SCOPING OPTIONS FOR BENCHMARKING HIGHWAYS ENGLAND'S CAPITAL EFFICIENCY

OFFICE OF RAIL AND ROAD

2ND MARCH 2018

FINAL

MAIN REPORT

Submitted by:

Cambridge Economic Policy Associates Ltd

in association with:

Institute for Transport Studies of the University of Leeds

CONTENTS

EXECUTIVE SUMMARY

Introduction

This report constitutes advice from Cambridge Economic Policy Associates (CEPA) and the Institute for Transport Studies (ITS) of the University of Leeds to the Office of Rail and Road (ORR) on potential options for benchmarking Highways England's capital efficiency.

ORR's key requirement for this project was to identify and assess the feasibility of a range of methods/approaches to benchmarking capex cost efficiency. ORR requested that our assessment of each potential option included a consideration of each option's feasibility (both for RIS2 and also in the longer term), robustness, data requirements, and effort levels required to undertake the option.

Approach

Our approach involved the following steps:

- *Selection and analysis of case studies*. We considered capex benchmarking approaches used in a number of other regulated sectors, focusing on sectors with a high level of network infrastructure, sectors subject to economic regulation and/or efficiency assessments, and those with publically available information.
- *Identification of potential options*. Based on our case study analysis, input from ORR and our wider knowledge, we drew up (and agreed with ORR) a shortlist of potential benchmarking options for assessment.
- *Assessment*. We collected evidence and assessed the different options on a consistent basis, focusing in particular on the key areas of interest from ORR's perspective: the applicability of the option to renewals and/or enhancements; the feasibility of undertaking the option in time for RIS2; the level of robustness of the method (including the level of risk associated with ORR relying on that method); the future potential of developing each method in the longer term; and the effort that would be required, either by ORR itself and/or by external parties.
- *Conclusions*. We reviewed our analysis/assessments for the different options to take a view on their relative merits, and thereby to set out our final conclusions.

Case study analysis

We considered the following UK sectors: rail, energy, water, airports, air traffic control and telecoms. In our analysis of each sector we considered the benchmarking approaches used, pros and cons, and their transferability to Highways England. Our case study analysis is summarised in Section 3, and the detailed case studies are in Annex A.

The key messages that emerge from assessing the case studies as a whole are as follows:

- Most regulators employ a mix of methods and efficiency assessment draws on a range of evidence.
- Some regulators use top-down econometric benchmarking between companies at a totex level. This approach is most evident in sectors with a large number of companies, e.g. water/sewerage, electricity and gas distribution, etc.
- Others such as ORR maintain a split between operations, maintenance and renewals and enhancements (large scale capex projects). This approach is consistent with Highways England's current policy and Government Accounting.
- Some regulators assess enhancements in a bespoke manner, e.g., Ofwat does not include all enhancements in its econometric benchmarking of wholesale costs.
- Related to this, the benchmarking approach taken is interlinked with the regulatory process, e.g. as part of the CAA's approach to capex governance, capex projects are assessed on an iterative basis, creating scope for bottom-up analysis.
- Some benchmarking approaches focus more on a company's processes or activities (i.e. inputs) than on the costs of those activities. For example, the capability reviews commissioned by ORR. These reviews may identify potential efficiencies from a different perspective to capex benchmarking, so both analyses should be cross-checked against one another to avoid the risk of double counting.
- Companies may be able to undertake internal bottom-up unit cost benchmarking but this can be more difficult for regulators given their limited access to data.
- The extent of capex efficiency benchmarking varies considerably between regulators. For example, CAA's approach to settings NATS' capex allowance has historically included very little capex benchmarking.
- Regulatory capex efficiency assessments can involve other pieces of information, e.g. company narratives on why an investment is required; consideration of the benefits from investment alongside the costs, etc.

Identification of potential options

We identified the following as potential options for capex benchmarking.

Table A: List of capex benchmarking options considered in this study

Title of option	Description of option
Top-down benchmarking across companies in the same sector	
International benchmarking	Top-down benchmarking against other overseas comparators in the same sector to assess a company's relative efficiency. In the case of Highways England, the comparators would be overseas National Road Authorities. Benchmarking could involve different techniques, e.g. econometric analysis and unit cost analysis.
National benchmarking	Top-down benchmarking against comparators in the same sector and country. In the case of Highways England, the comparators would be Local Road Authorities in the UK. A database of maintenance and renewal costs, as well as associated cost drivers, is currently in place, known as the CQC Efficiency Network (CQCEN) – see Annexes B and C for further details. As per international benchmarking, this can involve different techniques.
Internal benchmarking	Top-down benchmarking of different divisions or areas within the same company against each other (between Highways England's six regions in the case of Highways England). 'Adjacent highways networks' can also be compared (i.e. Scotland/Wales in the case of Highways England). As per international benchmarking, this can involve different techniques.
Project-level benchmarking	
Bottom up unit cost analysis	Companies with large capex programmes build detailed knowledge of unit costs over time and can use this to drive bottom up 'bills of quantity' type estimate of costs. Major engineering companies hold similar data and this can facilitate external comparisons. In the case of Highways England, POPE reports contain assessments of unit costs for certain aspects of major schemes.
Capex cost composition	Top-down review of the breakdown of pre-construction cost estimates for schemes. Costs are broken down by proportion into categories such as design, project management, overhead, contingency, risk, delivery, etc. The proportions can be compared across projects to locate projects out of line for further examination. Or an aggregate can be created and compared with external benchmarks, e.g. based on the experience of surveyors and engineers.
Other benchmarking options	
Productivity trend analysis versus other sectors	Top-down analysis of changes in costs over time for companies in different sectors, where this alternative sector has characteristics that are *somewhat* similar. Given comparators are in a different sector, it is not meaningful to compare *levels* of costs, so the analysis considers *changes* in costs over time. Includes two different options: 1) *Factor Productivity analysis (Total or Partial)*: Review of historic productivity improvements for other sectors, over business cycles. 2) *Top-down unit costs (or Real Unit Capital Expenditure)*: Review of changes in unit costs in other regulated sectors, over price controls.
Delivery	Comparison of outturn costs versus project budgets. This is typically done at a high level, i.e. it is a top-down measure.

Assessment

Below we summarise our assessment, focusing on the key areas of interest for ORR:

- *Feasibility*: The likelihood of whether an option can be undertaken in time for RIS2.
- *Robustness*: Whether the method would produce useful benchmarks for Highways England, including the level of risk associated with ORR relying on the method.
- *Effort*: The effort/cost that would be required to undertake the method.

A 'Red-Amber-Green' scale is used to highlight the desirability of each option from ORR's perspective. This rationale for this high level assessment is provided in Sections 5 and 6.

Table B: Summary of assessment of different potential options

Title of option	Feasibility for RIS2	Robustness for RIS2	Effort for RIS2
International benchmarking	Medium	Medium*	High
National benchmarking	Medium	Medium*	High
Internal benchmarking	High	Medium*	Medium
Bottom up unit cost analysis	High	Medium	Medium
Capex cost composition	Unclear **	Medium	Medium
Productivity trend analysis versus other sectors	High	Medium	Low
Delivery	High	Medium	Low

* The robustness of international, national and internal benchmarking may be constrained for RIS2 by limitations in data availability/consistency, but robustness for these methods is likely to be 'high' for RIS3.
** Despite uncertainties, our experience from other sectors suggests this option may be feasible for RIS2.

