be gained, even if such organisations are less efficient in overall terms.

5.4.3 Summary table of comparator organisations

We summarise in the table below the comparator organisations with which the benchmarking workstreams has contact, or with which contact is planned.

Bottom up benchmarking workstream	Heavy rail comparators	Non heavy rail comparators
Track	France, Italy, Switzerland, Sweden	Strukton Rail.
Signalling	Norway, Netherlands, Switzerland (bilateral workstream); 7 x members of LICB ERTMS benchmarking group (cost workstream).	SSL, Siemens, Invensys.
Telecoms	Sweden, Netherlands, Italy, (Belgium)	Refer Telecom, (Highways Agency, BAA, Siemens, ATOS, Frequentis, UK shopping centres).
Electrical Power & Fixed Plant	France, Netherlands	London Underground, OfGEM, (UK electricity DNOs ²⁰ , Balfour Beatty, Siemens).
Buildings & Civils (IP-led workstream)	Italy, Sweden, Netherlands, France	BAA, Asda, National Grid.
Buildings (AM-led workstream)	New York City Transit, SNCF, Netherlands	BAM Nuttall, Defence Estates, University of British Columbia, (London Underground, Blue Water, DLR, BAA, UK local authorities).
Structures (AM-led workstream)	Sweden, Belgium, SNCF, Italy, Ireland, Netherlands, New York City Transit, (Norway, Switzerland, Amtrak)	Hampshire & Sussex County Councils, Asda, (London Underground, Highways Agency, British Waterways, Amey, Defence Estates, Thames Water, Yorkshire Water, National Grid).
Earthworks & drainage (AM-led workstream)	Sweden, SNCF, RailCorp Australia, Netherlands, Ireland, Kiwi Rail, Belgium, (Norway, Switzerland, Canadian National & Pacific)	London Underground, Highways Agency, (British Waterways, Yorkshire Water, Environment Agency).
Maintenance	France, Sweden, Netherlands, Italy, Norway	Balfour Beatty, (Ambulance service).

²⁰ DNO – Distribution Network Operator (electricity)

Note:

- (In brackets): organisations with whom contact is not yet formally established.

Table 6: bottom-up benchmarking comparator organisations

In addition to the organisations in the table above, Arup / Imperial have also provided NR with contacts within the following organisations:

- RailCorp (Australia) – light and heavy rail infrastructure operator.
- Hong Kong MTR – operator of metro rail network with some high density suburban rail routes.
- Nexus (UK) – operator of metro rail infrastructure north-east England.
- BAA – owner and operator of a number of major UK airports (and rail) assets.
- London Underground – operator of major metro rail network.

5.4.4 Reporter opinion

Principal observations

NR's process for identifying and establishing contact with comparators appears reasonable, given the constraints in available time and the lack of an effective, formal benchmarking process or information sharing protocol between European rail IMs at an international level. We have however have recommended NR extends its reach of comparator organisations to included non-EU peers.

For European rail comparators, we regard the LICB dataset as an appropriate initial reference from which to identify appropriate comparators – notwithstanding concerns about robustness and accuracy of the dataset expressed by NR. Workstreams have on the whole targeted data gathering from those railways identified as efficient for the given asset area. There does however appear to be some ambiguity in that certain workstreams have identified less efficient organisations for comparison, on the grounds that that NR can “still learn from” higher spending comparators. We reviewed the comparators selected by the respective workstreams further in Chapter 6.

As per our recommendation in Section 0.9, we consider that the work should be undertaken within the context of the efficiency gap – and that the relative efficiency data encompassed within the LICB dataset should help inform and highlight this.

For non-heavy rail organisation, s the process of identifying non-heavy rail comparators has on the whole taken into account the particular characteristics and requirements of the different asset areas; such comparators are particularly prominent within the signalling, civils and telecoms workstreams. We have found in general that the engagement of such comparators has been based on sound reasoning. There appears to be a clear orientation amongst the respective workstreams towards engaging with organisations with whom meaningful comparisons can be made, based on aspects of efficiency that are relevant and

material for that particular asset area. We discuss each workstreams' engagement with comparators in Chapter 6 of this report.

With regard to the level of cooperation and engagement with the comparators, we note (and NR also recognises) that many aspects of this programme are at an early stage and for the most part, comparator organisations are yet to provide significant volumes of benchmarking data. We consider a degree of uncertainty will always be present in terms of the level of data that NR will be able to secure from comparators.

To help mitigate this issue, we recommend that NR should aim to increase the level of data obtained as far as possible, both in relation to the quality, scope and depth of data from existing comparator organisations and in relation to identifying additional comparator organisations. We consider that establishment of a more effective international formal benchmarking framework is necessary for the longer term if the benchmarking activities and data exchange are to progress on a sustained basis. These timescales extend beyond the timeframe for PR13.

As with any exercise involving sharing of data between companies on a voluntary basis, it is worth noting that for NR's bottom-up programme, there is an element of opportunism in relation to comparator engagement. Establishing contact with comparators may not always follow structured, pre-determined selection process, and may instead be based on an opportunity to make contact through an existing link or relationship with the organisation, or by indications of the organisation's willingness to engage. We do not consider that this in itself undermines the potential meaningfulness of information obtained from the given comparator, as that will depend on the nature of the organisation and the data it provides, rather than the means by which contact was established.

Recommendations

- We recommend that NR develops a detailed engagement plan for each workstream, setting out specific steps through which it plans to initiate, establish and maintain contact. This should include both proposals for maintaining long-term, mutual engagement, and implementing procedures that ensure outputs shared and mutual interests served, thereby maximising the prospects for obtaining meaningful data. We outline several recommendations for establishing a benchmarking group in Section 4.3.
- We recommend that NR develops ties with further comparator organisations, looking more widely than the present comparator pool which is heavily oriented towards European rail organisations. In particular, we consider NR should focus on benchmarking with light-rail / metro organisations both within the UK and internationally, for which we consider there to be a significant level of potential comparability in spite of differences in infrastructure characteristics. We also consider NR should explore contacts with non-European heavy rail organisations. We provide specific recommendations for potential comparators for the individual benchmarking workstreams in the next chapter. (Note: recommended comparators and contact details have been provided by Arup/Imperial through feedback sessions provided for each workstream.)

5.5 Programme data

5.5.1 Renewals and maintenance expenditure levels

We set out in **Table 7** below NR's annual expenditure levels for each of the renewals asset categories, as well as total maintenance expenditure. These figures are based on the 2010/11 year-end figures reported in NR's 2010/11 Regulatory Accounts. We also indicate for which expenditure categories a bottom-up benchmarking workstream has been established.

Regulatory Accts 2010/11	Actual 2010/11 (£m)	% renewals spend	Benchmarking workstream
<i>Renewals expenditure</i>			
Track	605	27%	YES
Civils	356	16%	YES
Signalling	373	17%	YES
Telecoms	33	1%	YES
Electrical Power & Fixed Plant	78	3%	YES
Plant and machinery	99	4%	NO
Operational property	272	12%	YES
Fixed Telecoms Network & GSMR	215	10%	NO
Other renewals	203	9%	NO
Total renewals expenditure	2,234	-	
<i>Maintenance expenditure</i>			
Total maintenance expenditure	1,068	-	YES

Table 7: expenditure levels for benchmarked and non-benchmarked M&R asset categories

As indicated above, for renewals activities, only the categories “Plant and Machinery” and “Other Renewals” are not included within the bottom-up benchmarking workstreams²¹. The renewals categories for which a bottom-up benchmarking workstreams have been established account for 86% of total renewals expenditure.

The maintenance division has also established a benchmarking workstream to analyse maintenance expenditure.

5.5.2 Areas of focus and scope of data obtained

As stated previously, the benchmarking approach taken differs between individual workstreams. No detailed programme-level / top-down specification relating to scope of data or outputs required from the workstreams appears to exist. As a result, the specific areas of focus of the benchmarking analysis undertaken and the nature and scope of data obtained, differ across the programme.

²¹ NR has informed us that it is now establishing a benchmarking and efficiency workstream for Plant and Machinery.

Drawing on the analysis available, the professional judgement of senior staff appears to be the primary means by which scope for efficiency savings or performance improvements within each workstream is established. The approach and outputs from each of the ten bottom-up benchmarking workstreams are reviewed in detail in Chapter 6 of this report.

5.5.3 Benchmarking coverage within individual asset categories

We set out in the table below an overview of the areas of expenditure to which the numerical benchmarking data (provided to Arup for review) relate.

Expenditure category	Annual expenditure level (FY 11/12) (£m)	Numerically benchmarked activities	Indicative coverage level (£m / % overall spend)	
Renewals expenditure				
Track renewals	605	- S&C renewal - High-output track renewal	251 ²²	42% ²³
Signalling	373	- ERTMS renewal	18	5% ²⁴
Telecoms	248		0	0%
Electrical power & fixed plant	108		0	0%
Civils (structures & earthworks)	356	- Renewal of metallic underbridge ²⁵ - Renewal of footbridge - Earthwork renewals works	115	32%
Operational property	272		0	0%
Other renewals categories	272		0	n/a ²⁶
Total renewals expenditure	2,234		385	17%
Maintenance expenditure				
Total mtce expenditure	1,068	- Re-railing - Re-sleepering ²⁷	46	4% ²⁸

²² We note that Arup has not been given sight of quantified comparative data obtained for track renewals.

²³ Indicative coverage figure assumes that High-Output track renewal accounts for 25% of total Plain Line renewal expenditure.

²⁴ We note that NR expects a significantly increased level of quantified benchmarking data to be obtained in forthcoming signalling benchmarking activities, as discussed in Section 6.3.3.

²⁵ Indicative coverage based on estimated renewals expenditure on metallic underbridges of 40% of total underbridge renewals expenditure.

²⁶ Other renewals categories have not been reviewed by Arup, therefore the level of data obtained for such categories through the bottom-up benchmarking programme is unknown.

²⁷ NR has stated is that track inspections (£54m, 5%) and signalling maintenance (£85m, 8%) have also been benchmarked, although Arup has not been provided with specific details of these activities.

²⁸ We note that the findings from NR's bottom-up visits are relevant to a much wider range of activity than those directly compared. For example, NR has observed different working practices

Table 8 - indicative coverage level of bottom-up numerical benchmarking data

The coverage figures provided in **Table 8** are to be considered as indicative only, reflecting the quantified data that have been provided to us for review.

Although NR has provided its own spreadsheet containing indicative percentages of expenditure that are “covered” through the benchmarking analysis, we consider that further detailed analysis is needed to examine the degree to which data are sufficiently representative. This will require judgements to be made in relation to a number of important aspects including:

- The breadth of comparator data in terms of the number of cost elements / activities / projects reviewed relative to overall expenditure.
- The depth of data in terms of detail, granularity and supporting explanatory evidence, and the degree to which the data encompasses all key aspects / factors influencing cost.
- The level of comparability of data, and the degree to which adjustments / normalisation is required in order to gain “like-for-like” comparisons.

We have sought as far as possible to provide qualitative comments in relation to the breadth, detail and level of data obtained within our detailed review of individual benchmarking workstreams in Chapter 6.

However, we do not yet consider that we have been provided with a sufficiently detailed insight into the benchmarking data obtained to enable us to form definitive quantified judgments of coverage levels.

We recognise that for a number of workstreams the data gathering process is still in its early stages; an assessment of data coverage may be more appropriate once the workstream has progressed further over the next year. This will depend on the degree to which further data can be obtained within that timescale.

5.5.4 Reporter opinion

Principal observations

NR’s bottom-up benchmarking programme is still at a comparatively early stage and for the most part, the amount of quantified benchmarking data obtained has been limited.

Although benchmarking workstreams are in place for the vast majority of M&R expenditure categories (as indicated in **Table 7** above), in many areas the level of coverage and degree of representativeness is not clear from the information we have been provided with. We consider further clarification is needed on the precise nature and scope of benchmarking data obtained, including:

- the degree to which influencing factors are included / excluded, narrowing down focus on particular initiative rather than holistic understanding;

around possessions; if it is assumed that such costs account for one quarter of maintenance delivery costs, this relates to a cost base of £200m, which is 18% of total maintenance expenditure. We discuss the outputs from the maintenance benchmarking workstream further in Section 6.7.4.

- the level to which small number of cost observations are extrapolated to higher-level efficiency projections; and
- clarification as to the extent to which cost factors that are not specific to an individual workstream but which apply across all workstreams (e.g. HQ costs, corporate overheads, planning, PM) are encompassed within the benchmarking analysis.

We acknowledge that in many areas detailed benchmarking data have not yet been provided, or are limited. We propose to explore further with NR what further data can be obtained and used in such areas for the run up to PR13. For areas in which data have been obtained, we look forward to working with NR to gain more detailed insight into the benchmarking data.

Recommendations

- We recommend that NR should explore means through which the amount of data - and the level of detail - can be increased from existing comparators. This should include exploring means through which existing datasets can be deepened and more observations obtained, e.g. through greater number of activities, increased use of questionnaires / automated data sharing.
- We recommend that NR develops a set of coordinated benchmarking data specifications / criteria, taking programme level objectives (in relation to overall efficiency gap) and setting these out as explicit data requirements for each workstream. This should entail criteria to ensure a sufficiently representative dataset – e.g. criteria for breadth of data in relation to areas of expenditure, level of depth and detail, sample size, and nature of comparative data. This should also account for incremental progress as the benchmarking progresses and increasing level of data are obtained (timetable / “roadmap”).
- We recommend that each benchmarking workstream lead provides a detailed explanatory document setting out the key assumptions and rationale relating to the benchmarking data outputs. This should include:
 - details of incoming data adjustments and normalisation;
 - explanation of the process by which key cost / efficiency factors identified and prioritised, and those factors excluded / marginalised from the analysis; and
 - details of the application and extrapolation of the data to higher-level expenditure and efficiency proposals / projections.

5.6 Programme outputs

5.6.1 Introduction

As discussed previously in Section 5.1, NR's bottom-up benchmarking programme has been established not only to gain an understanding of the elements contributing to the efficiency gap but primarily to identify positive drivers and initiatives to improve the organisation's efficiency and performance.

NR's bottom-up benchmarking activities relate not only to CP5 proposals, but also to improvements and initiatives that have both a short-term / immediate impact and those that take effect in the longer-term (e.g. beyond CP5). NR has indicated that the programme will continue on an indefinite basis.

We would expect that both ORR and NR should regard any such programme to support both short and long-term efficiency improvements in a positive light. However, the particular focus of this mandate is to review the programme outputs specifically within the context the CP5 expenditure and efficiency proposals.

To this end, our review of the programme outputs is based around the following two NR submissions, which represent key milestones within the overall PR13 determination process:

- The IIP, published in September 2011, which sets out NR's current expenditure and efficiency proposals. The inputs from the bottom-up benchmarking programme into IIP form the principal focus of this initial report.
- The SBP, to be published in January 2013; this will set out in a greater level of detail, NR's expenditure and efficiency proposals as the analysis and development of evidence progresses. A review of inputs within the SBP will be completed in later stages of this assignment.

We set out our initial findings in relation to the IIP, as well as preliminary observations in relation to the SBP, in the sections 5.6.2 and 5.6.3 respectively.

5.6.2 Bottom-up benchmarking in the Initial Industry Plan (IIP)

Background and purpose

Network Rail's Initial Industry Plan (IIP), published on 30th September 2011 is a document produced by NR in cooperation with passenger and freight train operators, and overseen by the cross-party Rail Delivery Group (RDG)²⁹.

NR has also provided a supporting document to the IIP on 21st October, in (attached as **Error! Reference source not found.** to this document) explaining in qualitative terms the efficiency factors for each asset category, as well as provided quantified breakdown of CP5 efficiencies based on bottom-up benchmarking findings for track and buildings & civils (but not for other asset categories).

²⁹ RDG is "made up of the Chief Executives of the passenger and freight train operating owning groups and Network Rail, (and) has been established to provide leadership on cross-industry issues enabling a higher performing, more cost effective and sustainable rail network" IIP (p.4)

The purpose of the IIP is to “set out how the industry can deliver a more efficient and better value railway and how the railway can play a key role in driving sustainable economic growth.”

The IIP is based largely on “top-down” projections of NR efficiency for CP5, drawing upon target efficiency levels set out in the Rail Value for Money (RVfM) study and PR08 determination as the basis for efficiency proposals. The document includes qualitative assessments of a number of efficiency initiatives and improvements identified through the bottom-up benchmarking programme.

High-level CP5 efficiency projection

The IIP contains a projection of NR’s total CP5 costs, including the quantum of year-on-year efficiency savings to be achieved by the end of the 5-year control period. This equates to 16% efficiency for Network Rail over CP5 and represents a cost saving against NR’s total 2008/09 cost base of £1.8bn. This is reflected in the breakdown of projected costs that has been provided (see **Error! Reference source not found.**)

NR’s efficient CP5 expenditure totals set out in the IIP are based on the projections of work volumes, outputs and expenditure levels contained within its “Tier 0” and “Tier 1” strategic planning models.³⁰ NR has provided supporting data illustrating the levels of efficiency allocated across the various asset categories and sub-categories through the Tier 1 detailed cost modelling. This then feeds into the Tier 0 output model which provides a dashboard high-level overview of Tier 1 outputs and projected CP5 efficiency levels.³¹

We provide further, more detailed observations with regard to the IIP efficiency projections in **Error! Reference source not found.**

References to bottom-up benchmarking themes / findings

NR has indicated that the outputs from the bottom-up benchmarking programme have been used to validate a number of the key assumptions and inputs feeding into the Tier 1 calculations.

A number of factors or initiatives influencing efficiency levels are cited in the IIP in general, qualitative terms, which draw upon findings from the bottom-up benchmarking including:

³⁰ NR has provided spreadsheets showing the breakdown “pre-efficient” and “post-efficient” expenditure levels by asset category and sub-category, which feed into the CP5 total calculations. The spreadsheets illustrate how efficiencies have been allocated, taking into account both scope and unit cost efficiencies as well as other cost-reducing parameters for the respective categories.

³¹ We note that Arup has been commissioned under two separate Independent Reporter mandates to review in detail the Tier 1 models, including:

- an audit of accuracy and reliability of Tier 1 input data, including unit costs, asset data, intervention / alignment with asset policies and inputs for non-volume costs; and
- an audit of the computational accuracy and integrity of the modelling interfaces, calculations and outputs.

We recommend that further reference is made to the reports under these mandates (presently ongoing) for detailed analysis of the Tier 1 models.

- changes to asset policies to focus on targeted renewals, partial intervention and refurbishment;
- changes to contracting arrangements to make greater use of outsourcing, alterations to risk-sharing and liability arrangements;
- improved resource productivity, e.g. multi-skilling of staff, consolidation of possession arrangements, longer-term workbanks / planning;
- changes to standards and regulations to improve efficiency and risk orientation; and
- improved asset knowledge including condition-based activities, information systems and whole-life-costing approaches.

The above efficiency factors / initiatives are discussed (along with others) in qualitative terms within the IIP. No explicit quantified link is made within the document between these factors and the high-level efficient expenditure projections set out elsewhere in the document (see above).

However, as mentioned previously we have been provided with provided quantified breakdown of CP5 efficiencies based on bottom-up benchmarking findings for track and buildings & civils assets. We review these data in Sections 6.2.5 and 0 of this report.

We explore the various efficiency factors and initiatives for improvement being explored for the remaining asset workstreams in other relevant sections of Chapter **Error! Reference source not found.**

Further clarifications

As indicated above, NR has provided both documentation explaining the underlying assumptions relating to maintenance and renewals efficiencies projected in the IIP, and quantified breakdown of CP5 efficiencies for track and buildings & civils assets.

We are now proposing to gain further detailed insight into the way in which specific efficiency initiatives and improvements identified through the bottom-up benchmarking programme have been factored into NR's projected CP5 efficiency levels as the bottom-up benchmarking programme progresses.

5.6.3 Bottom-up benchmarking in the 2013 Strategic Business Plan (SBP)

The most important submission document from Network Rail in the PR13 determination process will be the Strategic Business Plan (SBP), which will be finalised in January 2013.

It is expected that SBP will contain more detailed and comprehensive proposals for its expenditure and efficiency during CP5. Whilst the IIP sets out proposals for the achievement of an efficiency level of 16%, it also states that "Network Rail will reflect progress on developing further its efficiency initiatives for CP5 in its Strategic Business Plan (SBP), to be published in January 2013." We understand that NR intends both to substantiate further the existing IIP efficiency project, and to identify opportunities for efficiency above and beyond the 16% level.

5.6.4 Reporter opinion

Principal observations

The benchmarking programme clearly has potential to support Network Rail in identifying significant opportunities for efficiency improvement. However, at this stage, it is not clear to what extent to which information obtained through the programme will be sufficient and appropriate to meaningfully inform projections of NR's CP5 expenditure and efficiency.

Our review to date has identified that the degree of progress in terms of data collection and analysis varies between individual workstreams (see Chapter 6). In general the level of detailed quantified benchmarking data obtained to date has been limited. Although NR has provided quantified bottom-up estimations of CP5 efficiencies buildings & civils and elements of track expenditures, the application of bottom-up benchmarking data to the CP5 efficiency projections has on the whole been limited to a validation role, based on qualitative judgements of the potential impact of efficiency factors identified, rather than any quantified, comparative analysis of bottom-up benchmarking data.

In order to acquire a body of bottom-up benchmarking data of sufficient scope, detail and quality to inform expenditure and efficiency projections in explicit, quantified terms, we consider that significant progress may well be required within a short space of time particularly with respect to data collection. The recommendations set out throughout this report support this objective.

We also consider such data are necessary to substantiate and strengthen business cases that NR is currently developing in relation to efficiency initiatives it is proposing to take forward.

6 NR bottom-up benchmarking – workstreams & data analysis

6.1 Introduction

We present in this chapter of our report a detailed review of each of the ten asset-based benchmarking workstreams, the data and outputs that they produce and how these are applied to NR's CP5 expenditure and efficiency proposals. Our analysis includes the following:

- **Background:** we review the type of activity and the current and projected levels of expenditure that the given workstream relates to.
- **Workstream approach:** we assess the approach followed for the workstream, and review how this links to the overall programme and other workstreams.
- **Workstream data:** this includes a review of comparator organisations with whom the workstream team has engaged, or plans to engage, and a review of the nature of data obtained.
- **Workstream outputs:** we analyse the way in which data obtained through the benchmarking workstreams have been interpreted and analysed, and their application to the development of CP5 expenditure and efficiency proposals and assessment of the efficiency gap.
- **Reporter opinion:** this includes our principal observations in relation to the given workstream, together with our view of potential additional comparators with whom the workstream team may wish to engage, together with any specific recommendations.

It is important to recognise that all of the workstreams reviewed are regarded by NR as ongoing projects which continue to develop and evolve. It is therefore very likely that the level of data obtained and degree of analysis undertaken will progress over the period leading up to PR13.

6.2 Track benchmarking

6.2.1 Introduction

NR's track benchmarking programme is led by the track Asset Management division, headed up by Nick deBellaigue.

6.2.2 Background

NR's total expenditure for track renewals in FY 2010/11 totalled £605m – which represents 27% of the year's overall renewals expenditure (£2,234m in total).

For CP5, NR's IIP projections indicate that for track renewals, a total efficiency of 21% will be achieved over the five-year control period, consisting of:

- 14% unit cost efficiency.
- 8% efficiencies to be achieved through NR's "Asset Information Strategy" (ORBIS).
- -1% (i.e. 1% inefficiency slightly offsetting the efficiencies listed above) relating to increased input prices.³²

NR indicates that it will "deliver around 14 per cent efficiency in CP5 in its S&C, conventional plain line and High Output delivery programmes." NR goes on to state that "[t]his efficiency is highly dependent on changing the approach to managing worksite safety and will therefore require a strong industry engagement." NR has also indicated that this improvement is "highly dependent on a further set of enablers that need to be delivered which will facilitate a significant increase in worksite productivity."

It is notable that the CP5 scope efficiency figure is nil. NR has stated that "no further scope efficiency to the efficiencies already embedded in our CP5 asset policies are proposed for track assets". However, NR has stated that scope efficiency will be achieved through its adoption of its asset policy for CP5, and it also expects inherent efficiency to be achieved through the adoption of its Asset Information Strategy (ORBIS).

6.2.3 Workstream approach

For track benchmarking NR has followed an activities based approach. This has involved site visits by teams of around half a dozen individuals totalling approximately fifty individual attendances³³ to observe and analyse specific activities first hand. The aim has been to identify efficiency opportunities through improvements in processes and best practice.

NR has developed a matrix setting out high-level topic areas to be focused on, including both specific, quantified observations of time, resources, productivity

³² NR's CP5 efficiency projections by asset category are included in **Error! Reference source not found. Error! Reference source not found.**

³³ Alongside its own personnel from its track Asset Management division, NR has also arranged for an engineer from Arup to attend the site visits, to make observations and write them up alongside the NR colleagues.

etc. and qualitative assessment of processes and policies shaping the work undertaken. This guides both the questioning of interview subjects and first-hand observations gathered on site visits. NR has sought to engage a wide range of track renewal and maintenance participants in its interviews on site (e.g. management, planners, contractors, engineers, etc.).

NR indicated the matrix-based approach has afforded greater versatility and flexibility to tackle often complex subject areas, compared to a fixed set of interview questions. These are seen as too rigid and less likely to produce meaningful answers.

Typically around 3-4 matrices have been completed for each site visit, which contain individuals' observations of the given activity and feedback gained through interviews. Following each visit, NR representatives draft 'site visit reports' to summarise observations, and the information from the visit reports is collated into a single matrix and visit report for analysis.

The benchmarking data obtained form a pool of qualitative, process-related information that NR has drawn together to inform the development of business cases for specific efficiency initiatives.

Although data are captured to a considerable degree of detail within individual site reports and the matrices, we have not been provided with copies of the completed matrices, nor with details of how NR utilises the information obtained to "compare back" directly with its own activities or cost factors. NR has indicated that it developed basic quantified comparative calculations of whole-life-costs to support business cases for efficiency initiatives. However, these data and calculations have not been provided for review within the timescales of this assignment.

NR has indicated that it has developed a business case model, which it is using to apply the learnings from its benchmarking visits to its own operations, and to value the benefits of their adoption. We may be able to review this model in detail in due course as part of our progressive assurance process (subject to discussions with NR and the ORR).

6.2.4 Workstream data

6.2.4.1 Engagement with comparators

NR identified France, Italy, Switzerland, Sweden as peer organisations that it considered appropriate for benchmarking against its own track renewals operations.

NR has sought to focus on activities that can be most feasibly compared to its European counterparts. To date, NR has undertaken the following eight site visits with all four rail organisations named above:

- Sweden: replacement of S&C module at Trelleborg (April 2011)
- France high-output plain line renewal, Evry, Paris (April 2011)

- Switzerland: S&C replacements at Dietikon and Vengeron (May 2011)
- Italy: S&C renewal, Codogno (June 2011)
- Switzerland: discussion of S&C renewal costs, Bern (July 2011)
- Sweden: S&C renewal, Stockholm (October 2011)
- Switzerland: S&C renewal, Lugano (November 2011)

NR has stated that it plans to continue with a limited number of further visits to refine its knowledge and to build stakeholder buy-in to its proposed strategy relating to more efficiency infrastructure delivery through engagement with contractors, although details of further visits have not been provided.

NR has stated that it has been able to “highlight the importance of its suppliers and of developing a contracting model that is significantly different from that in operation in CP4.” NR has also stated that, in addition to the site visits listed above, it has conducted 12 supplier interviews both in the UK and overseas, as well as a benchmarking visit to one of National Grid’s alliances, to inform the development of its contracting model. First hand details of the above interviews and site visits were not provided for review, although we understand that this analysis informed NR’s quantified CP5 efficiency proposals, which we review further in Section 6.2.5.2 below.

6.2.4.2 Nature of data obtained

We understand that information from the track benchmarking site visits is captured both through matrices and individuals’ site reports.

For the matrices, NR has indicated that high-level topic areas include key elements and components of expenditure to be compared and analysed including methods, materials, resourcing and delivering – timings, resourcing (labour, plant, materials), operational arrangements, safety procedures & provisions, etc.

NR has stated that it is taking a whole-life-costing approach to its analysis of relative cost levels. The matrix is designed to capture the information necessary to gain a full understanding of costs over time, focussing not only on the execution of the specific activity itself but also the underlying intervention policies, periodicities, asset condition and treatment of risk. As indicated previously, we have not had sight of the completed matrices.

In parallel, individual site reports / papers have been produced which cover a range of topics and observations at an individual level. The information captured through these site reports varies in form and detail. As well as providing generally more expansive commentary and detailed observations to support general information captured within the matrices, some other cost-related information not captured within the matrices is included. Examples include information on tender prices from four bidders (available in Sweden), informal high-level cost information picked up through interviews / conversations on site (e.g. statements about total costs of a given piece of work) as well as company level information such as total volumes annual volume data (in Switzerland). However, these observations are dispersed across different individuals’ reported observations and

are not evidenced or referenced to any official or documented source. There is a risk that such data have not been captured and structured within a quantified benchmarking or cost analysis framework.

NR has indicated that it may be constrained in gaining cost-related data from comparator organisations, and has suggested that other rail organisations do not appear to have the same level of detailed, activity-specific data available within NR. Furthermore, NR has highlighted difficulties in ensuring data are comparable, given that there are likely to be differences in the way that costs are defined and accounted for. For example NR found that design costs are generally the same percentage of total project costs across comparators but that some countries' flexibility defining a 'design job' could affect baseline costs.

Nevertheless, NR has indicated that it aims to work towards obtaining a greater degree of detailed comparator data and drill-down to gain a more detailed insight into potential efficiency measures.

6.2.4.3 Review of quantitative data

Quantified observations of resourcing levels and timings for track renewals delivery were included within matrices filled out on-site by NR observers. A number of documents containing qualitative comparisons and observations were provided, together with reports containing qualitative analysis.

NR's benchmarking observations relate to S&C and high-output track renewals from five overseas site visits. Such activities account for 41% of NR's total track renewals expenditure. We have not been given sight of the quantified observations gained through this workstream, and are therefore not able to give an assessment of the comprehensiveness of the quantified data obtained (i.e. we have no visibility of the cost elements observed relative to total cost levels for the given activity type).

As we have not been given sight of the quantitative information secured within the timescales of this assignment, we are not able to provide an opinion with regard to data quality. However, based on our understanding of the data reviewed, we consider it likely that further comparative observations would be required across a greater number of activities, in order to ensure a reasonable level of data robustness, and to ensure results gained are not anomalous.

6.2.5 Workstream outputs

6.2.5.1 Identification of key efficiency themes

Drawing upon the analysis and interpretation of activities and policies rather than cost data, NR has developed business cases on the basis of number of specific efficiency themes identified through the benchmarking workstreams. These include the following:

- Improved contractor relationships: NR is focusing on an increased level of input from contractors, with earlier involvement in design / planning phases and a greater degree of contractor autonomy in resourcing and delivering works, thereby incentivising more efficient practices.

- Multi-skilled staff: NR is analysing measuring enabling staff to perform a wider range of activities, enabling greater labour productivity levels.
- Possessions strategy: a key area of focus is the delivery of key track renewals activities. This includes delivery of S&C renewals, within a shorter track possession timeframe, thereby raising productivity and reducing cost, as well as increasing resource utilisation through facilitation of mid week delivery opportunities, reducing the dependency on weekends and bank holidays.
- Enabling the supply chain to form operational and commercial relationships between each other, reducing NR's involvement where possible and appropriate. Developing alliancing framework relationships with suppliers
- Safety-related efficiencies: NR is reviewing how efficiencies can be realised through more efficient, less burdensome safety requirements, which includes an increased focus on individual responsibility.

NR has inferred some generalities from its analyses. In some cases, European comparators have similar costs to NR but the infrastructure quality / design specification is lower. In a number of cases, where comparators may appear to have higher costs, NR has stated it is certain that there are still some elements of good practice that can be adopted to drive down its own costs, and that these findings add to the general understanding that informs the development of efficiency proposals.

6.2.5.2 Quantified CP5 efficiency proposals for S&C renewal

On 16th November 2011 the track benchmarking team presented quantified CP5 efficiency proposals which included presentation slides depicting projected efficiency levels. The proposals made reference to activities undertaken by the track benchmarking team including both meetings held with track contractors, and comparative benchmarking analysis with other European rail operators. NR indicated its proposals draw upon the areas of best practice identified from comparators including in relation to logistics, contractor arrangements, possessions and design processes.

NR's presentation slides include graphs depicting unit costs for Plain Line and S&C renewals comparing NR with four European comparators (France, Sweden, Italy and Switzerland). These two cost categories collectively account for 60% of NR track renewals expenditure. The source data and calculations from which these unit rates were derived have not been provided to Arup for review.

NR's presentation also identifies in qualitative terms a number of factors considered relevant to development of efficiency opportunities, which include:

- Various factors around the improved management of supplier relationships (improved interfaces, workbank stability, better packaging of activities, longer-term engagements / contractual terms)

- Improved risk sharing and greater control of the process by suppliers (direct planning, plant ownership, incentivising supplier innovation)
- Streamlined delivery and logistics (speedier possessions, greater productive time).
- Improved workforce productivity (multi-skilled staff, lower levels of casual staff, staff mobility, consolidated “high performing” teams) greater productivity of workforce,
- Simplified design, testing and commissioning process, (less onerous standards, increased levels of standardisation).