Conclusions

Each of the potential options has limitations, as shown by all options scoring 'Amber' for robustness. It might be risky for ORR to overly rely on one method for RIS2, so we recommend exploring as many options as possible within the time available to develop a broad evidence base. The benchmarking results from different methods will provide partial pieces of evidence that will need to be 'triangulated', i.e. assessed in the round.

For RIS2 it would be cost effective to undertake options that are highly feasible and low effort, i.e. productivity trend analysis versus other sectors and delivery (of outturn costs versus forecasts). Other options (involving slightly higher effort) should also be developed to ensure that ORR can assess the *level* of Highways England's costs – the most feasible being internal benchmarking and bottom up unit cost analysis. Further road sector benchmarking (international/national) could be explored either now or for RIS3.

1. INTRODUCTION

This report constitutes advice from Cambridge Economic Policy Associates (CEPA) and the Institute for Transport Studies (ITS) of the University of Leeds to the Office of Rail and Road (ORR) on potential options for benchmarking Highways England's capital efficiency.

This section provides context to this study, explains its objectives, and finally sets out the structure for the remainder of the document.

1.1. Context

ORR is responsible for monitoring and enforcing the performance and efficiency of Highways England. ORR holds Highways England to account for its management of the Strategic Road Network (SRN) and advises the UK Government on the levels of funding and performance requirements for future road periods to help frame challenging and deliverable performance and efficiency requirement.

With the next five year Road Investment Strategy (RIS2) approaching[1], ORR will shortly be providing advice to the Secretary of State for Transport to inform the development of RIS2. Specifically, later in 2018-19, ORR will advise DfT on the challenge and deliverability of Highways England's Strategic Business Plan, including an assessment of the efficiency proposals

ORR is therefore keen to understand the range of potential analytical tools that it might utilise to consider Highways England's capital efficiency. In this report we consider a range of cost benchmarking options. Cost benchmarking – which involves comparing costs (or changes in costs) to other companies or sectors, or even to itself over time or between its different regions – is a tool that can help to inform ORR's understanding of Highways England's current performance and in advising on its future performance trajectory. As such, ORR has commissioned CEPA and ITS to provide advice on the feasibility and suitability of different benchmarking options.

This report does not consider benchmarking options in relation to either performance/outputs or processes/capabilities, this is because ORR has been developing its evidence base in these areas via a number of other studies and reports including:

- ORR's 'capability reviews' of Highways England – these are reviews (jointly commissioned with Highways England) to assess Highways England's portfolio and

[1] To cover the period 2020 – 2025

programme management, asset management, and procurement and contract management capability. These were published in late 2017 and early 2018.

- ORR's benchmarking plan for Highways England, published in 2016.
- ORR's end-of year benchmarking progress reports from 2016 and 2017.

1.2. Objectives of this report

ORR's key requirement for this project was to identify and assess the feasibility of a range of methods/approaches to benchmarking capex cost efficiency. ORR requested that our assessment of each potential option include a consideration of the following aspects (as well as others that we considered relevant):

- *Feasibility/Deliverability*: Whether the approach could be completed by September 2018, in time to inform the RIS2 Efficiency Review.
- *Robustness*: Whether the approach would produce useful benchmarks that could be applied to considering Highways England's potential capex efficiencies in RIS2, and the amount of risk in relying on a given method.
- *Data requirements*: Whether the required data (on costs and cost drivers) is publicly available, and if not, whether it could be obtainable, and at what effort.
- *Effort/Cost*: The effort or cost that would be required to undertake this option.

1.3. Structure of this report

The structure of the main body of this report is as follows:

- Section 2 describes our approach to undertaking this study.
- Section 3 contains our 'case study analysis'. We considered capex benchmarking approaches used in a number of other sectors, including rail, energy, water, airports, air traffic control and telecoms. The analysis of each sector was written up as individual case studies, considering the approaches used, pros and cons, and their suitability to Highways England. This chapter provides an overview of our analysis.
- Section 4 describes the list of potential benchmarking options. We define the various potential benchmarking options that we assess within this report. This list was developed primarily from our case study analysis in Section 3.

- Section 5 contains our assessment of the different potential benchmarking options. Taking the list in Section 4, we assessed the different options, considering their feasibility, robustness, etc. This chapter provides an overview of our assessment.
- Section 6 sets out our conclusions. In light of our assessment for each option (from Section 5), we provide our conclusions on benchmarking options for Highways England, both for RIS2 and further into the future.

In a separate document, we present three supporting annexes which provide the detailed analysis that underpins the main body of the report. The order and content of the annexes is as follows:

- Annex A contains our detailed write-ups of the individual case studies. This detailed analysis underpins the overview of our case study analysis that is presented in Section 3 in the main report.
- Annex B presents our detailed assessment of the different benchmarking options. This detailed analysis underpins the overview of our assessment that is presented in Section 5 in the main report.
- Annex C considers the potential for benchmarking Highways England against other road authorities. It contains a detailed discussion on the potential for top-down benchmarking of Highways England against both overseas National Road Authorities and Local Road Authorities in the UK. This annex highlights the various issues that emerge when considering these benchmarking options, and provides thoughts on how ORR might seek to develop these benchmarking options, particularly in the longer term.

2. APPROACH

In this section we describe the stages of work that we have undertaken.

2.1. Selection and analysis of case studies

The requirement to assess the efficiency of capital expenditure is clearly highly relevant to regulated industries, although importance varies according to the capital intensity of that industry. Some regulators, including ORR, have been considering efficiency of the companies or sectors that they regulate for a number of regulatory cycles, and therefore have useful experience which can be applied to Highways England where efficiency assessment is in the early stages of development.

To ensure that ORR is able to build on lessons learned from elsewhere, the first stage of this project was to consider capex benchmarking approaches used in a number of other regulated sectors. Ideal sectors are those with a high level of network infrastructure, sectors subject to economic regulation and/or efficiency assessments, and those with publically available data and information. In our case study analysis we considered the following UK sectors: rail, energy, water, airports, air traffic control and telecoms. The analysis of each sector was written up in an individual case study, which considers the approaches used, pros and cons, and their transferability to Highways England.

Our case study analysis is summarised in Section 3, and the detailed case studies are in Annex A.

2.2. Identification of potential options

Based on our case study analysis, input from ORR and our wider knowledge, we drew up (and agreed with ORR) a shortlist of potential benchmarking options for assessment. This list included options applied in other sectors (either in the past or currently), as well as some suggestions from ORR. This list is set out in Section 4, and we highlight where such options have been used in other sectors.

2.3. Approach to assessing options

We also drew up and agreed with ORR a template that we used to collect evidence and write-up our analysis of each benchmarking option on a consistent basis. These templates were then used as the basis for our assessment.

Table 2.1: Template for evidence, analysis and assessment

Item	Additional explanation of item, where relevant
Description of benchmarking method	Clear description what the method involves, and if there are ranges and options.
What type of expenditure it can be used to benchmark	Mention in particular suitability for: capital renewals, totex, enhancement, or whether actually this kind of method really is intended for something else like opex.
Absolute benchmark, or trend benchmark	For example, productivity trend analysis doesn't provide an absolute benchmark.
Frontier efficiency, catch-up efficiency, or something else	Generally speaking, catch-up efficiency is usually calculated with respect to an absolute benchmark, whereas frontier efficiency can be informed by either absolute or trend benchmarks.
Regulatory context appropriate or necessary to make such a method relevant	
Principal example(s) of past usage	Includes references to any relevant case studies
Lessons learned from past usage	e.g. usefulness. Clarify if the method is no longer used, and if so why not.
Practical options for applying such a method to Highways England	An overview of options, with more detail in boxes below. If there are no practical options, say so and stop here.
Robustness: would it produce useful benchmarks in the context of regulating Highways England	Issues to take particular account of, e.g. the differences between Highways England and other regulated companies.
What the benchmarks would be for Highways England	i.e. who are the comparators
What kind of data from Highways England would be needed for application	For example, some methods (e.g. factor productivity analysis) do not require any data from Highways England.
Existence of such Highways England data and/or prospects of making it exist	
What external data would be needed, does it exist, is it public, other methods of accessing it, prospects for making it exist	
Estimate of timescale for carrying out such work	Note in particular the deliverability for September 2018, or if it would be a longer term aspiration
Estimate of effort levels	
Summary assessment of method for application to Highways England	

2.4. Assessment of options

Our 'summary assessment' (the last box in the table) was made by taking into account the various pieces of evidence/analysis in the template, and then focusing in particular on the key areas of interest from ORR's perspective:

- *Applicability to different types of expenditure* – whether this option is well-suited to renewals and/or enhancements.

- *Current feasibility* – whether the timeframe and level of effort required for carrying out this option is sufficient to allow the results to inform ORR's RIS2 efficiency assessment for Highways England. ORR indicated that the option would need to be undertaken by September 2018.
- *Robustness* – whether the method would produce useful benchmarks in the context of regulating Highways England, and including the level of risk associated with ORR relying on that option for the purposes of assessing Highways England's capex efficiency.
- *Future potential* – whether this option is sufficiently relevant that it would merit ORR pursuing this method in the long term, such as for RIS3, e.g. via data collection, stakeholder collaboration, the establishment of 'benchmarking clubs', etc.
- *Effort/Cost* – The effort or cost that would be required to undertake this option, by ORR itself and/or by external parties.

2.5. Development of conclusions

The case studies and option assessments together provide the evidence base from which we draw the conclusions that are presented in section 6. In preparing conclusions we have considered options that might be feasible in the short term and those that require a longer time period for development. We think that short term options should be used, wherever possible, as the early stage of developing more complex approaches that will take time to establish.

3. CASE STUDY ANALYSIS

As noted above in Section 2, for this project we considered capex benchmarking approaches used in a number of other sectors. The analysis of each sector was developed into an individual case study. This chapter provides a list of the case studies that we developed, and then sets out the key findings and observations that emerge from them.

3.1. List of case studies

As discussed and agreed with ORR, we undertook an analysis of the approach to capex benchmarking in the following sectors:

- **Water**: Ofwat's approach to wholesale cost assessment for the England and Wales water and wastewater companies.
- **Energy**: We considered three separate energy networks in Great Britain (GB) – electricity distribution, gas distribution, and transmission (both electricity and gas). We reviewed Ofgem's approach to assessing cost efficiency.
- **Rail**: ORR's approach to assessing Network Rail's capital cost efficiency.
- **Air transport**: We focused on CAA's approach to regulating NATS (air traffic control), although we also utilised our broader regulatory experience of CAA's approach to regulating Heathrow Airport's capex.
- **Telecoms**. Ofcom's approach to cost assessment for BT Openreach.

3.2. Summary of individual case studies

Below, we provide an overview of each individual case study.

Water - England and Wales water and wastewater companies (Ofwat)

The current price control (PR14) is focused on top-down benchmarking of totex through the use of regression models and unit cost models. Ex post the regression modelling, there was then a process by which companies could make claims for increased cost allowances based on company specific characteristics (special factors) or bespoke projects - claims were assessed on a case-by-case basis.
In PR09, benchmarking had two stages. The first, was for Ofwat to estimate the costs of each company. Ofwat then applied an efficiency challenge to these costs for each company. A cost base comparative tool was then used to assess the relative efficiency of capital projects, with regards to procurement and delivery. Also, an efficiency challenge was introduced to account for increases in efficiency in the economy (frontier shift). Capital costs were benchmarked separately to opex.

Energy - Network operators in GB energy networks: Gas and electricity; transmission and distribution (Ofgem)

For the distribution networks (gas and electricity), there are multiple companies within each sector. For transmission there is a single operator. This affects the benchmarking options.
Distribution: current price controls (e.g. RIIO-ED1), Ofgem benchmarked the companies' costs against each other, in order to derive a view of the efficient level of total costs for the industry. A number of different models were used - disaggregated activity level models, and bottom-up and top-down totex models. For the previous price controls (e.g. DPCR5), capex and opex were assessed separately. For capex, a range of different techniques were used to compare the costs of the DNO's against each other, e.g. age based asset replacement modelling, unit cost benchmarking, etc.
Transmission: Ofgem's approach varied by the type of expenditure. Some international benchmarking was undertaken, e.g. comparison of unit costs versus the Scottish transmission operator, although Ofgem also considered trend analysis (comparing company performance over time) and company justification.

Rail – Network Rail (ORR)

Currently ORR is using three formal data-driven cost benchmarking procedures ('Levels'), two of which address renewals costs; the third only covers maintenance costs. All three of these benchmarking procedures are top-down cost benchmarking, taking the total relevant cost and trying to explain it econometrically by a number of drivers. Costs are not unitised. Level 1 is an international benchmarking of maintenance and renewal costs (i.e. totex) using an international dataset from of around 15-16 countries from UITP. A long period of data is available. Level 2 is based upon comparing Network Rail's routes maintenance and renewals costs against each other. NR is divided into 10 routes. Data is available for 5 years. Level 3 only addresses maintenance costs. These components have long been part of ORR's approach but the weight placed on each has varied overtime.

Air traffic control – NATS; Airports – Heathrow (CAA)

Air traffic control (NATS): CAA has only undertaken very limited benchmarking. For the current control period (RP2, 2015-2019) CAA undertook a simple unit cost comparison of NATS' historic capex versus ANSPs (air navigation service providers) in different countries. For the previous control period (CP3, 2011-2014), CAA undertook simple comparisons of NATS' capex-to-opex ratio versus ANSPs in different countries, and simple comparison of changes in NATS' capex level over time versus ANSPs in different countries.

Heathrow Airport: A protocol by which Heathrow Airport and the Airlines discuss / agree capex proposals, assessing at various stages through project development. CAA reviews the projects and can disallow inefficient expenditure ex post.

Telecoms – BT Openreach (Ofcom)

Based on Ofcom's currently published consultation documents, Ofcom has not identified any external benchmarks (i.e. other companies) that will be used as benchmarks for assessing BT's proposed capex. Instead, Ofcom primarily intends to consider BT representations, capex efficiency targets set using BT's historical capex data, and simple efficiency assumptions. Ofcom notes that the scope for capex efficiency gains may be lower than opex, with large proportions of capex being related to external contracts that are not always negotiable and often rise with construction prices.

For further details on each case study see Annex A.