NR emphasised the challenging nature of the proposed changes being faced, but expressed a strong determination to achieve the proposed improvements, citing a number of key enablers that it considers are required to deliver the quantum of efficiency, as set out in *Figure 4* below:

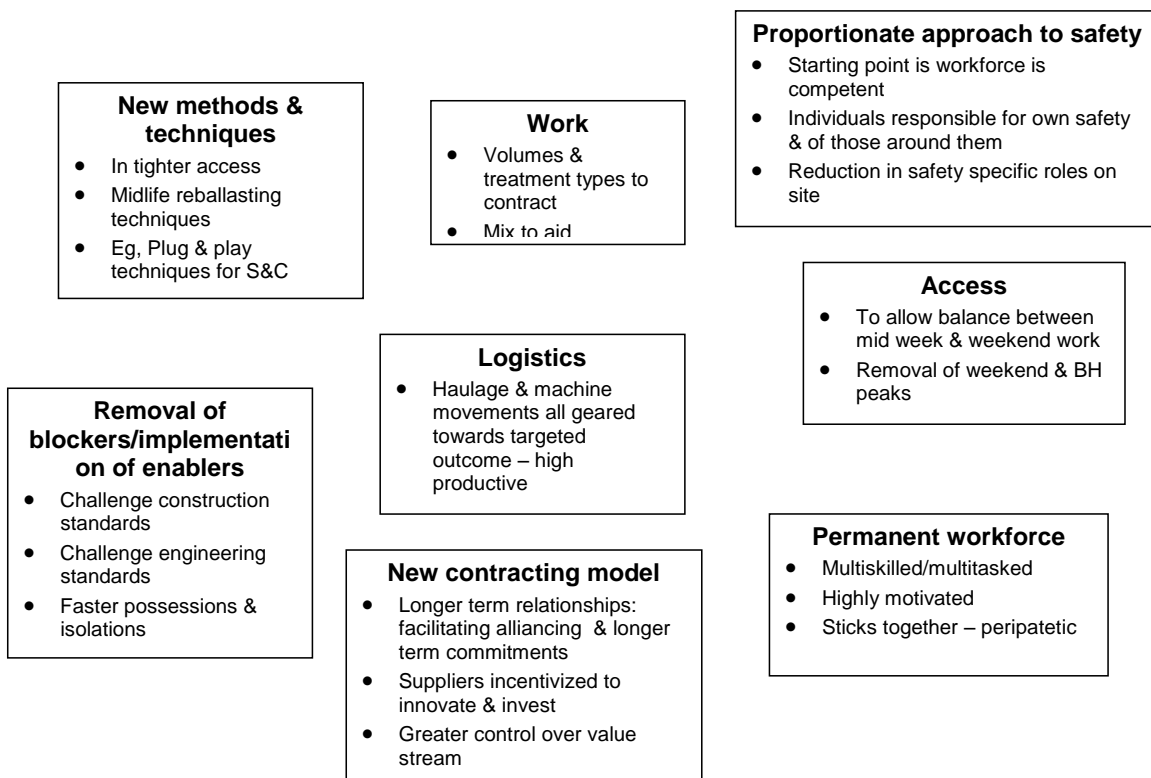


Figure 4: NR CP5 efficiency proposals: key enablers (source: NR)

For S&C renewals, NR has provided a quantified assessment of proposed 23.4% efficiency resulting upon the factors listed above. NR has provided a breakdown of current and future S&C unit cost, which we reproduced in *Figure 5*, below.

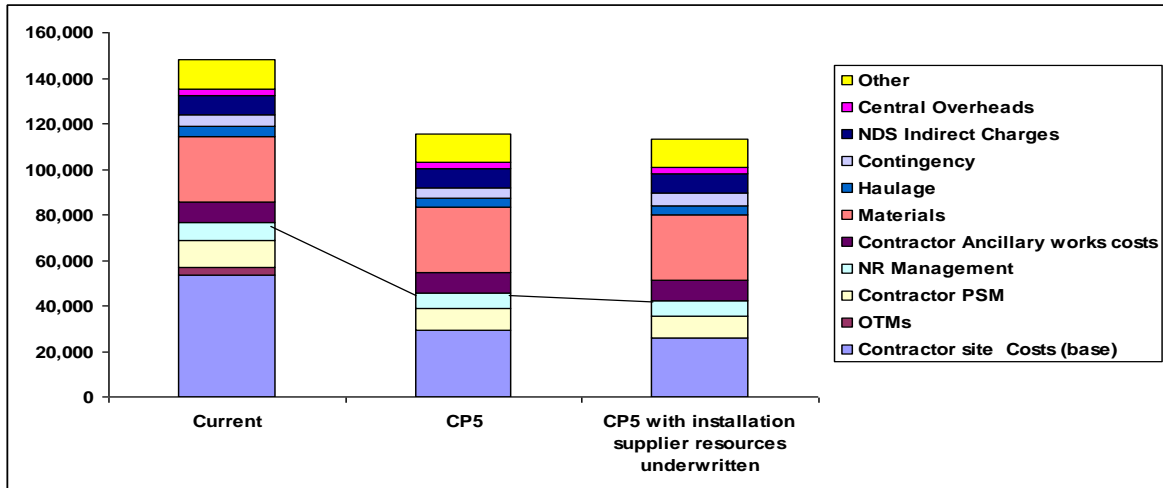


Figure 5 - Breakdown of projected S&C efficiency for CP5 (source: Network Rail)

As indicated above, the cost efficiency is achieved almost entirely through reduction in contractor site costs, with other cost factors largely unchanged. NR lists, in qualitative terms, the factors feeding into the contractor cost efficiency (although details of the quantified calculations have not been provided within the timescales of this review).

In terms of overall IIP efficiency projections, the proposals for S&C unit cost efficiency account for a significant proportion of NR’s overall projected unit cost efficiency of 14%, to be achieved for track renewals overall.

We understand that NR is proposing to develop and / or further elaborate proposed efficiency drivers relating to other track renewals categories.

For Plain Line track renewal, which accounts for a larger proportion of expenditure (combining conventional and high-output renewal accounts for %), NR has projected a slightly lower of 18% over CP5. To substantiate this projection, we understand NR is presently developing proposals for both conventional and high-output Plain Line renewals (which collectively account for the majority of remaining expenditure) on a similar basis to S&C.

For other areas of efficiency feeding into the IIP projections, NR indicated that it is identifying efficiency opportunities, and that developing proposals is a continual process. NR states that it proposes to develop proposals and progress measures to substantiate and implement proposed high-level efficiencies in the run-up to Control Period commencement.

6.2.5.3 Other efficiency proposals and degree of quantification

For a number of efficiency-related areas NR has begun the process of developing concrete efficiency proposals:

- With regard to safety related efficiencies, NR provided a report which quantifies the level of potential efficiency savings that NR considers may

be achievable through implementation of more efficient safety management.

- NR also provided a paper setting out discussions between NR and a number of key contractors (anonymised in the version provided), whereby proposed changes to supply conditions and increased involvement of the contractors in the delivery process are discussed. This involves providing suppliers the freedom (and incentive) to innovate and invest. NR has found that when it is acting in a client capacity it should ‘withdraw as far as possible’ from operational interfaces and suppliers should have direct relationships between each other. Although a number of concrete proposals for improved efficiency are set out, NR has not provided a quantified assessment of the efficiency impact.
- Following the review of practices through the benchmarking site visits, NR is in the process of developing concrete proposals for undertaking S&C replacements during overnight possessions on the UK network. NR recently carried out a trial overnight S&C replacement at Beeston, and has reviewed the extent to which the 2012/13 workbank might be delivered midweek. In addition it is holding meetings to review enablers and blockers in the change process³⁴. However, we note again that to date no quantified assessment of the efficiency impact of this initiative has been provided.

The utilisation of low-cost / standard designs for S&C components by comparator organisations is one factor identified as driving lower design and installation costs. However, NR indicated that although the installation of such items may reduce cost, a significant level of “upfront” investment would be required to facilitate such a change, which is not at present seen to be justified..

For other factors identified as key efficiency themes such as multi-skilled staff and reorganisation of delivery teams, NR has not provided any quantified analysis of the impact of such factors, or specific proposals for initiatives to increase efficiency in these areas to date.

Overall, NR’s activity-based analysis has produced a wide range of qualitative observations relating to various factors influencing cost and relative efficiency levels. These have fed into concrete efficiency proposals, for which NR has developed quantified, specific business cases (although we note that the detailed calculations informing these business cases were not provided within the timescales of this review).

6.2.6 Reporter opinion

Principal observations

We consider NR’s track benchmarking approach as a means through which useful comparative information relevant to the assessment of relative cost and potential efficiency measures can be gained. Overall, the workstream appears to be well

³⁴ We note that a two members of the Arup assignment team attended a workshop meeting on 20th October 2011, in which the implementation of such possessions and the associated challenges were discussed.

resourced with clear focus on identifying relevant and practical efficiency opportunities.

We consider that the scope of efficiency-related factors analysed through the benchmarking workstream to be sufficiently wide-ranging to identify relevant and meaningful drivers of cost that influence relative efficiency levels. The approach taken relies on first-hand observations from individual participants on site visits, which we consider should provide transparency in terms of data gathered and its utilisation for analysis. Observations are then collated to supporting in-depth analysis of a number of key factors influencing cost and efficiency levels including infrastructure characteristics, policy and planning process, and resourcing and delivery on-site.

We understand NR has captured quantitative information through completion of matrices with numerical observations, which have informed quantitative analysis of relative cost levels. However, we have not yet been provided with any of the matrices containing quantified observations, nor with any associated quantified data or analysis. Therefore the degree to which we are able to comment on the nature of data obtained is limited.

CP5 quantified efficiency proposals

Track is one of only two asset areas (alongside Buildings & Civils) for which NR has provided a quantified breakdown of CP5 efficiencies, drawing on bottom-up benchmarking findings.

NR's proposals for S&C renewals efficiency represent the most substantive quantified proposals that we have been provided with in relation to NR's CP5 efficiency projections. We consider that the proposals are based on reasonable logic, and are more specific in nature than any other asset area.

Arup has not been provided with any quantified source benchmarking from comparators that have fed into the S&C efficiency calculations within the timescales of this review, nor with the quantified calculations resulting in the 23.4% figure itself. We cannot, therefore, provide a direct assessment of the application of bottom-up benchmarking outputs to the efficiency proposals presented.

Overall, NR is projecting the achievement of efficiency savings in the delivery of S&C renewals that we consider to be significant, and which are – as acknowledged by NR – likely to represent a significant challenge to the organisation to achieve.

Recommendation: we recommend further analysis is needed to substantiate the proposals and mitigate potential risks, including:

- **Implementation costs:** we recommend that NR undertakes detailed analysis of the implementation costs of the factors identified (e.g. costs for NR team reorganisations, costs for training / recruiting skilled staff, implementation costs for new processes / systems / standards / working procedures, capex / equipment, e.g. track machines, site-testing equipment, etc.)
- **Consultation of proposals and buy-in from contractor organisations:** we recommend that NR explores the proposed measures in detail with contractor organisations. We consider it would be beneficial for NR to

continue consulting closely with contractors to confirm acceptance and secure buy-in as early as possible in the process.

- Assessment of risk: we recommend that NR undertakes a detailed assessment of risks associated with the efficiency proposals, develops mitigations, and makes a quantified assessment of the potential risk impact, in terms of the level of uncertainty associated with the efficiency factors currently projected.

NR has indicated that it is proposing to increase the level of quantified analysis as the benchmarking activities continue, including specifically in relation to Plain Line renewals. We look forward to reviewing these outputs further as the programme progresses.

Potential additional comparators

For the purposes of analysing relative efficiency for track renewals activities, we recommend that NR should aim, indicatively, to complete a benchmarking dataset which entails:

- engagement with approximately five European rail organisations in total (as long as at least three are classed as efficient), plus at least two other comparable IMs.
- analysis of the costs associated with approximately five activities per comparator.

We have recommended that NR explores engagement with the following as potential additional comparators:

- UK light-rail / metro system operators: e.g. London Underground and Newcastle Nexus.
- Non-European rail infrastructure managers: e.g. Hong-Kong MTR, New York City Transit, NSW Railcorp (Australia).

Recommendation: it has been recommended that NR undertakes comparative analysis between itself and its peers by costing work banks seen overseas as if they were undertaken in the UK. NR has indicated that it has attempted this approach and that it is working towards such analysis, but the initial results were unsatisfactory and could be misinterpreted.

6.3 Signalling benchmarking

6.3.1 Introduction

NR's signalling benchmarking is being centrally coordinated by the Asset Management team, and is being led by Rob Ireland.

6.3.2 Background

NR's total expenditure for signalling renewals in FY 2010/11 totalled £373m – which represents 17% of the year's overall renewals expenditure (£2,234 m in total).

For CP5, NR's IIP projections indicate that for signalling renewals, a total efficiency of 17% will be achieved over the five-year control period, comprising:

- 13% unit cost efficiency.
- 3% scope efficiency.
- 2% efficiencies to be achieved through NR's "Asset Information Strategy" (ORBIS).
- -1% (i.e. 1% inefficiency slightly offsetting the efficiencies listed above) relating to increased input prices.³⁵

NR attributes the achievement both of the 13% unit cost efficiency and the 3% scope efficiency to "asset management process improvements; optimised whole-life renewals requirements; whole rail-system optimised operational requirements and asset type standardisation." NR also cites "early locking down of workbanks" as a key enabler to realise the efficiencies described.³⁶

NR has indicated that its signalling benchmarking workstreams are focused on signalling delivery for the whole asset lifecycle.

6.3.3 Workstream approach

NR's signalling benchmarking is based on the following three defined workstreams, which combine quantitative and qualitative data gathering and analysis:

- Cost workstream, focused around on comparative quantitative data to support NR's understanding of signalling unit costs for assets.
- Bilateral workstream based on detailed and structured benchmarking on a "one-to-one" basis with other European rail organisations.

³⁵ We reproduce the NR's CP5 efficiency projections by asset category in **Error! Reference source not found.**

³⁶ We note that NR also applies this reasoning to the achievement of unit cost and scope efficiencies for both telecoms and Electrical Power & Fixed Plant renewals, for which exactly the same scope and unit cost efficiency levels (3% and 13% respectively) are projected. See Sections 6.4.2 and 6.5.2.

- Case studies workstream on suppliers' comparative analysis of projects delivered within the UK and elsewhere.

In terms of programme timing, an initial stage of work (Stage 1) was completed between February and May 2011. This work culminated in the production of a "Stage 1 Opportunities" report, setting out a range of potential efficiency opportunities NR has identified through the three signalling benchmarking workstreams. We review the Stage 1 Opportunities report further in Section 6.3.5 and each of the three workstream approaches in turn below.

NR has indicated that a further stage of work (Stage 2) entailing more detailed data gathering and analysis for each of the three workstreams with a greater level of quantified information is proposed up to the end of 2011. We understand that further workstages (3 and 4) are proposed for completion during 2012 although we have limited detailed of these latter workstages.

6.3.3.1 Cost workstream

The costs workstream is based on the sharing of signalling cost data with other European rail infrastructure operators .

The data shared thus far has been focused around ERTMS (European Rail Traffic Management System) infrastructure. ERTMS is a new type of signalling system that has been developed on the basis of Europe-wide standards. FY10/11 figures indicate that NR's ERTMS-related expenditure accounted for £18m - less than 5% of total signalling renewals expenditure. Although NR's utilisation of ERTMS to date has been limited to a single pilot scheme on a lightly used section of the rail network, NR is proposing to roll-out ERTMS deployment during CP5 to two major routes – the Great Western route during 2016-18, and the East Coast Main Line during 2018-20 – and is planning further deployment of the system beyond CP5. It is therefore anticipated that ERTMS will account for a higher proportion of expenditure than 5% in future (although we do not have a precise estimation of this figure).

This benchmarking activity is facilitated by the UIC, through the following two benchmarking groups: "On-board" ERTMS benchmarking and "Infrastructure" ERTMS benchmarking.

- "On-board" benchmarking relates to sharing of cost data for in-cab ERTMS signalling. NR has indicated that a significant amount of project data are available within this group and there are very few on-board signalling projects amongst the comparators' rail networks that are not included within the dataset.
- "Infrastructure" data relates to lineside ERTMS signalling infrastructure. NR has indicated that fewer projects are included within this dataset and that the normalisation of ERTMS infrastructure benchmarks can be challenging due to the network-specific characteristics of contributing projects.

In addition, we understand that as part of this workstream, NR has also undertaken further scrutiny and analysis of LICB cost data relating to other areas of signalling, including a review of normalisation assumptions.

NR has indicated that a key challenge will be to agree standard definitions for signalling projects between the comparators. As a means by which signalling costs can be measured on a unitised basis, the UIC's ERTMS benchmarking group is now adopting the SEU (Signalling Equivalent Unit) formulated by NR as the standard unit of measure, rather than the cost/km previously used. We understand that NR is involved in work to normalise the understanding of the SEU definition and gather metrics.

For Stage 1, NR has indicated that the outputs from the cost workstream have informed its cost and efficiency assumptions relating to on-board ERTMS. However, for confidentiality reasons, NR has been unable to share with Arup the actual cost data obtained through this workstream.

For Stage 2, NR has indicated that it will aim to improve further its understanding of the LCIB data, and continue with the ERTMS benchmarking. We understand that NR is proposing to work on a direct exchange of data with comparators to mitigate what NR considers to be the LICB's shortcomings. NR has also indicated that it proposes to make use of certain deliverables from the work undertaken by INESS (Integrated European Signalling System), which aims to standardise interfaces for signalling equipments across Europe by adopting ERTMS. The output from this work can support normalisation of benchmarking data.