3.3. Key messages from across the case studies

The key messages that emerge from assessing the case studies as a whole are as follows:

- Most regulators employ a mix of methods and efficiency assessment draws on a range of evidence.
- Some regulators use top-down econometric benchmarking between companies at a totex level, e.g. Ofwat. This approach is most evident in sectors with a large number of companies, e.g. water/sewerage, electricity and gas distribution, etc.
- Others such as ORR maintain a split between operations, maintenance and renewals and enhancements (large scale capex projects). This approach is consistent with Highways England's current policy and Government Accounting.
- Some regulators assess enhancements in a bespoke manner, e.g. Ofgem split investment between core and non-core at DPCR5, Ofwat does not include all enhancements in its econometric benchmarking of wholesale costs, CAA has a separate governance process for approving Heathrow Airport's capex, etc.
- Related to this, the benchmarking approach taken is interlinked with the regulatory process, e.g. as part of the CAA's approach to capex governance, capex projects are assessed on an iterative basis, providing greater scope for bottom-up efficiency assessments on a project-by-project basis. Price controls tend to require at least some top-down benchmarking to assess the order of magnitude.
- Some benchmarking approaches focus more on a company's processes or activities (i.e. inputs) than on the costs of those activities. For example, the capability reviews commissioned by ORR. These reviews may be able to identify potential efficiencies, but do so from a different perspective to capex

benchmarking. Therefore, the potential efficiency targets implied by both types of analysis should be cross-checked against one another to avoid the risk of double counting.

- Companies may be able to undertake internal bottom-up unit cost benchmarking but this can be more difficult for regulators given their limited access to data.
- The extent of capex efficiency benchmarking varies considerably between regulators. For example, CAA's approach to settings NATS' capex allowance has historically included very little capex benchmarking, and Ofcom's recent determination for BT Openreach did not use other companies as benchmarks.
- Regulatory capex efficiency assessments can involve other pieces of information. For example: (1) company narratives on why an investment is required; (2) age profiling of assets, e.g. low historic investment could support the case for higher future investment; (3) benefits – a proposed investment might be justified in terms of high benefits, even if its costs are high; and (4) changing technology, e.g. Smart Motorways involve increased capex relative to historic trends.

In Section 4, we provide definitions (and a further explanation) of the benchmarking options identified, and highlight the sectors in which such methods are applied.

4. LIST OF BENCHMARKING OPTIONS

In this chapter we list the benchmarking options that we identified from the case studies and our own experience and provide definitions for these options. Our assessment of these options is set out subsequently in Section 4.

4.1. List of options

In the table 4.1 below we set out the options that we have identified as being potentially relevant to ORR's work on Highways England. For each of the options, we provide the following:

- A title, so it can be clearly identified later in this report;
- A definition, so it is clear what we mean; and
- Examples of where an option has been used/applied by regulators in other sectors.

To provide greater clarity on the different potential benchmarking options, we have grouped them into the following categories:

- Top-down benchmarking across companies in the same sector;
- Project-level benchmarking; and
- Other benchmarking options, including comparisons to other sectors.

4.2. Qualifying remarks

In the table below, we present the different options as distinct alternatives in order to make our analysis as simple as possible to understand. However, we note that benchmarking options can overlap and precise techniques vary between sector and case study. For example, if ORR were to assess Highways England's major scheme costs on a bottom-up basis, and compare these cost components to other overseas highway authorities, this could in theory be described as 'bottom up' benchmarking or 'international' benchmarking.

Table 4.1: Benchmarking options: Definitions and examples of usage from case studies

Title of option	Description of option and examples of usage based on case study analysis
Top-down benchmarking across companies in the same sector	
International benchmarking	Top-down benchmarking against other overseas comparators in the same sector to assess a company's relative efficiency. In the case of Highways England, the comparators would be overseas National Road Authorities. Benchmarking could involve different techniques, e.g. econometric analysis and unit cost analysis. Examples of international benchmarking in other sectors are: • ORR's regulation of Network Rail. Applied to renewals. • CAA's regulation of NATS (the provider of National Air Traffic Services in the UK). Only very limited benchmarking undertaken.
National benchmarking	Top-down benchmarking against comparators in the same sector and country. In the case of Highways England, the comparators would be Local Road Authorities in the UK. A database of maintenance and renewal costs, as well as associated cost drivers, is currently in place, known as the CQC Efficiency Network (CQCEN) – see Annexes B and C for further details. As per international benchmarking, this can involve different techniques. Examples of usage from our case studies are: • Ofwat's regulation of England and Wales water/wastewater companies. Applied to renewals and some enhancements. • Ofgem's regulation of GB energy networks. Applied to totex (total expenditure), which includes renewals and enhancements.
Internal benchmarking	Top-down benchmarking of different divisions or areas within the same company against each other (between Highways England's six regions in the case of Highways England). 'Adjacent highways networks' can also be compared (i.e. Scotland/Wales in the case of Highways England). As per international benchmarking, this can involve different techniques. Examples include: • ORR's regulation of Network Rail. Applied to renewals. • ORR's analysis to date of Highways England's maintenance and renewals costs.
Project-level benchmarking	
Bottom up unit cost analysis	Companies with large capex programmes build detailed knowledge of unit costs over time and can use this to drive bottom up 'bills of quantity' type estimate of costs. Major engineering companies hold similar data and this can facilitate external comparisons. In the case of Highways England, POPE reports contain assessments of unit costs for certain aspects of major schemes. Examples in other sectors include: • Ofwat's regulation of England and Wales water/wastewater companies. Applied to large and/or company-specific schemes. • ORR's regulation of Network Rail. Engineering led reviews of specific enhancement proposals and costs. • CAA's regulation of Heathrow Airport's capital expenditure.
Capex cost composition	Top-down review of the breakdown of pre-construction cost estimates for schemes. Costs are broken down by proportion into categories such as design, project management, overhead, contingency, risk, delivery, etc. These proportions can be compared across projects to locate projects out of line for

Title of option	Description of option and examples of usage based on case study analysis
	further examination. Or an aggregate can be created and compared with external benchmarks. External benchmarks might be based on the experience of surveyors and engineers. • For example, work commissioned by ORR to understand the detailed costs underpinning Highways England's major projects. (Only limited information is available.)
Other benchmarking options	
Productivity trend analysis versus other sectors	Top-down analysis of changes in costs over time for companies in different sectors, where this alternative sector has characteristics that are *somewhat* similar. Given comparators are in a different sector, it is not meaningful to compare *levels* of costs, so the analysis considers *changes* in costs over time. Includes two different options: 3) *Factor Productivity analysis (Total or Partial)*: Review of historic productivity improvements for other sectors, over business cycles. 4) *Top-down unit costs (or Real Unit Capital Expenditure)*: Review of changes in unit costs in other regulated sectors, over price controls. Examples of use in other sectors are as follows: • Ofgem's regulation of GB energy networks. Applied to totex (total expenditure), which includes renewals and enhancements. • ORR's regulation of Network Rail. Applied to renewals and enhancements. • CAA's regulation of Heathrow Airport.
Delivery	Comparison of outturn costs versus project budgets. This is typically done at a high level, i.e. it is a top-down measure. • Work commissioned by CAA to inform analysis of Heathrow Airport's opex. • Work commissioned by ORR to inform analysis of Highways England's opex.

5. ASSESSMENT OF OPTIONS

In this Section we provide a summary assessment of the potential options (see Section 4) against the key criteria noted in Section 2. Each option has its own sub-section below. Further detailed assessments are in Annex B.

5.1. International benchmarking

Renewals focused. International benchmarking would be done on a top-down basis, so it would be very difficult to obtain any meaningful modelling results for enhancements, given that they tend to vary significantly over time, between organisations, and are relatively bespoke in nature. So the relevance of this method would likely be restricted to renewals. This is consistent with ORR's approach to Network Rail, where ORR is not currently undertaking benchmarking of enhancement costs.

Current feasibility. This option may be applicable in time for RIS2, but only for a narrow range of comparators at best (no more than two or three), given the lack of an existing dataset or a 'benchmarking club' to collect and share data (for costs and cost drivers) on a consistent basis. ITS has been involved in research using renewals data for the Swedish strategic network – this could be explored for use in RIS2 benchmarking, subject to agreement. Other potential comparators could be explored (ITS has links with Norway).

Whilst this provides a potential option, international collaboration is not straightforward, in terms of setting up cost-sharing arrangements (e.g. confidentiality), determining robust definitions and obtaining data on explanatory variables. Therefore there would be a risk of relying on this method providing robust benchmarking results in time for RIS2.

Finally, the OECD International Transport Forum (ITF) is in the process of establishing an international database of road construction costs, but this is at a very early stage.

Robustness. In terms of its robustness of approach, international benchmarking allows analysis of a wide range of companies and approaches, which enlarges the efficiency frontier. It has limitations, notably that data collection can be problematic, both in terms of physically collecting the data and then ensuring consistency across companies and is therefore time consuming. It may be difficult to incentivise multiple countries to collaborate on benchmarking, given there are different regulatory environments, i.e. regulated and public sectors. Given international datasets are not yet available for highways, any international benchmarking undertaken for RIS2 should not be solely relied upon, but should be considered alongside the results from other methods.

Future potential. Despite its limitations, we recommend that ORR considers the development of international benchmarking over time, given the promise shown by the BEXPRAC 2010 study, and the potential to enlarge the efficiency frontier beyond the UK. ORR has also acknowledged its aspiration to develop an international 'benchmarking club' of national road authorities, e.g. The OECD-ITF project to establish an international dataset. However, developing such international datasets/benchmarking clubs will be a long term venture (e.g. for RIS3). There may be value in discussing with ORR's rail team the lessons learnt its international benchmarking in the context of Network Rail.

The long term potential may also depend on how the results are used. For example, if infrastructure managers know that benchmarking results could potentially be used to identify them as 'bad' performers, there is limited motivation to participate and/or to provide robust data. On the other hand, if more emphasis is placed on using the benchmarking results to understand the drivers of performance, there may be greater incentive for participation. Lessons could be learnt from the experience of (national benchmarking facilitated by the CQCEN which is discussed further below.

Effort/Cost. The method would involve a high level of effort/cost, and it is not something that can be done in time for RP2. A large sample of high quality data would be required to undertake meaningful international benchmarking. This would require buy-in from overseas regulators and/or National Road Authorities, collecting data, checking its consistency over time and between countries (e.g. with respect to cost allocation), making refinements, ensuring process are in place to ensure consistent data collection in future years, etc. For example, Ofwat's previous annual data collection process (the 'June Returns') took many years to develop and refine before high quality data was being collected, e.g. companies take time to reach consensus on cost allocation approaches.

5.2. National benchmarking

Likely to be renewals-focused. The CQCEN collects data on capex, and its analysis to date has focused on maintenance and renewals. It is not yet clear whether it could be used to benchmark Highways England's enhancement costs. If feasible, national benchmarking would be best suited to renewals as it is a top-down method.

Current feasibility. Some initial benchmarking could be undertaken on renewals in the short term, as capex data is collected for local authorities (LA) via the CQCEN. Comparability with HE would be limited because LA cost data is currently aggregated across all road types (including minor roads), and CQCEN benchmarking to date has covered maintenance and renewals. There is, however, a possibility that A-road only cost

data could be collected from a sample of LAs in the short term, although confidentiality agreements might be required. There is also the issue of ensuring that potential cost drivers and explanatory variables are also collected on a consistent basis. Overall, if this option were pursued, it would likely provide relatively contextual benchmarking results, as it would take longer than six months to integrate Highways England into the CQCEN on a consistent basis. We also note that CQCEN also collects data on a number of variables related to 'quality' and 'customer satisfaction', and it is not clear if Highways England would want/be able to provide this additional data.

Robustness. If Highways England were to be benchmarked against LAs using the CQCEN dataset, ORR would benefit from being able to compare Highways England against a large sample size (both the number of years and the number of organisations). Benchmarking is already being undertaken using CQCEN data, so existing CQCEN members are already a long way up the learning curve in terms of generating consistent data for costs and cost drivers. National benchmarking (versus internal benchmarking) would have the benefit of widening the range of comparators to beyond Highways England.

However, there would be limitations due to differences in the types of activities undertaken between comparators, e.g. Highways England's network has more major roads, Highways England is a larger organisation, and its 'centrally managed expenditure' would need to be allocated across regions. Although data consistency issues between LAs are relatively well understood (i.e. 'lessons learned' could be applied to Highways England), consistency takes time to achieve, so any benchmarking for RIS2 would be limited and should be considered alongside results from other methods rather than alone.

Future potential. In light of the pros and cons above, we would recommend exploring national benchmarking as an option for RIS3. National benchmarking would be a useful compliment to international and/or internal benchmarking, and could even be combined for the latter, i.e. if Highways England is compared on a region-by-region basis. The most straightforward option would be to seek to integrate Highways England into the CQCEN but there would be challenges in terms of collected data from Highways England and ensuring it is consistent/comparable with data collected from the LAs.

The CQCEN has sought to place an emphasis on benchmarking being used in a positive way, by seeking to understand the drivers of best practice, by sharing this knowledge between LAs, and by allowing LAs to seek to justify why their costs might different from other LAs. This has been beneficial in terms of encouraging 'buy-in' from LAs and increases the likelihood of national benchmarking being a viable benchmarking tool in the future in the context of HE, although assuming that ORR were to use the data in the same 'spirit'.

Given that DfT is currently consulting on the establishment of a Major Road Network (MRN), any LA data that could be obtained that is specific to the MRN would provide a very direct comparison for Highways England, given the roads are relatively major. Given the MRN is at Government consultation phase at present, this is an option for RIS3 rather than now.

Effort/Cost. The CQCEN is already established and data is collected on an annual basis, so this benchmarking option likely has a lower cost than international benchmarking. However, the level of effort/cost required is still reasonably high: (1) there would be a cost of HE joining the CQCEN; (2) work would be required to ensure data consistency across HE and the comparators; and (3) the CQCEN dataset would need to be widened if ORR wished to benchmark enhancement costs.

5.3. Internal benchmarking

Likely to be renewals focused. As noted previously, enhancements are generally difficult to model using top-down approaches, so even in the longer term internal benchmarking at a top-down level is likely to be suitable for renewals only. Bottom-up analysis would likely be more applicable for enhancements, although there is a question of how much of Highways England's internal cost databases ORR will be able to access.

Current feasibility. Compared to international and national benchmarking, this option offers greater feasibility in the short term, e.g. ORR's 2017 benchmarking report includes some very high level unit cost comparisons between Highways England's routes for renewals and maintenance costs. ORR is also currently developing some econometric (multi-variable) models to assess the efficiency of maintenance and renewals expenditure between regions. However, ORR has been open that the renewals benchmarking is subject to limitations (e.g. limited comparability between regions), and if its econometric benchmarking models are not developed sufficiently in time for RIS2, any internal benchmarking would likely be restricted to being fairly high level and basic metrics such as renewals cost per km laid.

Enhancement costs in each region will vary considerably depending on the major scheme programme, and therefore the relevance of this option for RIS2 would likely be restricted to renewals. ORR has not given any indication that it intends to benchmark enhancements at a regional level.