NR has also indicated that alongside ERTMS, other areas of signalling infrastructure are captured by this workstream, although details of other areas have not been provided for review.

6.3.3.2 Bilateral workstream

The bilateral workstream is based around detailed, one-to-one sharing of benchmarking information between NR and other rail organisations.

For Stage 1, NR developed a detailed set of questions, which guided the process of data collection carried out through bilateral contact with the three other European rail organisations with whom NR has engaged through this workstream.

NR's questions related to overall characteristics of the comparator organisations, its delivery of maintenance and renewals activities, and the implementation of new technology. (We review the areas of focus in further detail in Section 6.3.4 below).

NR's data gathering approach involved telephone discussions and face-to-face meetings with the three comparator organisations.

Bilateral reports

For each comparator, a report that comprehensively detailing the results obtained has been produced. Each report contains an Executive Summary for the given comparator, containing the following:

- high-level cost comparison, comparing NR and the comparator's relative levels of total signalling maintenance and renewals expenditures.³⁷

³⁷ The high-level comparative expenditure data for the three comparators is obtained from "top-down" high-level LICB expenditure figures. NR provided us with a separate paper explaining how

- summary of key findings from comparator data given under the four key areas of focus (business drivers / renewals / maintenance / technology).
- “opportunities”: amongst the findings from the given bilateral benchmarking activities, NR highlights what it considers to be areas of potential opportunity; these are highlighted in bold within the text and assigned a reference number, which links the given opportunity to the Stage 1 Opportunities report (see below).
- “further work” proposed with the given comparator going forward.

Overall, NR’s bilateral approach based on direct one-to-one interaction has enabled qualitative comparator information relating to a wide range of factors associated with provision of signalling infrastructure to be obtained, to a high level of detail.

Stage 2 bilateral benchmarking

We understand that for Stage 2 of the benchmarking activities, NR intends work through the bilateral workstream to complete a “common infrastructure case study” with the three comparator organisations. NR’s benchmarking workstream will focus on areas where the greatest cost differences exist. NR’s benchmarking workstream also proposes to examine the project testing and commissioning phase and to study the inspection regimes.

6.3.3.3 Case studies workstream

The case studies workstream has the stated objective to “obtain comparative examples of work delivered by suppliers in the UK and abroad to understand cost and procedural differences.”

NR has indicated that the degree of success in obtaining data during Stage 1 of the signalling benchmarking has been limited. Constraining factors have related to issues of commercial sensitivity, the ongoing framework process and the ability to incentivise supplier engagement.

Nevertheless, we understand that NR requested suppliers to define particular opportunities that can be identified from overseas projects. NR also indicated that on the basis of this interaction it defined a number of opportunities itself.

Although NR has not provided reports or documentation detailing the actual engagement with the suppliers and comparisons made, a table has been providing summarising efficiency opportunities, of which around half have been provided by suppliers with the other half identified by NR itself. The table contains a brief description of the opportunity, together with an indication of associated action required and responsible person within NR.

For each opportunity listed in the table, an indicative percentage saving is also given; we have not been provided with an explanation of how this estimated saving is derived.

the high-level signalling expenditure comparisons were obtained from the raw data within the LICB dataset.

As with the bilateral benchmarking workstream, each of the opportunities set out from the case studies workstream feeds into the Stage 1 Opportunities report, which review further in Section 6.3.5 below.

For Stage 2, NR has stated it plans to formalize the approach with the three supplier organisations, and that it has identified specific case studies for which it proposes to undertake further benchmarking with the suppliers.

6.3.4 Workstream data

6.3.4.1 Cost workstream

Engagement with comparator organisations

We understand that seven European comparator organisations (including NR) participate in the two UIC ERTMS benchmarking groups. However, for confidentiality reasons NR has been unable to disclose the identity of the six comparator organisations.

Nature of data obtained

As indicated above, NR has been unable to share any first-hand data obtained through the cost workstream for confidentiality reasons. As a result, we have no insight into the scope, breakdown and level of detail of this dataset.

NR has stated that its Stage 1 report contains LICB data only in raw format, and that further analysis was not undertaken due to concerns regarding the robustness of the data obtained.

NR has provided a short report summarising in brief, qualitative terms the high-level conclusions drawn from the cost workstream. This states that NR considers the “on-board” data can be utilised to validate cost assumptions that inform efficiency projections, but that the “infrastructure” data are of insufficient quality for such a purpose.

However, it is not possible for us to comment further without visibility of the data involved and an explanation as to which specific assumptions this refers to, and how this relates to NR’s CP5 signalling expenditure and efficiency assumptions.

NR has indicated that it may be possible for further details of the data gathered to be shared for review, subject to agreement from the UIC. We consider this may be relevant for further review in due course as part of our progressive assurance process (subject to discussions with NR and the ORR).

6.3.4.2 Bilateral workstream

Engagement with comparator organisations

For the bilateral workstream, NR has engaged with the following three European rail infrastructure operators:

- SBB (Switzerland);
- Prorail (Netherlands); and

- Jernbaneverket (Norway).

NR has indicated that these three peers were not selected on the basis of cost alone, but also on their comparability, ability to support the work, and whether or not existing relevant relationships or studies existed.

Nature of data obtained

The bilateral workstream has produced by far the most detailed and comprehensive data of all the three workstreams. NR was able to obtain detailed information through responses gained to its questions. The questions were based around the following four key areas:

- “Business drivers”: this includes questions relating to institutional, regulatory and financial context within which the comparator organisation operates, with an emphasis on how such factors influence signalling asset policy and delivery.
- “Renewals”: this includes questions relating to renewals planning and prioritisation, interventions policy (including the dynamic between maintenance- and renewals-driven policy), typical scope and approach to delivering signalling renewals, procurement and contractor arrangements including in-house vs. outsourcing activities, design and supply chain, and organisational issues.
- “Maintenance”: this includes questions relating to maintenance policy, organisation and planning, delivery and asset monitoring
- “Technology”: this includes questions regarding deployment of ERTMS, measures to make the comparator organisation more efficient in signalling provision, comparator’s approach to replacement of legacy systems and specification of new technology requirements, and the organisations’ policy and proposals for ERTMS implementation.

NR has recorded detailed answers received from each of the three comparator organisations. This has enabled a significant degree of qualitative explanatory information to be obtained in relation the three comparators’ approach to management and delivery of signalling infrastructure, which has informed NR’s Stage 1 Opportunities report.

6.3.4.3 Case studies workstream

Engagement with comparator organisations

For the case studies workstream, NR has engaged with the following three rail signalling suppliers:

- SSL: benchmark project delivered in Holland.
- with Siemens: benchmark project in Germany delivered for DBNetz.
- Invensys – benchmark project from UK.

NR indicated it also proposes to engage a supplier organisation with project data from Finland for Stage 2 of the benchmarking.

Nature of data obtained

As set out in Section 6.3.3, a number of opportunities deriving from the case-studies benchmarking interaction have been identified and assigned an indicative percentage potential cost savings. However, we have not gained visibility of any actual benchmarking data or underlying analysis feeding into the opportunities stated. Therefore no comment on the nature of data obtained can be provided for this area.

6.3.4.4 Review of quantitative data

Quantified observations of relative cost levels relating to ERTMS infrastructure have been obtained through a UIC-based benchmarking group.

We understand that initial quantitative observations from the costs workstream relate only to ERTMS signalling infrastructure, which we understand relates to around 5% of current signalling renewals expenditure. We have not been given sight of the quantified observations gained through this workstream, so are not able to give an assessment of the comprehensiveness of the quantified data obtained (i.e. no visibility of the cost elements observed relative to total cost levels for the given activity type).

Quantified comparator information is shared between seven comparators (including NR), across a large number of schemes. We consider it likely that this dataset would be of a breadth sufficient to ensure reasonably robust data. However, data quality and robustness also depends on actual form and format of input costs, granularity and breakdown of costs and process for normalization. Because NR has not been able to show Arup the data itself, we are unable to form a definitive judgment at this stage.

NR has indicated that findings from the bottom-up signalling benchmarking have informed the CP5 signalling efficiency projection presented in the IIP. However, no documentation has been provided to explain how the bottom-up benchmarking results were used, in conjunction with other analyses, to arrive at the figures stated in the IIP. We understand that the CP5 efficiency proposals presented to date have been effectively high-level estimations, but details of the rationale behind the judgements made have not been provided. NR's programme plan defines a number of workstreams that relate to the "quantification of benefits", which were due for delivery in November 2011. We have not yet been provided with the data implied by that milestone.

6.3.5 Workstream outputs

6.3.5.1 Stage 1 Opportunities Report

NR has drawn together the findings from across the three workstreams during its first stage into the single Stage 1 Opportunities Report.

This document collates in a single table the 33 "efficiency opportunities", of which 22 were identified through the bilateral workstream and the remaining 11 through the case study workstream. The opportunities are only set out in very high-level, indicative terms within the table. No quantification is included, with NR simply categorising whether the given opportunity falls into the "embedded scope" or "value engineering" category, or both.

The opportunities appear to cover a broad range of different themes; this is reflective of the broad scope of the benchmarking approach, in particular for the bilateral workstream (see above). Areas of focus within the opportunities include organisational structure and project management, asset policy and risk, works planning, contracting and procurement including design and supplier arrangements, technological deployment.

NR then sets out a table a list containing 16 “priority” opportunities out of the original 33. The table contains two columns setting out the following:

- “Maximum potential benefits”: NR utilises a low / medium / high indication of estimated potential efficiency impact, whereby:
 - “Low” represents less than 1% impact on total signalling expenditure.
 - “Medium” represents an impact of between 1% and 3% on total signalling expenditure.
 - “High” represents an impacts of over 3% total signalling expenditure.
- “Estimated implementation difficulty factor” (low / medium / high).

The final part of the Efficiency Opportunities Report sets out as a next step, “further and more detailed analysis to quantify Stage 1 efficiencies identified above.” We understand this analysis will be undertaken by the efficiencies workstream, based on Stage 2 outputs from the signalling benchmarking activities.

6.3.5.2 Signalling benchmarking input into IIP expenditure and efficiency proposals

NR has indicated that, as a general principle, its focus has been on identifying efficiency opportunities from comparators that can be practicably adopted and delivered.

However, the information provided in relation to signalling benchmarking activities has not precisely defined or explained precisely how in quantified terms the benchmarking outputs or efficiency factors identified have fed into CP5 signalling expenditure and efficiency proposals set out within the IIP.

NR provided Arup with a short paper in relation to the relationship between the outputs from Stage 1 signalling benchmarking activities and the development of the IIP proposals. However, from the information provided it would appear that the input from the programme has been purely in relation to proposals and assumptions relating to signalling asset policy. NR states that programme outputs formed part of the process by which proposals for Signalling Asset policy supporting IIP were developed. No documentation or explanation has yet been provided explaining exactly which specific information / outputs were part of this process, and precisely how this input process was carried out.

6.3.6 Reporter opinion – signalling benchmarking

Principal observations

We consider the signalling benchmarking one of the most well developed and wide-ranging of the benchmarking areas, with three coordinated workstreams approaching comparative analysis from different angles.

Initial work through all three workstreams was undertaken according to a defined “Stage 1” timeframe, with results collated and combined into a summary “Stage One Opportunities” report. We understand that “Stage 2” of the work, involving continuation of all three workstreams, is ongoing.

The focus of analysis appears to have been wide-ranging. The bilateral workstream set out a defined and detailed set of questions through which information was collated from the three European rail comparators, taking into account a range of perspectives associated with asset / technological specification requirements, intervention policy, maintenance and renewals performance and delivery processes.

The “case studies” workstream entailed a more informal / opportunistic approach of engaging with suppliers to provide an outline of what they to be potential areas of efficiency. NR’s proposals for Stage 2, based on comparative cost analysis using a case studies, appears reasonable for furthering the scope of this benchmarking area.

We understand that through the “costs” workstream useful comparative cost information was obtained through the comparison of “on-board” ERTMS projects, although the quality of “infrastructure” ERTMS data were limited. Whilst ERTMS-related expenditure currently represents only a small proportion of NR’s overall signalling expenditure, its importance is likely to increase in the medium- to long-term as NR rolls-out ERTMS systems across major routes on the UK rail network during CP5 and beyond. For confidentiality reasons, NR was unable to share the results from this benchmarking, therefore we are unable to comment further with regard to the data obtained in this area.

Overall, a range of qualitative data have been obtained through the signalling benchmarking, from which NR has identified a defined set of efficiency opportunities. A report has been produced, which includes an assessment in qualitative terms the potential opportunities of a selection of the opportunities identified. However, the outputs we have been able to review from the signalling workstream are almost entirely qualitative in nature.

We consider that NR’s initial engagement with comparator organisations through the bilateral workstream in particular forms a reasonable basis for deepening the relationship and supporting advancement of the benchmarking activities to a more quantitative basis. To further this process it will be essential for NR to base its activities on the basis of cooperative engagement, based on common interest, and the sharing of results and analysis.

We consider that the challenge will be to progress the level of data obtained, to include a greater level of quantified information. On this basis NR will be able to combine quantitative data with the qualitative insights already gained to develop a more complete and tangible overview of relative efficiency levels and opportunities for improvement. Groundwork has been done that lays foundations for effective engagement, but closer and deeper cooperation necessary to develop quantified data to a greater level of depth and detail.

Potential additional comparators

For the purposes of analysing relative efficiency for signalling renewals activities, we support NR's proposals to engage with 3-4 further IMs. We understand NR is already in discussions with Germany and France, and is also considering engaging with Sweden and possibly Hong Kong and North America). In terms of the scope of project data, we consider NR should aim, indicatively, to complete a benchmarking dataset which entails approximately five comparator projects per IM.

We have recommended that NR explores engagement with UK light-rail / metro system operators: e.g. London Underground and Newcastle Nexus as potential additional comparators.

6.4 Telecoms benchmarking

6.4.1 Introduction

NR's telecoms benchmarking is being led by the newly formed internal telecoms organisation Network Rail Telecoms (NRT), part of the Asset Management division. The telecoms workstream is being headed up by Fraser Allen.

6.4.2 Background

NR's total expenditure for telecoms renewals in FY 2010/11 totalled £32.5m – which represents 1.5% of the year's overall renewals expenditure (£2,234 m in total). For CP5, NR's IIP projections indicate that for telecoms renewals, a total efficiency of 15% will be achieved over the five-year control period, consisting of:

- 13% unit cost efficiency.
- 3% scope efficiency.
- -1% (i.e. 1% inefficiency slightly offsetting the efficiencies listed above) relating to increased input prices.³⁸

NR's unit cost and scope efficiency projects of 13% and 3% respectively match the efficiencies in these categories for signalling, and NR attributes their achievement to exactly the same factors as for signalling (asset management process improvements, optimised whole-life renewals requirements, etc. - see Section 6.3.2).

6.4.3 Workstream approach

NR has set up the telecoms benchmarking workstream to follow the same format as that used for signalling benchmarking. However, the telecoms benchmarking only commenced in September 2011, with NR estimating it to be 10-12 months behind signalling. As a result, the telecoms benchmarking is one of the least advanced.

As with other workstreams, the aim of the telecoms benchmarking is to identify both efficiency factors that can feed into CP5 efficiency proposals within the SBP, and long-term future business planning beyond CP5.

NR is proposing to focus on GSMR network operations with individual best practice case studies highlighting possible efficiency benefits, rather than on the maintenance and renewal of a set of fixed, linear assets. Whilst NR provision of telecoms infrastructures has traditionally been through its fixed telecoms networks (FTNs) a reliable asset bespoke for rail use, the recent rapid developments in communication technology and growing demands for information have led NR to prioritise its focus for telecoms provision on GSMR network operations.

NR now considers technological obsolescence an area of risk, and it intends NRT to function as "service provider", to ensure telecoms provision and assets are oriented to service-based requirements.

³⁸ We reproduce the NR's CP5 efficiency projections by asset category in **Error! Reference source not found.**

This change in focus is reflected in NR's benchmarking workstream, which is orientated towards high-cost, large-scale cost elements with GSM-R network provision identified as a principle area of focus.

NR has identified a number of issues / challenges likely to arise in obtaining comparable benchmarking data including:

- Differing technical characteristics, scope and structure of comparator telecoms networks.
- Level of cost data availability, as well as categorisation and breakdown (i.e. how cost has been accounted for, for the given comparator network / data).
- Underlying factors influencing nature of network provision and its costs, such as policy and regulatory requirements, licensing.

Given the challenge and potential complexity associated with benchmarking telecoms network provision, NR has indicated it proposes to draw upon information from the benchmarking study to build-up a "fictitious" project based on comparators' cost factor data, but using its own unit costs.

We propose to review the further approach taken in forthcoming stages of this assignment, as it develops and evolves.

6.4.4 Workstream data

Engagement with comparator organisations

The telecoms benchmarking has initially focused on establishing contact with organisations involved in provision of GSM-R (for rail) purposes.