Robustness. In general, the main advantage of this option is that it is highly controllable from ORR's perspective and offers a single source of benchmarking data. ORR has also indicated that its (work-in-progress) econometric models for Highways England's regional

maintenance/renewals costs are generating plausible results. However, ORR's analysis has flagged that there are limitations to the modelling work to date, so we would recommend a cautious approach to interpreting the results of this (relatively early stage) internal benchmarking. Finally, it is important to recognise that internal benchmarking does not necessarily set a particularly ambitious efficiency frontier as it does not include any benchmarks which are external to Highways England, although this could be mitigated if Highways England's regions were also to be compared with LAs.

Future potential. This option is highly controllable from ORR's perspective and offers a single source of benchmarking data, so should be pursued. Regional benchmarking overcomes a number of the limitations of international benchmarking, because there is similarity in accounting methods, design standards, currency, etc. So ORR should focus efforts on establishing a robust dataset for RIS3, ensuring that data is recorded on a consistent basis across regions and/or over time, and that Highways England provides data on potentially relevant cost drivers. This would also facilitate trend analysis, i.e. comparing different regions over time, as well as at a fixed point in time.

Highways England could seek inclusion in the CQCEN on a region-by-region basis, although this may not be straightforward due to differences in network type, organisation structure, activity portfolio, approach to data recording, etc. A further option would be to consider bottom-up benchmarking between regions – this is discussed below.

Effort/Cost. The level of effort/cost of this option would depend on the technique applied within the analysis. For example, unit cost analysis is relatively straightforward and could be undertaken in a matter of weeks (given that ORR already collects data annually from HE on a regional basis), but robust econometric analysis would require months of work. However, if unit cost analysis is pursued in the short term, any potential data limitations should be acknowledged. Regional level data needs to be as consistent as possible, both between regions and over time (e.g. in terms of cost categories/definitions, cost allocation, etc.) and improving consistency would likely incur a cost, e.g. joint engagement between HE's regional teams, possibly with oversight/encouragement from ORR.

5.4. Bottom up unit cost analysis

Likely to be renewals focused. A combined method covering renewals and enhancements is unlikely to be practicable given the lumpy nature of capex and the bespoke make up of many capex projects. These comparisons are likely to be more effective if combined with the benchmarking of maintenance. Rail companies tend to compare these costs both together and separately given the trade-off that companies can chose to make between,

maintenance and renewals, as well as the difficulty in defining the distinction between 'heavy maintenance' and renewals. Benchmarking of specific activities and/or cost components using the 'bill of quantities' could be undertaken for renewals, but also potentially for enhancement costs, e.g. HE has used this approach to seek to demonstrate the efficiency of its scheme costs for Smart Motorways.

Current feasibility. This method considers disaggregated costs by activity i.e. it is bottom up rather than top down. Highways England is separated into regions and our understanding is that work is devolved into the regions which are responsible for delivery and cost. It should therefore be practical to undertake a degree of bottom up unit cost based benchmarking of some renewals across the regions. This would need to commence as soon as possible because we know that obtaining reliable and consistent data over time is time consuming and the work would be of most value if data is sourced for a number of years. In rail renewals comparators include for example track maintenance and renewal. In roads we would expect it to be possible to compare as a minimum pavement renewal, but comparison might also extend to bridges, drainage etc. assuming that a dataset of sufficient size can be developed within ORR's timescales.

It may be practical to provide external comparators for some renewals benchmarks. We are aware that some of the large QS practices hold such databases and access to those to provide an external view of Highways England regions' relative efficiency would be beneficial. However ensuring data comparability with Highways England is challenging.

Robustness. As with other forms of benchmarking there are limitations. The work is only as good as the data set underpinning it and initially this might be limited, particularly for some types of renewal. In the early stage of development the efficiency projections might best be used to identify areas of best practice within the business that can be shared and result in a reduction to costs rather than as a firm target. Quantitative efficiency gaps identified could however form part of a wider evidence base.

Future potential. This approach is likely to have most application/benefit in the short to medium term and could be replaced by more sophisticated econometric approaches in the longer term as the data set grows and becomes more robust. It is however an important catalyst for encouraging the business to consider the drivers of regional differences and what might need to change for Highways England to deliver renewals at costs that are comparable with the best internationally.

Effort/Cost. As a very rough estimate it would take 4-6 months and a not insignificant budget/level of resource to get a data set together and produce initial results assuming that HE is committed to the work and willing to spend time assembling the required data.

While a third party could assist and advise on the process, ultimately HE would need to extract the data from its systems in the required format.

5.5. Capex cost composition

Renewals or Enhancements focused. This method can be applied to both renewals and enhancements. Enhancements are likely to be more diverse, so less intelligence may result. However, given that most of the other benchmarking methods discussed in this report primarily apply to renewals, analysis of capex composition/cost breakdown could provide at least a high level/initial view of efficiency for enhancements.

Current feasibility. Currently unclear – requires further research in cooperation with ORR, and possibly Highways England. However, Nichols (March 2017) carried out a study for ORR of a 10% sample of major projects, and reported that costs are well recorded, so clearly had sight of such data. However whether the data is in a suitable form for these purposes would require further investigation. For example, ensuring that indirect costs are consistently defined is not a straightforward task.

Robustness. Moderate. Comparing various Highways England projects may raise avenues for further query as to why certain projects are out of line, and may result in some rectification. Comparing with external organisations and noting material variance from expected project spend breakdown may also be productive. Given special circumstances this method is unlikely to lead to a robust estimate of overall inefficiencies but it would provide evidence to support wider analysis of efficiency.

Future potential. In the event that suitable information is not currently available, it would be very feasible to agree a data protocol with Highways England to ensure that all projects have costs recorded in a suitable to facilitate such benchmarking.

Effort/Cost. This would depend on the depth and detail of the study. But we would expect a high level study deliverable in the short term could be delivered in circa 3 months. However, this assumes that HE would be able to provide data in a timely fashion.

5.6. Productivity trend analysis versus other sectors

5.6.1. Factor Productivity analysis

Capex-focused (renewals and enhancements combined). The EU KLEMS database doesn't distinguish between different types of capital, so we would not necessarily recommend applying this method separately to renewals and enhancements. However, we note that

it could be possible, i.e. by varying the 'comparator' sectors that are selected when developing the composite productivity measures for renewals versus for enhancements.

Current feasibility. Productivity growth measures (TFP and PFP) could certainly be calculated in time for RIS2 – the dataset (EU KLEMS) is publically available and the approach is well-established. Calculating the metrics does require some familiarity with the data and formulas that underpin the measures.

Robustness. These measures have a number of limitations – they are relatively high level and generic, they provide an indication of frontier shift efficiency only, and they only provide a proxy measure of productivity because it is not possibly to directly observe 'quantity' measures for outputs and capital inputs. On the other hand, they are able to provide a very high level (top down) starting point/cross-check for the level of efficiencies that might be expected from Highways England over time. Its acceptance by regulators in different sectors suggests that it does have value as a method. Given that this method is often used due to its feasibility (as opposed to its precision), this method should be considered alongside methods if it is being used to inform setting efficiency targets.

Future potential. The underlying dataset is administered via an EU-wide initiative, and data is only collected every few years. It seems unlikely that ORR would be able to influence how this data is collected, so the potential for this metric at RIS3 is unlikely to be any greater than it is currently.

Effort/Cost. Overall, the effort is low relative to some other benchmarking methods, because the data inputs are readily available and the approach is well-established. The timescale would likely be circa one month for an organisation familiar with this method – this would include updating the dataset, selecting/weighting the sectors (and discussing these with the company/regulator to obtain buy-in), calculating different measures, and analysing results. However, the level of effort required would be higher for organisations (or third parties) with less familiarity with this method.

5.6.2. Top down unit cost analysis

Renewals- or Capex-focused. As noted above, there are limitations of applying this method to capex benchmarking. If this method were applied to capex, it may be better suited to renewals, as these costs tend to be more consistent over time than enhancements. However, it may not be straightforward to collect 'renewals only' data, as different industries categorise costs in different ways, and differences in categorisation might also apply to whichever cost driver was chosen to calculate the unit costs. In light of this, the analysis might need to be undertaken at a capex (or even totex) level.

Current feasibility. Calculating/assessing changes in real unit capex across a range of companies and sectors could be feasible in time for RIS2. It would require starting reasonably soon, because data would need to be collected over a long time series (see the point below) and would likely need to be interrogated to understand the magnitude of capex programmes in different sectors.

Robustness. The main benefit of this method is that it is a top-down approach that provides a high-level sense check on whether the changes in Highways England's unit costs are broadly in line with other sectors. However, this method is typically applied to opex as it is more consistent over time, so efficiencies are easier to identify. Capex trends on the other hand are often cyclical, driven by Government programmes, so this method may have only limited potential for capex benchmarking.

Future potential. Having more time to develop a long term capex dataset would inevitably improve the quality of the analysis, so this could be worthwhile for RIS3. However, this method only focuses on changes in costs over time (as some comparators are only partially comparable), so it would probably be of greater value for ORR to prioritise benchmarking Highways England against other road/highway comparators, which we refer to elsewhere as 'international', 'national' and 'local' benchmarking. This type of benchmarking would enable ORR to analyse levels of unit costs, not just changes.

Effort/Cost. As noted above, analysis of changes in capex unit costs could likely be undertaken in time for RIS2, i.e. within the next 3-6 months. The main time requirement would be the collection and analysis of consistent capex dataset over time across comparators.

5.7. Delivery

Broad application. This method could be applied to capex as a whole, or enhancements and/or renewals separately.

Current feasibility. This is certainly achievable in time for RIS2. ORR already monitors outturn costs in relation to previous forecasts using a range of approaches, including the cost performance indicator (CPI). It would likely require minimum effort from all parties involved in RIS2.

Robustness. This method can be used to quantify cost overruns, and therefore to inform the scope for additional future catch-up efficiencies. It can be applied either by providing a broad sense of any inefficiencies in delivery, or could be used more precisely by

requiring that a certain proportion of cost overruns are added to the efficiency challenge in the next RIS.

However, this method needs to be carefully applied. Firstly, cost overruns do not necessarily indicate inefficiency, because they could be due to other factors, e.g. increased scope, uncontrollable external factors, etc. Secondly, this method does not reveal *definitively* whether costs at a point in time are efficient, but rather it highlights *potential* areas of historic inefficiency. This is because it takes as given that cost forecasts were set accurately and at an efficient level, both in terms of the original costing estimates by Highways England, and the efficiency assumptions that were subsequently applied to these costs. In effect, this 'Delivery' method therefore measures the *sum* of effectiveness/efficiency of delivery *and* the accuracy/reasonableness of the cost forecasts. As such, this method cannot be relied upon alone – it should be used in combination with other benchmarking methods.

Future potential. In the longer term, ORR should seek to develop its understanding of the source of cost overruns: The more that ORR can distinguish between inefficiency and other factors (e.g. scope increase), the more the ORR will be able to use this 'delivery' method to identify scope for additional catch-up efficiencies that can be achieved in the future. Whilst this method can be applied to renewals and/or enhancements, it would likely be most valuable for large one-off projects were it would be possible to dig down into the detail of the source of particular cost overruns – these could act as case studies from which general lessons could be drawn.

It would also be useful to compare Highways England's delivery performance against other companies and/or sectors, to get a sense of what is 'normal' in terms of the level of cost overruns.

Effort/Cost. Only minimal additional analysis would likely be required, given that ORR already monitors Highways England's outturn costs. The main effort/cost involved in this method is to interrogate outturn performance and to understand the driver of differences to forecasts. Based on our experience from other sectors, this process can be a source of disagreement, e.g. a company may claim that a cost overrun is due to scope change and/or external circumstances, which makes it difficult for the regulator to determine the extent to which inefficiency may have been a driver.

6. CONCLUSIONS

In this Section we provide our conclusions considering:

- **Benchmarking options for RIS2:** The benchmarking options that could feasibly be utilised by ORR in time for RIS2, i.e. in time to inform ORR's assessment of Highways England's capex efficiency that will underpin ORR's recommended capex efficiency targets for Highways England for RIS2.
- **Additional benchmarking options for RIS3:** The further benchmarking options that would be worthwhile to develop further so that they can be utilised as part of ORR's efficiency assessment for RIS3.

In our conclusions, we distinguish between top-down and bottom-up benchmarking options. We clarify whether each option applies to renewals and/or enhancements. We also provide some final qualifying remarks.

6.1. Benchmarking options for RIS2

6.1.1. Top-down approaches

The following benchmarking options could feasibly be undertaken for RIS2.

- *Limited international benchmarking for renewals for a small number of comparators*. Given the lack of an existing dataset or an 'international benchmarking club', this option is not feasible on a wide scale in time for RIS2. However, data for some overseas networks could be made available (particularly Sweden, and possibly Norway, although this would be subject to wider discussions), so at least some international comparisons may be feasible in time for RIS2. However, ensuring data consistency would likely be difficult in the short term.
- *Initial national benchmarking for renewals*. Highways England's renewal costs could be compared against LAs. However, as noted earlier, LA costs are currently aggregated across all road types, and it is not clear whether they could be disaggregated by road type in time for RIS2, so costs would not necessarily be directly comparable to Highway England's network. Also, as per international benchmarking options, data collected in the short term may not be consistent.
- *Internal benchmarking for renewals*. ORR is likely to be able to further develop its econometric modelling of maintenance/renewals costs between Highways

England's different regions. If ORR is not satisfied with the robustness of the models developed in the short term, some simplistic unit cost comparisons between Highways England regions could be undertaken as alternative. Such simplistic results would at least provide a starting point for discussion of projected costs.

- *Cost composition analysis for renewals or enhancements*. Engineers' experience of efficient costs could be utilised to assess cost composition for major schemes, assuming that Highways England shares cost data with ORR in a timely fashion. However, we note that this approach is relatively high level/broad brush, so should be considered alongside other methods.
- *Monitoring of Delivery for renewals or enhancements.* ORR already monitors whether proposed efficiencies are delivered, via the CPI metric within the RIS1 Performance Specification. This metric can give a high level sense (or potential even some quantified estimates) of the extent of efficiencies delivered to date, and therefore the scope for additional future catch-up efficiencies.
- *Productivity trend analysis versus other sectors, for total capex*. The change over time in productivity in other sectors can be used as a high level benchmark for Highways England, specifically in relation to a potential 'frontier shift' efficiency target for Highways England. The different potential options – factor productivity analysis and top down unit cost analysis – could both be feasibly calculated and analysed in time for RIS2.

6.1.2. Bottom-up approaches

Bottom-up unit cost analysis for renewals or enhancements. Assuming that Highways England provides data to ORR in a timely fashion, it would be feasible for ORR to benchmark unit costs for aspects of schemes/renewals against detailed cost databases. It is likely that an engineering assessment would be needed alongside this analysis, to consider whether variations in unit costs are driven by differences in (1) efficiency, (2) the scheme specification chosen by Highways England, and/or (3) the scheme standards and regulations determined by ORR and safety bodies.