We understand NR has already held telephone discussions with the following three organisations, and that initial meetings are planned to begin the process of sharing data up to the end of 2011:

- ProRail (Netherlands) – NR has identified that ProRail has a similar network structure for GSM-R provision, but very different approach to procurement, with GSM-R management & operation fully outsourced (in contrast to NR which now manages GSM-R in-house).
- Comtest – Comtest is the supplier of GSM-R to RFI (Italy).
- Refer Telecom – the Portuguese rail telecoms supplier which has implemented new technology and control centre consolidation programmes.

NR has indicated that it is exploring a number of further potential contacts, including:

- Other European rail organisations: Belgium and Sweden.
- Other transport infrastructure operators including London Underground, Hong Kong metro, Highways Agency, BAA and Frequentis (air traffic communications).

- Other (non-transport specific) operators / suppliers of telecoms infrastructure such as Siemens, ATOS, UK shopping centres / retail sectors operating CCTV / public communications systems.

Nature of data obtained

We understand that for telecoms no detailed benchmarking data have yet been obtained from the comparator organisations, but we understand that NR is proposing to gather initial data over the next 4 – 6 months, which we expect to be able to review in later stages of this assignment.

Review of quantitative data

NR indicated that it was expecting to obtain initial comparator data by the end of 2011, as reflected in the “Analysis and Draft Reports” workstreams set out in the programme plan during this period. However, no quantitative data have been provided to date.

NR’s programme plan indicates that it is proposing to develop unit costs to feed into the SBP submission during early-mid 2012.

6.4.5 Workstream outputs

At the time of writing, no data collection or analysis for buildings has yet been provided, but we expect to be able to review workstream outputs in later stages of this assignment.

6.4.6 Reporter opinion – telecoms benchmarking

Principal observations

The telecoms benchmarking workstream is the most recent of all workstreams in terms of its commencement date, and as a result one of the least advanced. The process of engagement with comparators began only recently, and no benchmarking data or analysis has yet been provided for review.

Overall, the workstream is heavily oriented towards comparative analysis of telecoms provision from the perspective of service provision and bandwidth availability (rather than in relation to costs for a set of fixed assets). NR is engaging with both rail and non-rail in order to compare both approaches taken to telecoms provision, and comparative costs. The most significant component of cost being focused on relates to GSM-R system implementation.

We consider this to be an appropriate approach, given the evolving nature of telecoms provision and technological advancement. However, given the recent nature of this benchmarking workstream, we consider there to be uncertainty at this stage of the extent of benchmarking analysis and findings that will be achievable within timescales of the PR13 determination.

We expect to be able to comment further on the telecoms benchmarking as it progress in later stages of this assignment.

Potential additional comparators

We have recommended that NR explores engagement with the following:

- GSM-R benchmarking group set up through the LICB.
- SNCF has implemented its GSM-R under a PPP concession arrangement which may provide interesting insights into this procurement method.
- Non-European rail comparators, e.g. Hong Kong MTR, NSW Railcorp (Australia), and metro rail organisations, e.g. Newcastle Nexus.

6.5 Electrical Power & Fixed Plant benchmarking

6.5.1 Introduction

NR's Electrical Power & Fixed Plant benchmarking is being led by the Asset Management team, and is being headed up by Adrian Murray.

6.5.2 Background

NR's total expenditure for Electrical Power & Fixed Plant renewals in FY 2010/11 totalled £78m – which represents 3% of the year's overall renewals expenditure (£2,234 m in total).

For CP5, NR's IIP projections indicate that for Electrical Power & Fixed Plant renewals, a total efficiency of 22% will be achieved over the five-year control period, consisting of:

- 13% unit cost efficiency.
- 3% scope efficiency.
- 7% efficiencies to be achieved through NR's "Asset Information Strategy" (ORBIS).
- -1% (i.e. 1% inefficiency slightly offsetting the efficiencies listed above) relating to increased input prices.³⁹

NR's unit cost and scope efficiency projects of 13% and 3% respectively match the efficiencies in these categories for both signalling and telecoms, and NR attributes their achievement to exactly the same factors as for signalling (asset management process improvements, optimised whole-life renewals requirements, etc. - see Section 6.3.2).

6.5.2.1 Workstream approach

NR's Electrical Power & Fixed Plant (E&P) benchmarking undertaken to date has been based around a desktop analysis of cost data relating to UK power distribution network operators (DNOs). This relates to comparative cost data for elements of E&P infrastructure.

We understand NR proposes to widen its benchmarking approach. This includes participation in a European E&P benchmarking group involving other European

³⁹ We reproduce the NR's CP5 efficiency projections by asset category in **Error! Reference source not found. Error! Reference source not found.**

rail infrastructure operators, Electrical Power & Fixed Plant infrastructure suppliers (e.g. Siemens), rail contractors (e.g. Balfour Beatty). We have not been provided with details of this group, and the nature or scope of information that it is proposing to share.

NR has also indicated that it is investigating further potential sources of benchmarking data.

At the time of writing, no documentation relating to E&P benchmarking programme has been provided.

We understand that NR is taking significant measures to gain a more complete overview and understanding of its own internal electrification and power expenditure, including, in particular, its asset requirements and stability in forward-looking workbanks for CP5. We would regard NR's ability to gain an accurate overview of its own current and prospective E&P expenditure to be an essential precondition for comparative analysis of costs with external comparators.

6.5.3 Workstream data

Engagement with comparator organisations

As indicated above, NR is proposing to engage with the European E&P benchmarking group. NR indicated it is unsure of the degree of information it may be able to obtain through its involvement.

NR also indicated it is proposing to investigate EU sources further to gain more detailed information, and that visits are proposed in the coming weeks; however, no details of the visits have been provided.

Nature of data obtained

NR has undertaken a desktop-based benchmarking analysis of unit costs, using information relating to UK DNOs available through OfGem. Using this data, NR has compared its own costs to those of the DNOs for the following components of E&P infrastructure:

- HV cable renewal
- Circuit breaker renewal
- Transformers
- OLE

NR presented the results of this benchmarking at the meeting on 17th August 2011. However, no documentation relating to the input data or calculation methodology has yet been provided.

We expect to be able to review data obtained in later stages of this assignment, as NR progresses in exploring and establishing contact with further sources of benchmarking data in Europe.

Review of quantitative data

Within its programme of bottom-up benchmarking, NR undertook a high level comparison of the cost of electrification components, drawing on publicly-available information provided by Ofgem in relation to electricity DNOs. The quantitative data related to transformers, OLE, switch gear and high voltage cables.

As we have not been given sight of the quantitative information gained, we are not able to provide an opinion with regard to data quality. However, based on our understanding of the work undertaken by NR, we consider it is likely that further comparative observations would be required across a greater number of activities, in order to ensure a reasonable level of data robustness, and to ensure results gained are not anomalous.

In the course of our work, we have reviewed NR's programme plan, which suggests that further progress with regard to data quantification can be expected prior to the preparation of the SBP. At this stage, we have not been provided with information to describe how the data available were used, alongside other relevant analyses, to derive the numbers presented in the IIP.

6.5.4 Workstream outputs

No data collection or analysis for E&P benchmarking has yet been provided, but we expect to be able to review workstream outputs in later stages of this assignment.

6.5.5 Reporter opinion –Electrical Power & Fixed Plant benchmarking

Principal observations

Electrical Power & Fixed Plant benchmarking is one of the least advanced workstreams. Although NR indicated during benchmarking meetings that it was proposing to engage with European IMs – including at the European E&P benchmarking group – it is not clear whether this engagement has started, and no documentation presenting data or analysis obtained through the benchmarking activities has been provided for review.

NR's desk-based comparative study of relative costs for E&P infrastructure components enabled a quantified comparison to be made; however, we consider that order for an appropriate scope and range of comparator data to be obtained beyond this initial study, direct engagement with comparator organisations is necessary on a similar basis to other bottom-up benchmarking workstreams.

We expect to be able to comment further on the Electrical Power & Fixed Plant benchmarking as it progress in later stages of this assignment.

Potential additional comparators

For the purposes of analysing relative efficiency for E&P renewals activities, we recommend that NR should aim, indicatively, to complete a benchmarking dataset which entails:

- engagement with approximately five European rail organisations in total (as long as at least three are classed as efficient).
- analysis of the costs associated with approximately five activities per comparator, using detailed drill-down unit-cost or activity-based benchmarking.

We support NR's proposals to engage with 3-4 further IMs (which we understand may include Germany, France and Sweden and possibly Hong Kong and Canada). In terms of the scope of project data, we consider NR should aim, indicatively, to complete a benchmarking dataset which entails approximately five comparator projects per IM.

We have recommended that NR explores engagement with the following as potential additional comparators:

- Non-European IMs with conventional electrification assets: e.g. Hong-Kong MTR, New York City Transit, NSW Railcorp (Australia).
- IMs that operate "third-rail" based electrification assets: e.g. London Underground, MTA Long Island Railroad and MTA Metro North (which both operate out of New York.)

Recommendation: We suggest the merits of a Bayesian approach for the benchmarking analysis. NR considers a probabilistic approach more suitable, because the Bayesian method is seen as a "black box." Such perceptions might be overcome through transparency, if inputs & assumptions are clearly laid-out.

6.6 Civils benchmarking (buildings, structures, earthworks)

6.6.1 Introduction

NR's civils benchmarking work comprises five individual workstreams.

- Firstly, NR's Investment Projects (IP) division headed up a workstream that began its activities in mid-2010, completing an initial benchmarking study in January 2011 comparing costs for four civils schemes.
- The IP team has since commenced a further study relating to best practices across rail and non-rail sectors; this work is presently ongoing.
- NR's AM division has also initiated its own benchmarking workstreams for each of the three main civils asset sub-categories:
 - structures;
 - buildings; and
 - earthworks.

For each of the five workstreams set out above, we review in the following sections of this chapter the specific approach taken, the nature of data obtained and the analysis applied in relation to NR's overall CP5 efficiency projections across all civils assets.

6.6.2 Background

The first of five civils-related benchmarking workstreams that we review in this chapter is the comparative civils cost benchmarking workstream led by NR's Investment Projects (IP) division. This work began in March 2010, with a full report produced in January 2011 comparing costs.

NR's total expenditure for civils renewals in FY 2010/11 was as follows:

- For civils assets (including structure and earthworks) expenditure totalled £356 m – which represents 16% of the year's overall renewals expenditure (£2,234 m in total).
- For buildings assets (labelled within the regulatory accounts as "Operational property") expenditure totalled £272 m – which represents 12% of the total renewals expenditure.

For CP5, NR's IIP projections indicate that for renewals in both the buildings and civils categories, a combined volume and unit cost efficiency of 16% will be achieved over the five-year control period. This will consist of the following:

- 11% unit cost efficiency: NR indicates that this will be achieved through the same efficiency factors that are applicable to signalling efficiencies (see Section 6.3.1): "asset management process improvements; optimised whole-life renewals requirements; whole rail-system optimised operational requirements and asset type standardisation." NR states that "(t)he key

enablers to unlock these efficiencies are improved contractual relationships with our suppliers, locking down of workbanks as well as improved working windows through the optimisation of disruption and construction costs.”

- 5% scope efficiency: NR indicates that “(t)hese scope efficiencies will mainly be delivered through the B&C transformation programme already shared with ORR.”⁴⁰

For both the buildings and civils categories, NR projects a negative input price efficiency impact of -1% (in line with all other renewals categories).

The efficiency attributed to AIS system implementation for buildings is 1%, whilst no efficiency for civils is attributed to the AIS system.

6.6.3 Civils comparative cost benchmarking(IP team)

6.6.3.1 Approach

The purpose of the first phase of benchmarking work led by NR’s IP division was to establish the scale of the cost gap for specific rail-related civil engineering activities, and to identify the factors driving these cost gaps.

The work began in mid-2010 with a report completed in February 2011. NR indicated that the report represents the first known attempt to employ bottom-up benchmarking for real rail schemes.

This initial study resulted from NR’s collaborative working group, the “Civils Benchmarking Alliance” described by NR as “an informal group of companies including Network Rail, Birse, Bam, Vinci and Faithful+Gould”. The respective participants were asked to provide indicative cost data for projects undertaken within France, the Netherlands and Germany.

The first part of the review entailed a comparative analysis of factor price rates for labour, plant and materials. This helped NR to understand underlying cost differences.

NR then undertook a detailed project cost comparison exercise, whereby comparator organisations were asked to virtually bid for the following three specific civils infrastructure projects:

- Reconstruction of an underbridge.
- Refurbishment of a station footbridge and lifts.
- Stabilisation works on an embankment.

The comparator organisations were provided with an equivalent technical specification for each of the three projects, for which they provided costs set out on a bottom-up basis. NR utilised Eurostat Purchasing Power Parity figures updated on 1st October 2010 to establish the comparative price level index through which source cost data were harmonised.

⁴⁰ We reproduce the NR’s CP5 efficiency projections by asset category in **Error! Reference source not found. Error! Reference source not found.**

Using the data provided, NR analysed and collated a comparative breakdown of cost for the three schemes, with total costs for each sub-divided into one of the following five categories: labour, plant, materials, sub-contractor costs and on-costs (overhead).

NR then analysed key cost differences for each of the three schemes in detail as well as the sum of costs per category when added together for the three schemes. A range of factors driving differing cost levels are discussed and the report concludes that a number of areas of potential efficiency can be identified from the study. We discuss the interpretation and application of the results obtained in further detail below.

Faithful & Gould independently verified this benchmarking work.

6.6.3.2 Workstream data

Engagement with comparator organisations

Using the approach described in section 6.6.3, NR engaged with the following three contractors responsible for delivery of European infrastructure projects to obtain comparative cost data:

- Vinci: cost data based on projects delivered in France;
- BAM: cost data based on projects delivered in the Netherlands;
- Balfour Beatty: cost rate data based on projects delivered in Germany; and
- A UK civils contractor engaged in delivery of highways infrastructure.

NR selected the following three schemes for the detailed bottom-up cost comparison:

- Underbridge reconstruction: scope specification based on the “Victoria Road” project, the reconstruction of a four-track rail bridge over a two-lane road
- Station buildings renewal: scope specification based on the “Wrexham Access for All” project, comprising replacement of a station footbridge and installation of a new lift.
- Embankment stabilisation: scope specification based on the stabilisation of 2000 linear metres of embankment, toe loading the side slopes of the embankment.

NR considers the three projects to be “representative of typical B&C projects,” although the report does not explicitly state the motivation for these particular projects being selected.

Nature of information obtained

NR has presented its analysis and the results obtained in full in its February 2011 report.

For the initial analysis of comparative cost rates for labour, plant and materials, NR was able to obtain detailed comparative cost data from the benchmarking

contracts for France, Germany and the Netherlands, which comprised the following:

- Labour – hourly cost rates for 6 x typical project roles.
- Plant – hourly cost rates rates for 8 x types of machinery typically required for civils renewals works.
- Materials – unit prices for 7 x types of material typically required for civils renewals works.

For the detailed comparison of bottom-up cost elements for the three schemes, NR obtained comparator data from the UK civils contractor and the Netherlands, with France also providing data for the underbridge and embankment projects (but not the station project).

Review of quantitative data

Detailed quantified benchmarking data relating to three specific civils renewals projects (underbridge reconstruction, footbridge & lifts replacement, earthworks stabilisation) have been obtained from three European comparators, which include a breakdown of components of cost within the total amounts. NR has also obtained comparative cost rates for labour, plant and materials from four European contractors.

NR's benchmarking observations relate to the civils sub-categories; underbridge, footbridge and embankment renewals. Whilst in high-level terms these three asset sub-categories account for 53% of NR's total civils renewals expenditure, NR has not provided a detailed assessment of coverage that accounts for the level of representativeness of expenditure of the given project case study in relation to overall expenditure under the given sub-category. For example, whilst replacement of a metallic underbridge has been benchmarked, a significant proportion of expenditure under this sub-category (approximately 60%) relates to non-metallic structures, which were not benchmarked.

Our view is that the quality of data we have reviewed appears reasonable. The methodology and assumptions by which comparator figures were obtained, and the methodology by which comparator figures were normalized, are clearly laid out in the report provided. However, we consider that the robustness of data are limited by the small number of comparator organizations and projects compared. A greater number of activities need to be observed to increase robustness and ensure project data gained are not anomalous.

Although NR has indicated that initial outputs of the programme to date influenced the numbers presented in the IIP, we have not been able to verify how the outputs of this programme, in particular, were deployed in the development of those numbers.

6.6.3.3 Workstream outputs

Civils cost benchmarking report

NR's report presents comparative costs both for the three schemes individually, and the total sum of cost across all three schemes. Based on the combined results from the three schemes, the initial results suggest that NR is the most expensive of the four organisations within the dataset, with costs round 15 - 25% higher than

the comparators' costs – although the cost differentials vary across the three schemes.

NR's report contains some detailed interpretation and analysis of differences in cost elements. Differences in cost are discussed at some for each of the three project case studies and a number of specific factors that lead to comparatively high costs for NR in certain areas are discussed, as well as those factors for which NR's costs are comparatively lower.