6.1.3. Risks

All of the methods above have limitations, and therefore the main risk to ORR for RIS2 would be to be over-reliant on any one of these methods. The top down benchmarking options that use road network data (i.e. international, national and internal

benchmarking) can only provide precise comparisons if the underlying data is known to be robust and comparable, and that is not the case for RIS2. The most straightforward options for ORR to implement in time for RIS2 (productivity trends analysis and monitoring of delivery) provide useful context for an efficiency assessment, but they do not compare Highways England with other highway operators so it is difficult to use these to provide a definitive assessment of scope for efficiencies. Cost composition analysis and bottom-up unit cost analysis focus on certain schemes, so whilst they provide a snapshot of efficiency in certain cases, they cannot provide a top-down (company-wide) efficiency benchmark. In addition, bottom-up analysis would require that Highways England makes available to ORR detailed bottom-up cost data, and there is a risk that ORR may not receive the data it requires in time for it to undertake the analysis.

6.1.4. Overall assessment

In light of the discussion above, our high level of assessment of each option is illustrated below, focusing on the key areas of interest from ORR's perspective (feasibility, robustness and effort required to undertake each option for RIS2). A 'Red-Amber-Green' scale is used to highlight the desirability of each option from ORR's perspective. Full explanation on the rationale for our assessment is set out above and in Section 5.

Table B: Summary of assessment of different potential options

Title of option	Feasibility for RIS2	Robustness for RIS2	Effort for RIS2
International benchmarking	Medium	Medium*	High
National benchmarking	Medium	Medium*	High
Internal benchmarking	High	Medium*	Medium
Bottom up unit cost analysis	High	Medium	Medium
Capex cost composition	Unclear**	Medium	Medium
Productivity trend analysis versus other sectors	High	Medium	Low
Delivery	High	Medium	Low

* The robustness of international, national and internal benchmarking may be constrained for RIS2 by limitations in data availability/consistency, but robustness for these methods is likely to be 'high' for RIS3.

** Despite uncertainties, our experience from other sectors suggests this option may be feasible for RIS2.

Each of the potential options has limitations, as shown by all options scoring 'Amber' for robustness. It might be risky for ORR to overly rely on one method for RIS2, so we recommend exploring as many options as possible within the time available to develop a

broad evidence base. The benchmarking results from different methods will provide partial pieces of evidence that will need to be 'triangulated', i.e. assessed in the round.

For RIS2 it would be cost effective to undertake options that are highly feasible and low effort, i.e. productivity trend analysis versus other sectors and delivery (of outturn costs versus forecasts). Other options (involving slightly higher effort) should also be developed to ensure that ORR can assess the level of Highways England's costs – the most feasible being internal benchmarking and bottom up unit cost analysis. Further road sector benchmarking (international/national) could be explored either now or for RIS3.

6.2. Additional benchmarking options for RIS3

6.2.1. Top-down approaches

Top-down benchmarking against other road authorities, most likely for renewals. Benchmarking of Highways England against other road authorities is a long term aspiration for ORR. Other benchmarking exercises (e.g. CQCEN) show that comparisons are likely to be possible, although there will be issues to overcome (e.g. data collection) and there will inevitably be some limitations on comparability. There are pros and cons to using different benchmarks:

- *Internal benchmarking*. This is the most practical option as data consistency is highest, although it does not include any external benchmarks (unless the adjacent highway networks in Wales and/or Scotland are included).
- *National benchmarking*. This offers the greatest sample size, but integrating Highways England might be difficult, particularly given different network characteristics.
- *International benchmarking*. This would introduce benchmarks of the greatest relevance to Highways England (i.e. overseas national road authorities), although developing data consistency and buy-in from other countries would take time and effort.

All of these options – to a greater or lesser extent – will require the collection of robust data that is consistent across organisations and time, as well as collaboration and buy-in from external organisations. It is also worth noting that – as identified by ORR's internal benchmarking and CQCEN regional benchmarking to date – there are considerable issues to overcome (e.g. data consistency), and even after considerable time/effort it is likely that certain limitations will still exist. These challenges are likely to be greatest for

international and national benchmarking, given that Highways England is in a position to encourage its regions to provide data on as consistent a basis as possible.

6.2.2. Bottom-up approaches

Bottom-up unit cost analysis would likely provide the best long term benchmarking option for enhancements, given the considerable future investment in the SRN (making top-down benchmarking of enhancements very difficult). ORR needs to ensure mechanisms are in place to acquire detailed cost data on major schemes.

The potential for bottom-up benchmarking is linked to the regulatory cost assessment process. For example, CAA's capex governance process assesses the efficiency of scheme proposals on an iterative basis (rather than in 5-year cycles), which provides the regulator more opportunity for to interrogate bottom-up costings. ORR might wish to consider how processes have developed in other sectors (e.g. regulation of Heathrow Airport) as a means of expanding the range of benchmarking options that are available.

6.2.3. Risks

It is roundly five years before benchmarking will need to commence for RIS3. This provides considerable time and opportunity for ORR to test the potential for applying different options, to discuss and understand the risks associated with those options, and to consider mitigation strategies. None of the options that we have considered in this report need to be excluded because they present excessive risk to the regulator in terms of reliance. However, as different benchmarking options start to develop over the coming years (e.g. as international datasets become available), we recommend that ORR undertakes an analysis of risk so that it is well-placed to determine the weighting to give each method within its overall efficiency assessment.

6.2.4. Overall assessment

Highways England became a regulated company only relatively recently (in 2015) and the regulatory regime is still developing. Currently, ORR has the opportunity to take a long term perspective, and to put in place now those actions that are necessary to facilitate a greater range of benchmarking options in the future. Given the potential benefits of international and national top-down benchmarking, we recommend that ORR considers making resources available to develop the required datasets and benchmarking clubs.

6.3. Qualifying remarks

The benchmarking options considered in this report assess cost efficiency from different perspectives and therefore are likely to be relatively complimentary. For example, the top-down productivity metrics provide a high level sense check that could be considered alongside any bottom-up unit cost analysis or internal benchmarking.

However, there are other 'non-cost' benchmarking options that could be considered alongside the options highlighted in this report. For example, ORR's 'capability reports' (noted in Section 1.1) consider efficiency from the perspective of processes. These methods are unlikely to be mutually exclusive, because the level of efficiency demonstrated by a company's processes is likely to be exhibited subsequently in the company's level of costs. As such, ideally we would hope to observe broadly consistent conclusions across different types of benchmarking.

When interpreting cost benchmarking results, the 'quality' of a given activity (e.g. renewal) also requires careful consideration and is not something considered explicitly in this report. For example, the quality of a renewal could relate to the safety requirements which determine design standards, the specification (e.g. depth of treatment), the complexity of works involved, the importance/criticality of the asset being addressed, the level and type of traffic carried by the highway, etc. All of these will affect the importance, value, and frequency of renewal requirements, as well as the cost. For all of the benchmarking options discussed in this report, it is important that differences in quality are either taken into account in the quantitative benchmarking analysis (e.g. incorporating quality-related explanatory variables within an econometric modelling approach), or at the very least are considered qualitatively ex post. We recommend that further work is undertaken in this area, to consider how the level of 'quality' of capex-related activities undertaken by Highways England compares to the comparators used in the benchmarking analysis.