The most significant factor identified as leading to higher relative costs for NR is the costs of gaining access to the rail infrastructure – i.e. possessions costs. The report discusses the way in which the UK possession regime not only gives rise to additional compensation costs but also that it requires work to be undertaken within a comparatively shorter timeframe that in itself drives up costs for performing the work. NR applies an adjustment to the original results of the three-scheme comparison, to illustrate the effects on comparators' costs, if they were subject to the same compensation structure. This results in significant alteration of the total cost figure (based on the sum of costs across all three schemes), whereby the UK is shown to be marginally less expensive than the Netherlands (by around 8%), and only slight more costly than France before normalisation for possession and disruption costs (although as stated above, the data for France does not include any cost for the station project). NR states later in the report that the adjustment of cost to normalise the impact of possessions costs in this ways shows NR prices to be lower than both the Netherlands and France, once data have been normalised.

Other significant drivers of comparatively high NR costs identified by the report include:

- Contracting strategy: overseas comparators have longer tender periods to develop more efficient / lower cost designs.
- Design life: this is identified as generally higher in the UK than elsewhere.
- Construction parameters: a number of technical parameters likely to drive comparatively high UK cost are identified relating principally to typically more onerous design standards for bridges.
- Health and safety requirements: the UK tends to have stricter / more onerous requirements.
- Impact of Schedule 8 performance penalty regime: NR has indicated the risk of penalties associated with train disruption (levied through the Schedule 8 performance regime) drives risk-averse behaviour amongst contractors that can lead to inefficient practices. NR has identified that measures taken by contractors to mitigate potential risk in this area – such as procuring additional contingent plant to ensure timely delivery of works – is a factor that drives up overall costs.
- Costs for materials: NR indicates this is typically higher in UK than elsewhere.

NR sets out at the end of its report a number of recommendations. As well as setting out the need to broaden benchmarking analysis itself, the recommendations also relate to a number of specific areas of through which NR may be able to identify or establish efficiency measures.

NR indicates that improved understanding of the trade-off between activity costs and compensation costs (payable through the Schedule 4 regime) is needed. Although Schedule 4 penalties account for £250m cost to NR annually, the report indicates that NR's "schedule 4 calculator is limited in its ability to accurately calculate the Schedule 4 costs". We understand that an internal project is ongoing within NR to address this issue.

Other recommended areas of focus include:

- Establishing the degree to which NR's standards are reasonable or overly prescriptive and "gold-plated" (i.e. over-specified in terms of technical characteristics).
- Assessing the increase of supplier / contractor tender timescales.
- Review of the cost and necessity of NR's Christmas / bank holiday working.
- Gaining a greater understanding of comparator organisations' overheads.

In summary the report approaches the comparative benchmarking from the perspective of gaining a full overview of cost differentials and associated drivers contributing to the cost gap. In this respect, the work differs from the stated object of other workstreams to identify efficiency opportunities rather than analyse the whole cost gap (see other sections of this report). We consider that this aligns the workstream more closely to the programme-level objective of closing the "efficiency gap" (see Section 5.2) compared to the other workstreams. The conclusions relate predominantly NR's access requirements and regime, which not only gives rise to compensation expenses, but also to the requirement for shorter-term, more costly and intensive work practices.

Whilst the report does not set out any specific efficiency initiatives, we understand that the findings have informed a pilot project recently undertaken by NR at Nottingham station, to test the implications in terms of cost of the "relaxation" of possession compensation rules in relation to a station renewals project.

CP5 quantified efficiency proposals

Buildings & civils is one of only two asset areas (alongside track) for which NR has provided a quantified breakdown of CP5 efficiencies, drawing on bottom-up benchmarking findings.

On 21st November 2011, NR presented a numerical breakdown of the projected IIP volume and unit cost efficiencies which generally apply to both buildings and civils assets; (scope efficiency and unit cost projections of 5% and 11% respectively are projected in the IIP for these two asset categories – see Section 5.6.2) .

NR set out in high-level terms the proposed measures from which the projected scope efficiency of 5% would be delivered. NR indicated that this efficiency is based on optimised asset management processes. NR described its transformation programme, whereby the planning of B&C workbanks is optimised, with revised asset policies and RAMPs (route asset management plans) supported by improved

systems, tools and processes. A further breakdown of this figure was not provided.

For unit cost efficiencies, NR indicated that the 11% efficiency was to be achieved by improved delivery processes. A number of factors contributing to lower cost delivery and unit costs were discussed, and a list of eight specific factors contributing to the overall 11% efficiency were presented, which we reproduce below.

CP5 plan	Maximum benefit	Applicability	Planned benefit
1. Reduce NR project management overheads	2.03%	100%	2.03%
2. Improve working windows	12.30%	10%	1.23%
3. Reduce contractor overheads	0.78%	50%	0.39%
4. Efficient tender designs	1.91%	75%	1.45%
5. Improving contractor cash flow	0.94%	100%	0.94%
6. Performance specifications	3.39%	75%	2.55%
7. Procurement package improvements	2.55%	75%	1.93%
8. Reduce contractual damage	0.30%	75%	0.22%
Total			11%

Table 9 - Breakdown of projected B&C CP5 efficiency (source: Network Rail)

Arup has not been provided with the calculations and assumptions upon which the estimated maximum benefit and percentage applicability have been based. We cannot, therefore, give a direct assessment of the application of bottom-up benchmarking outputs to the efficiency proposals presented.

6.6.3.4 Reporter opinion - civils comparative cost benchmarking (IP team)

Principal observations

The civils comparative cost benchmarking (led by the IP team) is one of the more advanced bottom-up benchmarking workstreams. Drawing on detailed comparative analysis utilising data relating to three case studies from four comparator organisations, we consider that NR has identified and appraised the impact of a range of salient cost and efficiency factors, that relate not only to civils activities, but also to NR's maintenance and renewals of other asset types.

The types of activities for which detailed benchmarking was undertaken - underbridge reconstruction, station buildings renewal and embankment stabilisation – represent significant proportions of buildings and civils expenditure.⁴¹

⁴¹ Figures from Statement 9 of the FY11/12 regulatory accounts indicate that underbridges accounted for £26m (around 35% of total civils renewal expenditure), with earthworks accounting

We regard the approach to be wide-ranging, with a range of factors associated with the relative cost levels analysed. Alongside a comparative assessment of direct factor costs across comparators relating to the works – labour, plant, materials – NR also discussed factors relating to policy, planning, contracting and delivery processes as well as requirements and constraints in relation to both the infrastructure itself, and rules and regulations relating to the works.

NR's normalisation adjustment of results in the report to exclude the impact of the UK possessions regime (Schedule 4 costs), whilst providing an insight into potential for longer possessions arrangements, should not lead to exclusion of the Schedule 4 regime itself as a potential factor driving efficiency. We consider the overhaul of such contractual structures for the purposes of improving efficiency and cost-effectiveness across industry interfaces to be a valid area of focus from an efficiency perspective, and consider that NR, as a major rail industry player, has as much influence as any other party to initiate change or improvement of such regimes to this end.

We consider the benchmarking methodology, based on costing-up of three infrastructure schemes, is in itself an innovative and insightful approach. However, we consider that both the number of comparators engaged through this approach, and the number of case studies should be expanded to strengthen the robustness of results and the level of insight gained into the respective cost factors. This should include consideration of the issues set out below.

Although we note that NR's analysis suggests that France is a high-performing benchmark for this asset group, it is not clear that the European railways on which the comparator projects for the study were based (France, Holland) are likely to provide full access to 'best practice' (see Table 3). We therefore believe that engagement with more comparator organisations through efficiency analyses such as LICB could significantly strengthen the robustness and quality of the benchmarking dataset.

Experience suggests that single-scheme analyses will have numerous project-specific anomalies and, therefore, random effects. We therefore also recommend that NR undertakes undertaking additional case studies for the given civils activity areas (underbridges / station buildings / embankments).

We support NR's engagement of third party (Faithful & Gould) to independently verify the work. However, we consider that the robustness of the results obtained may be further enhanced through sensitivity testing to model additional scenarios affecting NR's cost levels. This could include:

- substituting price rates / unit costs used for the three schemes with prices from other NR "more expensive" projects,
- testing the impact of cost or possession overruns, factor price inflation or changes in contractor arrangements)
- assessing costs for similar works on "more difficult" assets, e.g. operational constraints / restricted access, poorer asset condition, etc.).

for £15m (20% of total civils renewal expenditure). Works on station buildings accounted for the vast majority - £39m, or 90% - of NR's total expenditure on operational property renewals.

CP5 quantified efficiency proposals

Apart from track, B&C is the only asset area that has provided any further breakdown of the projected high-level efficiency numbers projected in the IIP.

We consider that the efficiency measures proposed, and the levels of efficiencies estimated appear plausible. However, we have not been provided with the detailed calculations and assumptions upon which the estimated maximum benefit and percentage applicability have been based, therefore it is not possible at this stage to assess the robustness of the projected efficiencies.

Recommendation: As with track, we consider that for B&C efficiency projections the treatment of risk is important; we would recommend that NR clarifies the way in which risk, if at all, has been factored into its projections, (e.g. whether this has been taken into account within NR’s “Applicability” estimations for the respective efficiency factors). For the sake of clarity, we would recommend that for presentation of efficiency projections, risk contingencies or uncertainty ranges are itemised separately.

Potential additional comparators

For the purposes of analysing relative efficiency for civils renewals activities, we recommend that NR should aim, indicatively, to complete a benchmarking dataset which entails analysis of the costs based on at least three examples of performance of a given activity type.

We have recommended that NR explores engagement with the following:

- IMs that have, through top-down analysis, been identified as efficient comparators (e.g. Italy, Sweden);
- Non-European rail organisations: e.g. New York MTA Metro North and Long Island Railroads, Railcorp New South Wales (Australia) and Amtrak (US – focussing in particular on the north-east corridor route which similar in nature to some of NR’s main routes).

Recommendation: We suggest widening the scope of analysis and extending the methodology to enrich the analysis. We recommend including three schemes for each sub-asset group, including three known-to-be-efficient railways. We suggest that NR could identify probable efficient comparators from the existing top-down benchmarking undertaken by both NR and ORR (e.g. Italy and Sweden – see **Table 6**). Finally, we recommend considering non-EU comparators.

An alternative, less resourceful approach to this work might be to follow the method undertaken by other asset groups of estimating the cost of completing works undertaken overseas as if they were undertaken in the UK.

6.6.4 Civils best practice benchmarking (IP team)

6.6.4.1 Approach

In early 2011, the IP initiated a second benchmarking project, relating to civils best practices. Arup has been provided with a report from March 2011 setting out

the scope, approach and initial findings from this workstream. We understand that this work continues. A more recent version of the report document has not yet been provided.

The Best Practice Benchmarking Review (IP2) sought to identify best practices for civils renewals from public and private sectors organisations based in the UK.

NR's approach was based on collection and analysis of qualitative data, with a focus on comparing asset management and project delivery processes with both rail and non-rail comparator organisations.

NR developed sets of descriptive survey questions, considering a range of factors relating to buildings and civils infrastructure delivery, asset management and policy.

NR surveyed nine comparator IMs, alongside itself, under a set of survey questions. The questions were structured under the following three sections:

- Project management: this included the following six sub-categories:
 - Safety (relating to safety monitoring practices.)
 - Safety initiatives (specific measures and workforce initiatives to improve safety).
 - Construction design and management (roles of IM and its contractors in relation to construction and design).
 - Design (degree to which design undertaken in-house).
 - Specifications (degree to which contractors operate under performance and/or functional specifications).
 - Workbank budget (length of time (years) that workbank and budget defined by the IM).
- Commercial management: this included the following two sub-categories:
 - Unit costs (whether the IM collates unit costs).
 - Procurement method (how far the IM utilises alliances, framework agreements or competitive tendering).
- Asset management: this included the following four sub-categories:
 - Asset condition (method by which the IM scores asset condition).
 - Examinations (comparative frequency of examinations).
 - Assessment (degree to which IM performs a reassessment of structural strength of its structures assets, and frequency of assessment (if applicable)).
 - Policy & prioritisation (general focus and whole-life-costing approach taken by IM to managing asset condition and expenditure).

NR analysed the results of the survey questions to identify key themes and factors relevant to NR's efficiency. We review the data obtained and its interpretation and analysis in further detail below.

6.6.4.2 Workstream data

Engagement with comparator organisations

Following internal consultations with each internal head of asset management (earthworks, structures and buildings), NR developed four general criteria, on the basis of which comparator organisations were selected:

- “they operate in a very similar economic/political/regulatory environment to [NR]”;
- “[NR] wanted to see how [its] approach differed from private sector equivalents”;
- “[NR] believed they have a similar asset base;
- “[NR] was told [the comparator] was more efficient.”

On this basis, NR contacted thirty organisations, of which the following nine agreed to proceed with the NR benchmarking exercise:

- 4 x private companies: Amey (Highways Agency), Thames Water, National Grid and ASDA.
- 5 x public organisations: Hampshire County Council, British Waterways, Defence Estates, London Underground Limited and Staffordshire County Council.

Nature of data obtained

Our review of data obtained is based on the contents of the initial draft report of March 2010. We acknowledge work is ongoing and note that NR has indicated a greater degree of detail and related analysis is envisaged in updated versions of the report.

In its initial report NR was able to produce descriptive answers summarising in qualitative terms the nature of the comparators’ asset management approach relative to the questions raised. NR has then used the information to undertake a qualitative analysis of factors it considers relevant for efficiency in civils asset management, policy and project management.

The efficiency factors related to procurement contracting differences in the public and private sectors appear to be of particular importance. NR concludes from the information obtained that the private sector companies are more “hands off,” allowing contractors and designers to fulfil their roles without interference.

Whilst NR’s report provides a number of numerical details about its own operations, no equivalent data are evident from comparators to which NR can compare, and the focus remains high-level.

6.6.4.3 Workstream outputs

NR has utilised the results of its analysis to present a set of 12 key efficiency “questions”, representing “challenges to NR’s organizational potential”. These comprise the following:

- 4 x questions relating to project management. For example, “What are the blockers to having a five year rolling workbank?”
- 2 x questions relating to commercial / procurement arrangements. For example, “How can NR use frameworks to our advantage?”
- 6 x questions relating to asset management. For example, “Could we stop regular inspections of all or parts of buildings and rely on tenants to advise if there is a fault?”

In certain places within the report, NR suggests it is proposing to make changes based on the findings gained, e.g. changing contracting / procurement arrangements to a more “hands off” approach.

However, the version of the report provided (March 2011) leaves many of the questions open and it is not yet clear from this document how NR proposes to apply the findings in concrete terms to CP5 expenditure and efficiency proposals.

NR has suggested that the workstream data could form the basis of a gap analysis, although we understand such work has not yet been undertaken.

We look forward to working with NR in coming weeks to gain an understanding of how this workstream has progressed.

6.6.4.4 Reporter opinion –civils best practice benchmarking (IP team)

Principal observations

NR’s civils best practice workstream is based around a qualitative benchmarking data which relate to project management, asset management and commercial factors. We consider that this approach should provide relevant and meaningful insights into NR’s efficiency and effectiveness in these areas. As with the civils cost benchmarking workstream, we consider that the best practice workstream also involves themes – such as contracting strategy and asset management – that are relevant to asset areas other than civils, and which may contribute to a wider understanding of the constraints and structural factors affecting NR’s costs more generally.

We consider that the nine UK comparators represent a good selection of in terms of breadth, scope of activities and the nature of their assets.

However, as indicated above, we have not gained any visibility of how this workstream has progressed beyond the draft report, dated March 2011, that NR has provided for review. Although twelve specific “efficiency questions” relating to potential measures to improve efficiency are identified, it is not clear how NR proposes to respond to them, nor is there any quantified assessment or indication as to how such factors may impact NR’s expenditure or efficiency levels.

We recommend that NR provides update and clarification with regard to these issues.

Potential additional comparators

We recommend that NR should aim to extend the scope of engagement with comparators for this workstream. It is likely that this can be done most effectively by engaging with organisations involved in other civils workstreams (such as those mentioned for the civils cost workstream – see Section 6.6.3.4).

6.6.5 Structures benchmarking (AM team)

6.6.5.1 Workstream approach

The benchmarking workstream focusing specifically on benchmarking of structures is being led by NR's Asset Management division (headed up by Mark Evans). The aim is to engage both rail and non-rail comparators in the UK and elsewhere internationally, with a particular focus on asset management, planning and policy, thereby complementing the IP-led workstreams that focus more on asset delivery.

This workstream was initiated in September 2011 and is consequently one of the less advanced workstreams. Our review focuses on the proposed approach set out to us in recent meetings. We recognise that this is likely to evolve and develop in further detail as the workstream progresses.

NR is planning to develop an approach based on an interview matrix. This will guide the interview process which NR expects will comprise two key meetings (an initial interview and a follow-on session) and may be accompanied by site visits to review asset practices / application of policies first hand. Additionally, NR hopes to speak to contractors and consultants to gain a more comprehensive understanding of the data it collects.

NR's planned approach includes examination of the following six core themes:

- condition/defect assessment;
- lifecycle costs;
- asset fatigue;
- route criticality;
- technical specifications – in particular, the decisions around the use of steel or concrete materials for bridge construction; and
- contracting (i.e. in-house v. tendering out).

6.6.5.2 Workstream data

Engagement with comparator organisations

The first organisation that the structures workstream is planning to benchmark with is London Underground, with which a data sharing agreement is in place. We understand that this precludes the sharing of cost-related data, in order that the companies avoid any risk of contravening anti-cartel legislation.

NR also indicated it has initiated contact with both rail and non-rail comparator organisations:

- 4 x European rail operators (Sweden, Belgium, France and Italy)
- 9 x UK-based non-rail organisations (Hampshire and Sussex County Councils, British Waterways, Defence Estates, Highways Agency, Amey, Thames Water, Asda)

We understand that the level of progress in terms of meetings held varies and that efforts are ongoing to arrange meetings etc. We note that Arup/Imperial have also provided NR contact details for individuals within BAA, Nexus (operator of the Tyne and Wear metro) and RailCorp (Australian commuter and light rail operator). A number of organisations listed are common to this and other workstreams, including the best-practice workstream being led by IP (see Section 6.6.4.2), and the buildings and earthworks benchmarking workstreams (see sections 6.6.6.2 and 6.6.7.2). NR has indicated that contacts being made are being coordinated across the different workstreams

Nature of data obtained

Given that this workstream is still at an early stage, we understand that no detailed benchmarking data have yet been obtained from the comparator organisations. We expect to be able to review data obtained in later stages of this assignment, once the workstream has sufficiently progressed.

Review of quantitative data

NR indicated that it was expecting to obtain initial comparator data by the end of 2011, as reflected in the “Analysis and Draft Reports” workstreams set out in the programme plan during this period. However, no quantitative data have been provided to date.

NR’s programme plan indicates that it is proposing to develop unit costs to feed into the SBP submission during early-mid 2012.

6.6.5.3 Workstream outputs

We summarise in Section 6.6.3.3, above, the quantified breakdown of projected CP5 efficiencies based on bottom-up benchmarking findings that NR presented to us on 21st November 2011. The projected efficiencies apply to all categories of buildings and civils asset – including structures.

Alongside this, NR has stated that it plans to identify opportunities in the organizational set-up, embed these opportunities in its own organisation and realize efficiency gains required prior to CP5. At the time of writing, no data collection or analysis has yet been provided, but we expect to be able to review workstream outputs in later stages of this assignment.

We understand that a report has been completed by NR with regard to the costs of replacement of metallic structures, whereby NR’s cost data are compared with France. A copy of this report has not yet been provided for our review.

6.6.5.4 Reporter opinion – structures benchmarking (AM team)

Principal observations

Structures benchmarking is still at a comparatively early stage, and is therefore one of the less advanced workstreams. Our opinion is limited by the fact that no benchmarking data collection or analysis has yet been provided.

We consider that the six themes identified as areas of focus (condition/defect assessment; lifecycle costs; asset fatigue; route criticality; technical specifications; and contracting arrangements) are likely to be salient factors from the perspective of asset management efficiency; however the nature of data to be gathered and how these are to be factored into efficiency proposals is presently unclear.

Given that NR is only presently at the stage of establishing initial contact with comparators, we consider there is uncertainty at this stage about the extent of benchmarking analysis, and the findings that will be achievable within timescales of the PR13 determination.

Potential additional comparators

Please see our comments in Section 6.6.3.4.

6.6.6 Buildings benchmarking (AM team)

6.6.6.1 Workstream approach

The benchmarking workstream focussing on benchmarking of buildings is being led within NR by Mark Evans, supported by representatives from the asset group.

One of the less advanced workstreams, the buildings benchmarking activities began in late 2009. NR has indicated activity has been sporadic rather than continuous, however we understand that a dedicated team is now assigned to this workstream going forward.

To date the focus of activity has desktop research and reviews of information in the public domain and within NR's organisation, as well as previous studies. As well as seeking to analyse performance metrics and identify areas of best practice for buildings assets, NR proposes to use this process to establish topics for exploration and to inform the approach taken to the direct benchmarking comparator data. NR has engaged the consultancy company Opus to help support its analysis.

NR is aiming to focus its approach around asset management policy, with the aim of "establish[ing] insights into qualitative efficiency factors." NR has indicated that it is seeking to gain sufficient information to support business cases that will enable it to implement measures to achieve both efficiency requirements during CP5, and efficiencies within a longer-term timeframe.

Through its desktop analysis, NR has identified a number of potential areas of focus which we understand will inform the benchmarking approach going forward. These include:

- Contracting / procurement arrangements, including optimisation of specifications, responsibilities for planning and delivery and risk allocation;
- Asset management processes including asset knowledge, lifecycle planning and interventions, performance requirements and risk.

NR indicated that it is seeking to focus in particular on scope efficiencies – given that management of activity volumes and intervention policies are seen to be more relevant in terms of efficiency potential than unit cost-based efficiency.

To this end, NR has developed a qualitative benchmarking questionnaire framework setting out asset management and policy areas through which NR is seeking to identify efficiencies. NR indicated that it has tested the questionnaire internally, but we understand that the questionnaire has not yet been utilised through contacts with comparator organisations.

We propose to review the approach further once we receive the questionnaire documentation and gain an understanding of the precise approach for inquiring in different areas and the scope of the analysis.

6.6.6.2 Workstream data

Engagement with comparator organisations

The buildings benchmarking team is proposing to establish benchmarking contacts with a range of organisations. This includes organisations with whom NR has already established contact through other workstreams including:

- European rail operators including France, Italy, Belgium and Holland.
- Non-heavy rail organisations including London Underground, BAM Nuttall.

NR has also established contact with University of British Columbia (Canada) to share information and gain an insight into the university's research relating to the residual life of buildings.

We understand that NR has also met with UK Train Operating Companies, which undertake day-to-day operation of rail stations to discuss potential efficiency improvements relative to performance and output requirements in station provision. At this stage, we have not been provided with any documentation or details of these discussions.

Arup/Imperial discussed with NR the potential utilisation of further data sources including:

- Information from private concessionaires who have acquired and operated station facilities, e.g. Sacyr (Spain), that may be publically available.
- Analyses / reports containing information from the London Underground PPPs (Metronet and Tube Lines) including cost-related reports for the PPP Arbiter (OPPPA).

Nature of data obtained

As with the structures benchmarking, we understand that for buildings no detailed benchmarking data have yet been obtained from the comparator organisations but we expect to be able to review data obtained in later stages of this assignment.

Review of quantitative data

NR indicated that it was expecting to obtain initial comparator data by the end of 2011, as reflected in the “Analysis and Draft Reports” workstreams set out in the programme plan during this period. However, no quantitative data have been provided to date. NR’s programme plan indicates that it is proposing to develop unit costs to feed into the SBP submission during early-mid 2012.

6.6.6.3 Workstream outputs

We summarise in Section 6.6.3.3 above, the quantified breakdown of projected CP5 efficiencies based on bottom-up benchmarking findings that NR presented to us on 21st November 2011. The projected efficiencies apply to all categories of buildings and civils asset – including buildings.

Apart from this, as with structures, at the time of writing, no data collection or analysis for buildings has yet been provided, but we expect to be able to review workstream outputs in later stages of this assignment.

6.6.6.4 Reporter opinion – buildings benchmarking (AM team)

Principal observations

In common with structures benchmarking, the buildings benchmarking workstream is also at a comparatively early stage, and no benchmarking data collection or analysis has yet been provided for review.

We consider that, once again, NR has highlighted two key areas of focus – contracting / procurement and asset management strategy – that are, at a high-level for this workstream, likely to be key factors influencing relative efficiency levels.

With regard to desktop-based analysis of publicly available data sources, we consider that, given the nature of buildings and the fact that a wide range of infrastructure exist in which operational buildings are managed, there may be a greater range of potential cost- and efficiency-related information available, compared to other more rail-specific asset types.

Nevertheless, we regard such analysis is likely to be of limited value, compared to direct engagement with comparator organisations – in particular the European rail operators identified. We support engagement with comparators utilising structured and detailed sets of questions – although as indicated above, the questionnaire has not been provided for our review.

Given that NR, as with other workstreams, is only presently at the stage of establishing initial contact with comparators, we consider there is uncertainty at this stage about the extent of benchmarking analysis, and the findings that will be achievable within timescales of the PR13 determination.

Potential additional comparators

Please see our comments in Section 6.6.3.4.

Recommendation: we understand that NR operated stations are a major spend and include relatively homogenous costly assets such as canopies and escalators that we recommend could be benchmarked with peer IMs. Though less significant in its contribution to total costs, operating expenditures (e.g. energy use) might represent a ‘small win’ for NR’s AM—buildings benchmarking practice.

6.6.7 Earthworks benchmarking (AM team)

6.6.7.1 Workstream approach

The benchmarking workstream focussing on benchmarking of earthworks is being led within NR by Mark Evans, supported by representatives from the asset group

The earthworks workstream is one of the newest benchmarking workstreams, with work commencing in March 2011. However, the earthworks benchmarking can be considered to be more advanced than the structures and buildings workstreams.

We understand that the earthworks benchmarking workstream proposes to follow a similar approach to the AM-led structures workstream (see Section 6.6.6), with the utilization of an interview matrix and to question comparators in the UK and Europe, including both rail and road comparators.

The questionnaire will be used to gather qualitative analytical data, which we understand will then be analysed and applied to NR’s own efficiency proposals. Information to be collected includes the following:

- Comparator earthworks (number and type of slopes, budget / expenditure levels, etc).
- Condition management (KPIs, asset condition measures, AM databases, prioritisation etc.)
- Inspection regime.
- Innovations in earthworks monitoring and management.

Using the information obtained to date, NR has sought to identify specific factors that may support more efficiency asset management of NR’s earthworks assets. We review these in further detail below.

6.6.7.2 Workstream data

Engagement with comparator organisations

NR has undertaken benchmarking meetings with the following three organisations:

- London Underground

- Highways Agency
- SNCF.

NR is proposing to establish contact with other organisations responsible for management of earthworks assets, and has identified the following as organisations with whom it proposes to establish contact:

- Other rail organisations: Kiwi Rail (New Zealand), Dutch, Swiss and German railways
- Non-rail organisations: British Waterways, Yorkshire Water, Environment Agency,

Nature of data obtained

From the three meetings undertaken to date, NR has obtained qualitative information relating to the comparator organisation's management of earthworks assets.

From its discussions with SNCF, NR has found that the number of earthworks failures, geology and asset condition on the French network are "very similar" to NR. However, NR identified that SNCF does not periodically examine the works, instead relying on maintenance teams to monitor condition and identify issues relating to condition when they arise. SNCF has 115km of remotely monitored embankments, for which a signalling system stops trains if a failure occurs.

For the Highways Agency, NR reviewed the organisation's deployment of a geotechnical data management system utilised to capture earthworks data and inform activities on the asset.

We understand that information obtained from London Underground (LU) related to examination process, number of failures, amount of assets, and overall costs although we understand the report documenting these findings is not yet complete.

Review of quantitative data

NR indicated that it was expecting to obtain initial comparator data by the end of 2011, as reflected in the "Analysis and Draft Reports" workstreams set out in the programme plan during this period. However, no quantitative data have been provided to date.

NR's programme plan indicates that it is proposing to develop unit costs to feed into the SBP submission during early-mid 2012.

6.6.7.3 Workstream outputs

NR has drawn upon the findings of its qualitative benchmarking analysis to focus primarily on potential efficiencies in relation to the monitoring, inspection and risk mitigation relating to earthworks; indeed, such factors are themselves the principal cost elements relating to earthworks assets.

We understand NR is evaluating the utilisation of remote monitoring systems similar to those used in France for rural routes, e.g. in Scotland. We look forward

to reviewing the evaluation and projected efficiency impact of this initiative further with NR as the proposals progress.

We summarise in Section 6.6.3.3 above, the quantified breakdown of projected CP5 efficiencies based on bottom-up benchmarking findings that NR presented to us on 21st November 2011. The projected efficiencies apply to all categories of buildings and civils asset – including earthworks.

Apart from this, no detailed indication of the impact of the above factors on CP5 expenditure and efficiency of the measures in quantitative terms has yet been provided.

6.6.7.4 Reporter opinion – earthworks benchmarking (AM team)

Principal observations

Although earthworks benchmarking is, like structures and buildings benchmarking, still at a comparatively early stage, we consider that from a best-practice perspective NR has gained useful insights into significant factors affecting comparative efficiency levels, most notably in relation to monitoring and inspection practices.

Despite the fact that sharing of data to date has been limited to just two comparators organisations - SNCF and the Highways Agency - and that only a handful of meetings have taken place, we understand that NR has already drawn upon this analysis to develop concrete proposals for implementation of remote earthworks monitoring.

However, the focus to date has been on best-practice, and, similar to other workstreams, the benchmarking data obtained and analysis undertaken does not appear to relate directly to overall expenditure and efficiency proposals.

Meaningful bottom-up benchmarking analysis of a range of factors influencing relative efficiency levels would require engagement with a number of comparators. Although several potential comparators have been identified for the earthworks benchmarking, it is once again uncertain how far meaningful bottom-up benchmarking analysis can be undertaken, and the findings that will be achievable within timescales of the PR13 determination.

Potential additional comparators

Please see our comments in Section 6.6.3.4.

6.7 Maintenance benchmarking

6.7.1 Background

The maintenance benchmarking workstream is being delivered by NR's maintenance division, headed up by David Wynne.

NR's total expenditure for maintenance in FY 2010/11 totalled £1.07 bn – which represents 17% of the company's total annual expenditure (excluding financing costs and tax) of £6.15bn.

NR's CP5 projections in the IIP indicate that for maintenance a total efficiency of 14%⁴² will be achieved over the five-year control period, consisting of:

- 4% scope efficiency.
- 9% unit cost efficiency.
- 3% efficiencies to be achieved through NR's "Asset Information Strategy" (ORBIS).
- -1% (i.e. 1% inefficiency which slightly offsets the efficiencies listed above) relating to increased input prices.⁴³

NR proposes the combined scope and unit cost efficiency (13% in total) can be achieved on the basis of a number of factors including:

"... reducing reliance on directly employed staff for lesser skilled maintenance activities, increasing the scope of the projected roll-out of video inspection techniques and reviewing the deployment of rapid response personnel to align with criticality of individual routes.....

"Achieving these savings is also dependent on further implementation of risk based maintenance and successful management of the access regime at critical locations."

For the 3% efficiency attributed to the AIS system NR has stated that this will be achieved through improved asset information, enabling a reduction in maintenance rework and inspection requirements.

6.7.2 Approach

NR's maintenance benchmarking activities comprise two distinct but interrelated workstreams:

- Internal cost benchmarking between individual Maintenance Delivery Units (MDUs) within NR; and

⁴² We reproduce the NR's CP5 efficiency projections by asset category in **Error! Reference source not found.**

⁴³ We note that the percentage breakdown does not exactly add-up to the same total figure; this is likely to be due to rounding of the numbers by NR, although Arup has not been provided with the original calculation spreadsheet / formulae.

- External activity-based comparative benchmarking with other rail organisations.

Approach: internal cost benchmarking

NR manages maintenance of the rail infrastructure in-house, with 40 semi-autonomous MDUs responsible for delivery of maintenance across the national network. The recording of cost by MDUs – including the widespread utilisation of unit costs under the “MUC”⁴⁴ framework – is undertaken on a highly uniform basis, enabling NR to record and collect large amounts of data.

NR’s internal cost benchmarking is based on statistical analysis of costs, focusing primarily on the total cost per MDU and the possible factors that lead to differences in relative cost levels. The analysis is not based on a conventional cost function approach (ie costs as a function of prices and output levels), but rather on a hypothesis testing approach which regresses total cost on covariates representing some characteristics of the MDUs. These include: track miles, levels crossings, s&c unit density, traffic volumes (train km/track km), and electrified route. Other ‘exogenous factors’ are also specified within the model which are determined through surveys of the Delivery Units.

The model is estimated by Ordinary Least Squares (OLS) in log-linear form. The objective of the analysis is to estimate the effect of key DMU characteristics on costs and to use this information to form an efficiency ranking of MDUs which helps identify potential for cost savings. The selection of covariates for the model appear to be based on a type of ‘stepwise’ approach in which NR originally identified more than 15 factors potentially associated with total MDU cost (i.e. the dependent variables); and then retained or removed variables according to some goodness of fit criteria.

Arup has reviewed documents describing the regression analysis and has met with members of NR’s team. NR has explained that a log-log specification has been used, and that the independent variables enter (log) linearly, not in polynomial form – although documents provided do not provide detailed information about the OLS specification.

We explore the factors further in the next section.

Approach: external comparative benchmarking

NR has adopted an Activity Based Benchmarking (ABB) approach to compare its maintenance activities to other European railways. This is considered more robust than a direct comparison of activity costs given that comparator organisations are likely to account for various cost elements (e.g. labour costs, treatment of overheads, penalties payable for disruption, etc.) differently to NR.

The ABB approach entails the following three key elements:

- Comparison of policies and standards: NR has undertaken an office-based comparison to identify and assess factors likely to influence maintenance activities and associated cost levels.

⁴⁴ MUC – Maintenance Unit Cost: a standardised unit of measure adopted for approximately 60 repeatable activities, enabling volumes to be reported for the given activity type and a unit cost to then be derived. Approximately two thirds of total maintenance expenditure is captured under the MUC framework.

- Site-based observation of maintenance activities: NR has sent teams to observe first hand comparator railways carrying out defined maintenance activities. Details recorded include the level of resources used (e.g. labour, plant and materials), timings, associated management and delivery processes (e.g. planning, possession arrangements). NR has sought to observe the given activity type repeatedly over a five day period to ensure that it observes ‘typical’ methods.
- Calculation of comparator cost levels using NR unit rates; this has involved an estimation of the comparator’s costs for the given activity, which is done by multiplying the resourcing levels (labour, plant & materials) by NR unit rates. The comparator’s estimated cost is then compared to the costs incurred by NR for the same activity.

In addition to the ABB-based activities described above, NR is also proposing to contact three UK-based non-rail utilities companies – National Grid, Scotia and Enterprise – to compare approaches to the contracting of maintenance work, and the utilisation of multi-skilled staff.

Combining the outputs from the various bottom-up analyses, NR has begun to assess in qualitative terms the factors that influence the relative cost levels, which we explore in further detail below. NR has not however provided any detailed information or documentation setting out how it proposes to relate the outputs from the bottom-up benchmarking directly relate to the CP5 maintenance expenditure and efficiency projections.

6.7.3 Workstream data

Internal cost benchmarking

The internal benchmarking analysis gauges the relative performance of each MDU within the 40-strong sample. The sample dataset affords a wealth of cost factor data: NR performs over 80,000 maintenance jobs each week. NR currently uses an OLS regression approach to appropriate budgets across cost centres.

NR considers it has identified five key structural factors that relate to relative levels of operational expenditure per MDU. These are:

- track length (track miles per route miles);
- number of level crossings;
- S&C unit density;
- traffic volume (train Km / track Km); and
- absolute track geometry (ATG).

External comparative benchmarking

Engagement with comparator organisations

For maintenance benchmarking, as with track, NR identified “efficient” European rail comparators through analysis of costs captured through the LICB dataset. Although the LICB maintenance data were limited in detail, NR indicated that it was able to identify France and Sweden as comparators that are generally efficient in maintenance of all asset types, whilst Italy was identified as particularly

efficient in track maintenance. NR also indicated it plans to arrange visits with Netherlands and Norway, but that it has not been successful in trying to arrange visits to Switzerland and Denmark.

NR has provided the results obtained from the following two initial visits:

- France: review of rail changing.
- Sweden: review of sleeper changing.
- France: site visits during October 2011 (note that details of the findings from these site visits have not been provided for review).

NR has indicated that it proposes to undertake the further site visits, as follows:

- week of site visits to Sweden – spring 2012
- week of site visits to Italy –spring 2012

In the longer term NR also plans to facilitate sharing and exchange of data on an automated basis, enabling greater levels of data to be shared remotely.

NR also indicated it was proposing to initiate contact with three non-rail organisations - National Grid, Scotia and Enterprise – during October 2011.

Nature of information obtained

NR has provided reports from the two site visits, containing detailed observations of the activities undertaken, together with high-level supporting information relating to the comparator organisation's maintenance organisation and rail network as well as a brief description of NR's own processes for undertaking the same activity.

Arup has also been provided with the calculation spreadsheets utilised to compare costs for the two activities reviewed.

The tables are used to capture quantified observations obtained on-site for each individual cost element / factor that feeds into the total activity. The level of resource associated is multiplied by the relevant NR cost rate for each cost element; this produces a relative cost rate for the comparator's performance of the activity. NR's own cost rate, built-up in exactly the same format, is provided in the adjacent table, enabling comparisons for each element of cost within the total to be compared.

NR has indicated that this comparative analysis has informed its assessment of maintenance efficiency opportunities – which we discuss below. However no further detail of the application of cost analysis to overall maintenance expenditure and CP5 efficiency proposals has been provided.

Review of quantitative data

In the course of our work we reviewed quantitative data relating to two maintenance case studies. We understand from NR that quantitative data are also available in relation to a further four case studies.

The quality of data reviewed in relation to the case studies described above appears reasonable. The methodology and assumptions by which comparator

figures were obtained, and the methodology by which comparator figures were normalized, is clearly laid out in the report provided. However, we consider that the robustness of data are limited by the small number of comparator organizations and projects compared. A greater number of activities need to be observed to increase robustness and ensure project data gained are not anomalous.

Although NR has indicated that initial outputs of the programme to date influenced the numbers presented in the IIP, we have not been able to verify how the outputs of this programme, in particular, were deployed in the development of those numbers.

6.7.4 Workstream outputs

Internal cost benchmarking

NR has sought to utilise the results of the internal maintenance benchmarking to improve efficiency and encourage best practice across the maintenance organisation. MDUs are grouped according to key factors identified in the regression analysis. Measures to improve internal efficiency include:

- League tables: NR considers league tables incentivise MDUs to perform more efficiently relative to their peers and help to identify good and bad practices within the organisation.
- Extra management support and action plans provided for lowest-performing / highest-cost MDUs within a group, provision of support from top-performing MDUs in relation to critical delivery practices such as scheduling and staff rosters.
- Sharing of best-practice through internal communications, conferences, etc.

External comparative benchmarking

Alongside the two initial cost comparisons provided for rail changing and sleeper changing, (each of which compared NR cost with a single comparator, as documented above) we understand that NR has undertaken a number of additional comparative cost analyses; however, details of the results obtained and analysis undertaken have not been provided.

Although the size and scope of the benchmarking dataset is limited to observations from four site visits, NR has drawn upon the comparatively detailed level of information from the site reports and detailed cost comparisons to identify a number of key factors influencing relative cost and efficiency levels. NR indicated it is presently exploring a number of efficiency opportunities on this basis, including:

- Inspection and servicing frequencies: NR has found it services its assets more than comparator organisations, and is evaluating a lower-cost, more risk-based approach to maintenance activities, e.g. greater use of remote monitoring technology and track recording trains.
- Possession and safe system of work management: NR is exploring methods to make the processes for taking possession and ensuring worksite safety more efficient.

- Asset management information: improved knowledge of asset condition and development of more efficient maintenance activities on this basis.
- Other factors relating to efficient organisation and structure of the maintenance functions such devolved decision-making, and improvements in “rapid response” processes to deal with unplanned incidents more efficiently.

Although the above factors have been presented by in high-level qualitative terms, NR has not yet provided details of concrete initiatives or proposals that it plans to implement on that basis. No indication of the impact of the above factors on CP5 expenditure and efficiency of the measures in quantitative terms is yet provided.

6.7.5 Reporter opinion

Principal observations

External maintenance benchmarking

We consider that the maintenance external benchmarking is one of the most advanced workstreams in terms of quantified, comparative analysis undertaken, although the degree to which NR has developed concrete efficiency proposals in relation to maintenance activities is limited.

The methodology to estimate comparative costs for specific activities by drawing upon quantified observations and multiplying these by internal unit costs appears reasonable. In the absence of available comparator cost data, an activities-based comparative methodology can provide useful insights into the relative levels of resource intensity and productivity associated with particular activities.

The observations-based approach is by its very nature also likely to provide additional qualitative insights into the nature of maintenance processes and activities on the ground, alongside the quantified observations being recorded. As well as the quantified comparison of on-site resourcing, we consider that NR has retained focus on a reasonably wide range of other factors relevant to efficiency such as planning and inspection processes that drive maintenance volumes, staff skills and competencies, and factors influencing cost rates on-site such as contracting strategy, possessions and safety arrangements.

NR’s programme involves engagement with comparators including Sweden and Italy, which are considered to be efficient in comparative terms. Although we have not viewed in detail the process by which these rail organisations were selected, we consider that three comparators may be a reasonable number assuming these organisations are in comparative terms more efficient in delivering rail maintenance.

However, we have found the significance of the external maintenance benchmarking in terms of the number of sets of observations gained is limited; and detailed results and comparative cost analysis from only two visits were provided for review – although, as indicated above, we understand that further visits were planned for recent months.

Although we acknowledge that undertaking activity-based benchmarking depends on site-based observations is a resource- and time-intensive process, we would

consider that extension of the dataset to encompass both a greater number of observations (i.e. completed jobs for the given activity type). This would strengthen both the robustness of data obtained, in terms of reducing the incidence of outliers, and increase its explanatory power through encompassing a greater number of quantitative and associated qualitative observations.

Internal maintenance benchmarking

NR's external maintenance benchmarking activities are also complemented by its internal benchmarking analysis. This focuses principally on relative levels of total expenditure per MDU, and the causal and structural factors influencing differences in cost which are assessed using regression analysis.

We consider that the regression approach provides a useful tool to summarise the data and generate some empirical evidence on the nature of variation in costs at the MDU level. Over time, as more data becomes available, it will be possible to develop the model further and conduct some more in-depth analysis.

The model results are not presented in detail. However, the plots of actual versus predicted values do appear to show that the model does have some explanatory power. The results are used to identify key drivers of costs and to gauge the performance of each unit in the sample.

It is important to stress that the model estimated by NR is not a conventional cost efficiency analysis which, drawing on economic theory, considers costs in relation to prices and the volume of output. It is instead an approach on an intuitive understanding of what the key cost drivers may be.

The MUC framework provides the opportunity for around 70% of internal costs to be captured on a unitized, with more than 80,000+ datapoints captured every week. We consider this represents major scope for further internal benchmarking.

Development of efficient expenditure proposals

Although the maintenance benchmarking activities are amongst the most advanced of the benchmarking workstreams in terms of detailed bottom-up data obtained, NR has not yet provided details of concrete initiatives or proposals that it plans to develop or implement that draw upon the maintenance benchmarking analysis, other than the high-level themes identified, that relate to factors that influence efficiency. This is reflected in the IIP, whereby reference is made in qualitative terms to the range of relevant drivers (multi-skilled staff, access arrangements, inspection and monitoring etc. – see section 6.7.1 above) but without explicit, quantified linkage of the respective factors to actual efficiency numbers.

We consider that NR should provide clarification as to how it proposes to develop and substantiate in quantified terms its maintenance efficiency proposals, drawing on comparative benchmarking analysis and the insights it provides.

Potential additional comparators

For the purposes of analysing relative efficiency for maintenance activities utilising the activity based benchmarking approach developed for the benchmarking, we recommend that NR should aim, indicatively, to complete a benchmarking dataset which entails:

- engagement with approximately five rail organisations in total (as long as at least three are classed as efficient), plus at least two other comparable IMs.
- analysis of the costs associated with approximately five activities per comparator.

We have recommended that NR explores engagement with the following as potential additional comparators:

- Non-European rail infrastructure managers: e.g. Hong-Kong MTR, New York City Transit, NSW Railcorp (Australia), US railroads including MTA New York Long Island and Metro North Railroads.
- Other rail IMs that are known to have developed advanced approaches to risk-based maintenance.

Recommendations

External benchmarking

- To deepen its analysis, we recommend NR uses time series data from other railways to identify where these organisations have driven down maintenance costs. Doing so will allow NR to explore in detail whether these cost savings arise from efficiency, deferred maintenance or the impact of renewals or enhancement spending.

Internal benchmarking

- For present and future internal benchmarking analysis of efficiency across MDUs, we recommend that NR considers the following statistical techniques:
 - Semi/non-parametric estimation.
 - Confounding / omitted variable bias (OVB).
 - Normalised dependent variable.
 - Cost function with price data.
 - Use of temporal as well as cross-sectional analysis.
 - Analysis by category of spending.
 - ANOVA for grouping analysis.
 - The internal benchmarking is used to allocate a fixed budget.
- We present a full definition of each of the above statistical techniques in **Error! Reference source not found.**

7 Recommendations

We set out in *Table 3* below our level recommendations in relation to this assignment. We draw together recommendations from both within the Executive Summary, and throughout the remainder of the report. We categorise our recommendations under the following six sub-headings:

- Alignment of objectives
- Engagement with comparators
- Benchmarking dataset
- Benchmarking data analysis
- Workstream-specific recommendations
- Alternative benchmarking approaches for analysis by ORR

Ref.	Recommendation to Network Rail
<i>Alignment of objectives</i>	
2011.BUB.1	<p>We recommend that a requirement to obtain data to support and substantiate the analysis of NR's efficiency gap is defined as an explicit objective for each workstream.</p> <p>This should influence the approach taken, the scope of analysis and the utilisation of the data obtained by the given workstream. It should also help to improve visibility of efficiency factors identified by NR (including initiatives already identified) and may provide additional insight in relation to how and why a given efficiency factor / opportunity is being prioritised, thereby helping justify its implementation.</p>
<i>Engagement with comparators</i>	
2011.BUB.2	<p>We recommend that NR develops a detailed engagement plan for each workstream, setting out specific steps through which it plans to initiate, establish and maintain contact. This should include both proposals for maintaining long-term, mutual engagement, and implementing procedures that ensure outputs are shared and mutual interests served, thereby maximising the prospects for obtaining meaningful data.</p>
2011.BUB.3	<p>We recommend that NR develops ties with further comparator organisations and look more widely than the present comparator pool which is heavily oriented towards European rail organisations. In particular, we consider NR should focus on benchmarking with light-rail / metro organisations both within the UK and internationally, for which we consider there to be a significant level of potential comparability in spite of differences in infrastructure characteristics. We also consider NR should explore contacts with non-European heavy rail organisations. (Note: recommended comparators and contact details have been provided by Arup/Imperial through feedback sessions provided for each workstream. Further details to be provided in Chapter 6).</p>

Ref.	Recommendation to Network Rail
<i>Benchmarking dataset</i>	
2011.BUB.4	We recommend that NR should explore means through which the amount of data - and the level of detail - can be increased from existing comparators. This should include exploring means through which existing datasets can be deepened and more observations obtained, e.g. through greater number of activities, increased use of questionnaires / automated data sharing.
2011.BUB.5	We recommend that NR develops a set of coordinated benchmarking data specifications / criteria, taking programme level objectives (in relation to overall efficiency gap) and setting these out as explicit data requirements for each workstream. This should include criteria to ensure a sufficiently representative dataset – e.g. for breadth of data in relation to areas of expenditure, level of depth and detail, sample size, and nature of comparative data. This should also account for incremental progress as the benchmarking progresses and increasing level of data are obtained (timetable / “roadmap”).
<i>Benchmarking data analysis</i>	
2011.BUB.6	<p>Recommendation: We recommend that each benchmarking workstream lead provides a detailed explanatory document setting out the key assumptions and rationale relating to benchmarking data outputs. This should include:</p> <ul style="list-style-type: none"> • details of incoming data adjustments and normalisation; • an explanation of the process by which key cost / efficiency factors have been identified and prioritised, and those factors excluded / marginalised from the analysis; and • details of the application and extrapolation of the data to higher-level expenditure and efficiency proposals / projections.
<i>Workstream-specific recommendations</i>	
2011.BUB.7	Track: we recommend NR undertakes further analysis to support its CP5 efficiency proposals for S&C renewals. This should include analysis of implementation cost relating to the initiatives proposed, and a detailed assessment of risks associated with the proposals.
2011.BUB.8	Track: we recommend that NR undertakes comparative analysis between itself and its peers by costing work banks seen overseas as if they were undertaken in the UK.
2011.BUB.9	Buildings & civils: we recommend NR undertakes further analysis to support its CP5 efficiency proposals for Buildings & Civils activities. This should include analysis of implementation cost relating to the efficiency initiatives proposed, and a detailed assessment of risks associated with the proposals.

Ref.	Recommendation to Network Rail
2011.BUB.10	<p>Maintenance: To deepen its analysis, we recommend NR to use time series data from other railways to identify where these organisations have driven down maintenance costs. This should allow NR to explore in detail whether these cost savings arise from efficiency, deferred maintenance or the impact of renewals or enhancement spending.</p>
2011.BUB.11	<p>Maintenance: to support NR’s internal maintenance benchmarking activity we suggest the merits of a Bayesian approach. To ensure transparency, we recommend that inputs & assumptions are clearly laid-out to support such analysis.</p>
2011.BUB.12	<p>Maintenance: For present and future internal benchmarking analysis of efficiency across MDUs, we recommend that NR considers the following statistical techniques:</p> <ul style="list-style-type: none"> • Semi/non-parametric estimation. • Confounding / omitted variable bias (OVB). • Normalised dependent variable. • Cost function with price data. • Use of temporal as well as cross-sectional analysis. • Analysis by category of spending. • ANOVA for grouping analysis.
<p><i>Alternative benchmarking approaches for analysis by ORR</i></p>	
2011.BUB.13	<p>We recommend that ORR considers alternative comparative analytical approaches to support its analysis of NR’s efficiency level and CP5 targets for the PR13 review process. We recommend the ORR considers the following:</p> <ul style="list-style-type: none"> • Analysis of efficiency potential using internal NR cost data. • Qualification of renewals capex according to quality of proposals. • Bottom-up engineering models for technical comparison.

Table 10 - NR bottom-up benchmarking: recommendations